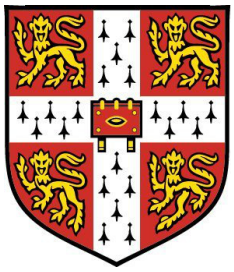
Differential Drug Response as a Function of Age



Rita Moita Santos

Girton College

University of Cambridge

A thesis submitted for the degree of

Doctor of Philosophy

September 2014

To my parents

Declaration of Originality

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

The text in this thesis does not exceed the limit of 60,000 words set by the Biology Degree Committee.

Acknowledgements

If I were to do a differential analysis of my life before and after my PhD, with the same scrutiny which I've employed to study differential drug responses, these would be the main findings: 4 babies born in the group, 1 wedding, 1 funeral, 3 office moves (including an upgrade from the portacabins), 1 new building on campus, 15 seconds with a UK minister, 2 conferences in the states, 3 new countries to add to the existing list of countries in which I've lived because of science, ∼650 free cups of tea, ∼110 working weekends, 13 formal dinners, and 8 "painful" months of writing. And the conclusion, you may ask? Well, the conclusion is, I don't think I would have been able to do it without the support of an amazing group of people, for which I would like to leave a very BIG thank you.

First, and foremost, my supervisor and mentor, John Overington, for all the support and freedom to become a scientist on my own, and for telling me at the interview (after I asked which kind of project he had in mind for me) "I don't quite know yet, you also don't quite know what you want as well, but as long as it helps humans you can do whatever you like". It made my day, and inspired me to be daring.

Second, my thesis advisory committee, Alvis Brazma, Anthony Davenport, Cornelius Gross and Helen Parkinson, for all the critical and constructive comments throughout my project. A special thanks to Cornelius for hosting me in Monterotondo, for the experimental part of my project, and his group - Bianca, Camilla, Elena, Francesca, Lukas, Noellia, Tamara and Yang - for all their patience in bearing with a bioinformatician in the lab.

Third, my collaborators and research advisers: Nenad Sestan, for providing me access to extra gene expression data; Bissan Al-Lazikani, for all the numerous discussions on drug targets; John Marioni for all the great statistical discussions; and Janet Thornton, who albeit being the busiest person at the EBI, still had time to chat with a PhD student about the lack of phenotypic information on gene expression data.

Naturally, though, one does not do a PhD isolated, and for that I'm very grateful to everyone in the ChEMBL group, for the hilarious and fun moments in the darkest hours (and in the sunniest too), besides allowing me to absorb some of their diverse scientific expertise. To all past and present group members, a gigantic thank you: Anna Gaulton, Anne Hersey, Ben Stauch, Felix Krueger, Francis Atkinson, Furkan Yegin, Georgios Papadatos, Gerard van Westen, Grace Mugumbate, Ines Smit, Jon Chambers, Jose Bach-Hardie, Kathryn Holmes, Kazuyoshi Ikeda, Louisa Bellis, Mark Davies, Michal Nowotka, Nathan Dedman, Patrícia Bento, Ruth Akhtar, Samuel Croset, Shaun McGlinchey, Yvonne Light, and all the trainees and interns which I had the pleasure to interact with - Raghd, Taj, Michael, Pele, David, Rodrigo, Ramesh, Simone, Anna, Abhik and Katie. With a very special thanks to Anna, Anne and Patrícia for all the laughs and tears while annotating the mechanism of action of 1600 drugs (still waiting to open the champagne bottle though).

But a PhD is not just about work, it is also about waking up with your housemates making a snow man, when the roads are too dangerous to drive to campus, right Leonor Quintais and Sergio Martinez Cuesta? It is about going to the cinema with your friend, seating up in the opposite side of the room, and leaving the theatre daydreaming about the same scene, right Ana Rita Gomes? It is also about going for drinks to forget frustrating meetings about science and sports, right Stephan Beisken? And, of course, let's not forget board game evenings, museums and musicals in London, and delicious food, right Afonso Guerra, Chen Li, Luís Figueiredo, Matthias Ziehm, Nidhi Tyagi, and Ursula Neu?

However, a journey to a PhD does not start at day zero, and therefore I own a big thank you to my professors in Lisbon for telling me to go abroad, and to my colleagues and friends, Carla Baltazar, Sara Porfírio and Michael Pereira, for dissipating my fears of leaving 'everything' behind, and always being there whenever I missed home (or lost my scientific confidence). Also, a big thank you to my master's supervisor, Chris Oostenbrink, for sharing his network with me, which allow me to have my first contact with industry, and spend six amazing months in the south of Europe again (sea, sun and science, what else can a person want?), and for being above all a friend, and a PhD instigator. Jozef Hritz, no words to thank you properly for all the patience and the hours which you've spend teaching me

while I took my first steps in the computational research world (let's 'never' talk about water molecules and docking though). Anita de Ruiter and Ruben Vosmeer, for all the great afternoons drinking tea in the lab, numerous cooking evenings, and for familiarising me with the Dutch culture. Eszter Császár for all the great long conversations about everything and nothing while climbing the Eiffel tower, and while relaxing at the Szecheny Spa. Susana Direito and Inês Narciso, I believe that "our portrait from 1850 in Netherlands" says it all. A big thank you for all the great trips and moments in Holland and ever since (anyone up for a "Bife" in Trindade?). Alice Glaettli, Sebastian Röhrer and Swetlana Derkesen, thank you for being there while I took my first steps in the big and 'scary' corporate world. Bojan Źagrovic, Mario Hlevnjak, Anton Polyansky, Ruban Zubac, Drazen Petrov, Antonija Kuzmanić, Daniela Kruschel, Pawel Janowski, and Maria Janowska for all the great moments in Croatia, specially Antonija and Daniela for all the cocktails and cappuccinos which took me to realise that I'm an applied science person.

Last, but not least, to my three "little sisters", Ana Isilda, Irís Bidarra and Liliana Martins, thank you for never giving up on me during all the years that I've been away, and for always finding time in your calendars every time I'm in Lisbon. Finally, to my family and to my parents for always being there for me, and never ever saying "No" to my many scientific endeavours: *Obrigada*.

Abstract

Medicines are generally developed to be effective and safe in adult populations, and only a small subset of drugs has been formally studied for efficacy and safety in paediatric or geriatric populations. As a consequence, 45% of prescribed drugs in paediatric hospitals are given off-label, and approximately 20% of drugs given to the geriatric population fall into the same category. However, it is not easy to extrapolate doses, therapeutic indications, safety and efficacy, from one age group to another. Physical size, weight, lipid fraction, metabolic state, drug target and metabolic machinery expression, *etc.*, are all factors that play a role in drug absorption, distribution, elimination, and ultimately efficacy, and these change across the typical human life-span. However, the lack of understanding on how some of these factors change, especially at a molecular level, has been a hurdle in the safe translation of medicines from one population to another, giving rise to a few documented cases of differential drug responses: increased number of side effects, or lack of efficacy. The aim of my project was, therefore, to identify which pharmacologically relevant genes change their expression with age in human populations, in five primary organs (liver, brain, kidney, lungs and heart), which could increase the understanding of some of the age-related differential drug responses reported in the literature. In parallel, I also investigated on how these changes translated to animal models (macaque, rat and mouse) commonly used in drug development.

Overall, a small fraction of pharmacologically relevant genes changed their expression with age in the five primary human organs analysed. However, it was generally observed, that for drug targets associated with reported age-related differential drug responses, when given to either paediatric or geriatric populations, there was an overall tendency for those targets to decrease their expression with age. This pattern was particular striking, among ion-channels, such as GABA A receptor, and interestingly this pattern was not seen in rodents, where the opposite trend was detected. This observation seemed to suggest that one of the reasons why a drug may lack efficacy in children, for instance, may be because the dose

given was not high enough to occupy all the expressed receptors, and thus fully activate the necessary downstream signals for an effective therapy. Likewise, this trend also seemed to suggest, that for geriatric populations, a smaller dose may be required to prevent over inhibition or over activation of these drug targets.

In conclusion, this systematic analysis on the ontogeny of pharmacologically relevant genes brings a new perspective on age-related differential drug responses, and challenges some of the approaches used when extrapolating knowledge from adult studies to other aged populations.

Contents

1 Introduction **1**

1.1 Age-related differential drug responses . . . 2
1.2 Medicines for children . . . 4
1.2.1 Dose extrapolation . . . 5
1.3 Medicines for the elderly . . . 7
1.3.1 Dose extrapolation . . . 8
1.4 Understanding age-related differential drug responses . . . 9
1.4.1 Physiological changes . . . 9
1.4.2 Molecular changes . . . 13
1.5 Aim of the project . . . 21

2 Compilation of age-related Differential Drug Responses **25**

2.1 Introduction . . . 25
2.1.1 Conventional literature searches . . . 25
2.1.2 Refining conventional literature searches . . . 26
2.2 Implementation . . . 27
2.2.1 Machine learning application for literature search optimisation 27
2.3 Results and discussion . . . 31
2.3.1 Overview of age-related differential drug responses . . . 31
2.3.1.1 In paediatric populations . . . 34
2.3.1.2 In geriatric populations . . . 38
2.3.2 Limitations of the dataset . . . 41
2.4 Summary . . . 45
2.5 Material and Methods . . . 46
2.5.1 Machine learning application for literature search optimisation 46

2.5.2 Validation of the NB classifier . . . 47
2.5.3 Chemical properties, approval year and ATC code of drugs in the DDR database . . . 48

3 Compilation of Drug targets, Metabolic enzymes and Transporters 49
3.1 Introduction . . . 50
3.1.1 Drug targets compendiums . . . 50
3.1.1.1 Historical perspective and existing resources . . . 50
3.1.1.2 Issues with existing resources . . . 51
3.1.2 Metabolic enzymes and transporters compendiums . . . 52
3.1.2.1 Historical perspective and existing resources . . . 52
3.1.2.2 Issues with existing resources . . . 52
3.2 Implementation . . . 53
3.2.1 A new "drug to drug target" map . . . 53
3.2.1.1 Problematic assignments . . . 55
3.2.1.2 Comparison with existing resources . . . 58
3.2.2 A new list of Metabolic and Transporters genes . . . 61
3.2.3 Linking drugs to the metabolic and the transporter system . 63
3.3 Results and discussion . . . 66
3.3.1 Distribution of drug target classes in the context of age-related differential drug responses . . . 66
3.3.1.1 In paediatric populations . . . 66
3.3.1.2 In geriatric populations . . . 69
3.3.2 Distribution of metabolic enzymes and transporters classes in the context of age-related differential drug responses . . . 72
3.3.2.1 In pediatric populations . . . 73
3.3.2.2 In geriatric populations . . . 76
3.4 Summary . . . 78
3.5 Material and Methods . . . 79
3.5.1 Drug list and target assignments . . . 79
3.5.2 Grouping of efficacy targets . . . 81
3.5.3 Comparison of drug-target assigments . . . 84
3.5.4 Extracting data from SuperCyp . . . 85

4 Analysis of gene expression datasets as a function of age 87
4.1 Introduction . . . 88
4.1.1 Data requirements . . . 89
4.2 Implementation . . . 90
4.2.1 Datasets from the public domain . . . 90
4.2.1.1 Pitfalls . . . 93
4.2.2 Dataset generated in-house . . . 96
4.2.2.1 Experimental design . . . 97
4.3 Results and discussion . . . 99
4.3.1 Analysis of human expression data . . . 100
4.3.2 Analysis of macaque expression data . . . 117
4.3.3 Analysis of rat expression data . . . 117
4.3.4 Analysis of mouse expression data . . . 120
4.4 Summary . . . 131
4.5 Material and Methods . . . 132
4.5.1 Datasets from the public domain . . . 132
4.5.2 Experimental methods . . . 132
4.5.3 Ethics statement . . . 132
4.5.3.1 Mice . . . 132
4.5.3.2 Tissue extraction . . . 133
4.5.3.3 RNA isolation . . . 133
4.5.3.4 Sample hybridisation . . . 133
4.5.4 Microarray data analysis . . . 134
4.5.5 NGS data analysis . . . 135
4.5.6 Linking drugs to gene expression data . . . 136

5 General discussion 137
5.1 Overview of gene expression changes in the context of age-related differential drug responses . . . 138
5.1.1 In paediatric populations . . . 138
5.1.2 In geriatric populations . . . 144
5.1.3 Clinical implications . . . 149
5.1.4 Pitfalls . . . 151

5.2 Looking beyond age 154
5.2.1 Gender-related differential drug responses 154
5.2.2 Race-related differential drug responses 155
5.3 Looking beyond genomics 156
5.3.1 Proteomics 156
5.3.2 Metabolomics 157

6 Conclusion and future plans **159**

Appendix A **167**

Appendix B **197**

Appendix C **253**

List of abbreviations **309**

References **311**

Chapter 1

Introduction

George Bernard Shaw once said that "If you have an apple and I have an apple and we exchange these apples then you and I will still each have one apple. But if you have an idea and I have an idea and we exchange these ideas, then each of us will have two ideas." [1]. The scientific enrichment gained by employing this simple principle opens the door to a world of possibilities. This is how, after attending a simple talk at the FIGON Dutch Medicine Days conference in 2008, by a pharmacist working in a children's hospital ward in Leiden, Netherlands, I got introduced to the issues surrounding medicines for children. In particular, the opening line from her talk, directed to some delegates from the pharmaceutical industry in the audience, will forever remain in my mind: "I'm going to let you know what we do with your precious medicines, in order to actually be able to administered them to our neonatal patients: we crush them, we mix them with solvents, and we have no idea about the impact of that procedure in the final therapeutic response."[1]. Naturally, when the time came to pick a research topic for my PhD project, this vivid memory swarmed back into my mind, filling it with questions: a) why is the pharmaceutical industry not formally testing medicines in children, in order to provide safe and effective dosing schemes for practitioners?; b) is it just because there is more risks involved, and thus a higher liability, besides ethical and technical constraints, to do clinical trials in children?; c) what are the regulatory bodies doing to address this gap?; d) will a medicine developed for an adult work in the same way in a child?; e) or are the disease mechanisms

[1] Quoting as accurately as I remember it.

in "kids" different than the ones in adults?; f) how does the maturation process affect absorption, distribution and metabolism of drugs, is it a linear process?; g) and finally, are there any differences at the molecular level between a child and an adult which could impact drug response, and help understand some of the behaviours seen in the clinic?. Despite being triggered by the issues about developing medicines for children, I rapidly expanded my curiosity to the other side of the spectrum as well, the elderly: a) how does ageing affect drug exposure, and response?; b) why does a medicine stop being effective once the patient reaches a certain age, is it due to a physiology difference or is it due to differences at the molecular level?. Being a biochemist by training, I quickly decided to focus on what one could learn by exploring the differences at molecular level between differently aged populations in the context of drug response. In the end, what started as a daring idea, became the fuel for the pages to follow, but as in the words of Johann Wolfgang von Goethe "Daring ideas are like chessmen moved forward. They may be beaten, but they may start a winning game." [2].

1.1 Age-related differential drug responses

As mentioned in the abstract of this thesis, medicines are generally developed and tested for safety and efficacy in adult populations, and only a minority of those, are formally tested in paediatric and geriatric populations. Ethical, technical, and logistical challenges have been the main drivers for this bias. As a consequence, 45% of current drugs prescribed in paediatric hospitals are still used off-label, without adequate understanding of appropriate dose, pharmacokinetic profiles, efficacy, and possible unseen side effects [3]. Similarly, approximately 20% of drugs given to the elderly in geriatric wards fall into the same category [4]. Pharmacists have then to extrapolate the efficacy and safety guidelines from the studies conducted in the adult populations. However, it is not a straightforward process. Size, weight, fat fraction, metabolic state, drug target availability, *etc.*, change with age, and they all play a role in the drug's ultimate safety and efficacy profile. The lack of understanding of how some of these factors change, especially at a molecular level, has been an hindrance when trying to transfer one medicine from one population to another. As a result, the practitioner is sometimes confronted with what we

call age-related differential drug responses, i.e., the drug shows more side-effects or unseen side-effects when transferred from one aged population to another, or lack of efficacy. For example, chloramphenicol, a broad-spectrum antibiotic, when administered to premature and newborn infants with high doses, causes the "Grey baby" syndrome[1] [5; 6], while in adults, chloramphenicol toxicity is responsible for aplastic anaemia[2] instead. On the other hand, paracetamol when given to children, despite being slowly metabolised, the incidence of reported cases with hepatoxicity is lower than in adults [7]. Another interesting drug class with known differential responses are the selective serotonin re-uptake inhibitors (SSRIs), a class of antidepressant drugs, such as paroxetine, fluoxetine, citalopram, *etc.*. Overall when children are exposed to SSRIs the incidence of adverse drug reactions is higher than in adults, and the outcome is not so effective [8; 9]. Aspirin, is another great example of differential drug response, it has long been tentatively associated with the incidence of Reye's syndrome[3] in children [10; 11], yet it is relatively safe when administered in adults. In case of age-related differential drug responses in geriatric populations, some of the most known cases are: the increase in sensitivity for benzodiazepines [12], and the fact that β-adrenergic receptor antagonists have a decrease in efficacy in this aged population [13].

There have been a few attempts to understand these reported differential drug responses, mainly from studies on physiology differences, but also, a few from the molecular level point of view. For instance, the hepatic glucuronidation machinery, which is estimated to have lower capacity in infants, might explain the "grey baby" syndrome caused by chloramphenicol in toddlers [6]. Additionally, immature oxidative pathways, in the case of paracetamol, might explain the reduced incidence of hepatoxicity seen [14], and the lower expression levels of the beta-adrenergic receptor in the human heart in the elderly, seems to be the cause for the lack of efficacy of beta-adrenergic inhibitors when prescribed to lower arterial pressure [13]. However, the lack of a central resource with information on drugs with reported differential drug responses, has probably prevented a thorough analysis on the

[1] This syndrome is characterised by cardiovascular collapse, cyanosis, hypothermia, hypotension, grey skin colour and vomiting after administration of high doses of chloramphenicol.

[2] Aplastic anaemia is a condition where bone marrow does not produce sufficient new cells to replenish blood cells.

[3] Reye's syndrome is a potentially fatal disease that causes numerous detrimental effects to many organs, especially the brain and liver.

differences between differently aged populations, specially at the molecular level.

In the next sections, a summary of the constraints on clinical trials in paediatric and geriatric populations will be given, as well as, a review of the latest efforts from the regulatory bodies to foster formal testing of medicines in these two populations. Also, an overview of what is currently known in terms of physiological differences and molecular differences between differently aged populations, relevant for drug response, will be presented.

1.2 Medicines for children

As mentioned in the previous section, ethical, technical, and logistical challenges have been the main barriers to develop and formally test medicines for children. These include for example, the debatable risk/benefit ratio from exposing children to the test of drugs for non-lethal diseases, the lack of technology to assay very small amounts of blood [15], and the necessity of special facilities to guarantee the well-being of the children and their parents while performing clinical trials. In addition, the lack of available age-appropriated formulations, and the absence of an established list of safe excipients to be used in children [16], hinders the development of new medicines. Likewise, the lack of incentives from the regulatory agencies, in Europe the European Medicine Agency (EMA), and in the USA the Food and Drug Administration (FDA), to foster the interest of the pharmaceutical industry to conduct clinical studies in children, in order to overcome the barriers due to high developmental costs, and a small and fragmented market, have also contributed to the low number of medicines formally tested in children [16].

In the last decade, however, both regulatory agencies have improved their policies, and created several infrastructures to foster a profitable work-environment. Thus entice pharmaceutical companies to invest in this area, and adapt medicines according to paediatric needs [3]. In Europe, the first legal document, entitled "Better medicines for children - proposed regulatory actions in paediatric medicinal products" was created in 2002, and brought some guidelines on how to perform clinical trials in children within ethical considerations [17]. After several revisions, inputs and impact assessment, the regulations were finally translated into law in 2007 [18]. Paediatric initiatives in the USA started earlier than in Europe. In

1997, the FDA offered a financial incentive to pharmaceutical companies: if the companies would conduct paediatric studies, an additional 6 months of market exclusivity would be granted. Later, the "Best Pharmaceuticals for Children Act" (BPCA) was created in 2002. The aim of the BPCA is to regulate the labelling of paediatric usage and enhance the paediatric drug development and testing of off-patent drugs. It was updated in 2007 to include additional patent exclusivity for on-patent drugs being tested for paediatric use [19]. However, some of the incentives created by the regulatory bodies might not be as profitable as envisioned. For example, the economic return on clinical trials performed under the paediatric exclusivity program in the USA was variable, found to be between -$8.9 million to $507.9 million, which may prevent other companies from embarking on the same route [3]. Nevertheless, both the BPCA and the paediatric regulation created in the European Union, have nourished an increasing number of paediatric clinical trials, which may aid in the future to decrease the number of drugs used off-label in this population [16]. Furthermore, both regulators, in a joint effort, have created a database of safety and toxicity excipients for paediatrics (STEP - Safety and Toxicity of Excipients for Pediatrics), to stimulate and improve paediatric drug formulations [20; 21]. Lastly, EMA and FDA have also agreed to share information on scientific and ethical issues related to the development of medicines for children, to foster global development plans.

1.2.1 Dose extrapolation

In the absence of guidelines, however, and until new policies have a direct impact in clinical settings, the medical practitioner and the pharmacist are left to extrapolate doses from adult studies to paediatric patients. In the past, simple rules were used to estimate the right dose based on the age or the weight of the child, such as Fried's rule (to estimate dosage for infants $\leq$ 2 years old, eq. 1.1), Young's rule (eq. 1.2) and Clark's rule (eq. 1.3), to name only a few [22; 23; 24].

$$Infant's\ dose = \frac{age\ of\ child\ in\ months}{150\ lb} \times adult\ dose \tag{1.1}$$

$$Child's\ dose = \frac{age\ of\ child\ in\ years}{age\ of\ child\ in\ years + 12} \times adult\ dose \tag{1.2}$$

$$Child's\ dose = \frac{weight\ of\ child\ kg}{70\ kg} \times adult\ dose \tag{1.3}$$

Nowadays, these rules solely based on age and/or weight are not so commonly applied [23], given their empiricism, and accumulating evidence reporting on their lack of accuracy when extrapolating the dose of certain drugs. For example, Penna's rule tend to overestimate the dose of paracetamol for a child between 1 and 5 years old, exposing the child to a higher risk of hepatoxicity [25]. Even rules which use body surface area (BSA) (eq. 1.4) are no longer routinely employed for the same reason, despite being preferable over the previous set. The current best practice is therefore, to consult paediatric formularies, such as the british national formulary for children (BNFC) [26], which summarise the best clinical practices based on paediatric expert recommendations and years of clinical experience.

$$Child's\ dose = \frac{adult\ dose}{1.7m^2} \times Child's\ BSA\ m^2 \tag{1.4}$$

Alternatively, pharmacokinetic and pharmacodynamic models (PK/PD) developed based on population studies with children, and consequently, more accurate than the previous empirical rules, can be used to develop rational dosing schemes with a predictable efficacy and safety profile for the individual child of varying age [27; 28]. These mathematical models describe the influence of developmental changes on drug disposition, efficacy and safety. Specifically, these models consist of a series of equations which describe the compartments of the human body, parametrised based on *in vitro* and *in vivo* data collected from the literature, such as, concentration of metabolic enzymes in the liver and in the gut, their kinetic parameters, concentration of transporters, *etc.* [29].

Lastly, allometric scaling based on the simple allometric relationship between body weight and pharmacokinetic parameters (eq. 1.5, where Y is clearance or volume of distribution, BW is the body weight, and a and b are the coefficients and the exponent of the allometric equation), can also be used to extrapolate an adult dose to a child dose, however, in practice it is mainly used to estimate the

first-children dose for clinical trials [30; 31].

$$Y = a.BW^{b} \tag{1.5}$$

1.3 Medicines for the elderly

The circumstances and barriers which contribute to lower the number of formally tested medicines in the elderly, are however, quite different to the ones enumerated above for a lack of formally tested medicines in children. In short, the presence of medical comorbidities, the uptake of on average two to nine different drugs per day [32], the increase in susceptibility for adverse drug reactions (ADRs), the heterogeneity of the geriatric population, i.e. elderly *vs.* very elderly, all contribute to make this age group unattractive for clinical trials. For example, the PREDICT study, which recently analysed ongoing clinical trials for heart failure, concluded that 25% of those have an upper age limit, and approximately 80% exclude patients with comorbidities [33]. In addition, it is hard to recruit patients older than 65 years old, specially above 75 years of age. Low mobility and sometimes lack of transportation, disruptive trial schedules unadapted to the patient's daily routine, and their caregivers, as well as, overly protective physicians, and sometimes inappropriate facilities to accommodate the very elderly, all contribute to lower the number of patients enrolled from this age strata [33].

Historically, regulatory agencies started to promote the inclusion of older subjects in clinical trials in the late 1980s [34], and in the last two decades several initiatives led to development of guidelines on how to foster the study of drugs likely to be used in older adults. For example, in 1989, the FDA, created the first guide on how to perform clinical trials in the elderly, and later in 1998 it approved the geriatric labelling rule, which guaranteed that safety and effectiveness data of a new drug in relation to age and the appropriate dosage requirements was reported in every prescribing information [34]. In Europe, the first document concerning the inclusion of the elderly in clinical trials followed from the International Conference of Harmonisation (ICH E7) in 1994 [35]. In 2006, the adequate representation of elderly patients in clinical development programmes was again raised by the European Commission (EC), following a letter drafted by the Euro-

pean Union Geriatric Medicine Societies (EUGMS), which resulted in the creation of the EMA geriatrics medicine strategy in 2011 [34; 36]. This strategy aims to improve the access to the best care to the ageing population in Europe, by liaising with the industry, geriatric expert committees, and patients groups, in order to address unmet needs of this population. For example, this initiative encourages future clinical trials to enrol more elder subjects, and monitor these patients for geriatric-specific adverse events, such as cognitive decline, balance, urinary incontinence, among others. In terms of pharmacovigilance, the new legislation also offers an opportunity to strengthen the system for monitoring the safety and benefit-risk balance of medicines post-authorisation, by optimising data collection and by using information from spontaneous ADR reports to fill existing knowledge gaps [33].

Overall, both regulatory agencies are working together to increase the number of elderly subjects enrolled in clinical trials, and insure that the trial reflects as closely as possible the population of patients for which the drug will be largely prescribed to.

1.3.1 Dose extrapolation

Meanwhile, clinicians are confronted with the task of dosing multiple medicines to the elder patient, without appropriate information on possible drug interactions, nor if the dose used in adults is safe and effective in this aged population, or if a dose adjustment is in fact required. Historically, for drugs eliminated through the kidneys, a decision to reduce the dose was made, for instance, based on the levels of creatinine clearance. In short, creatinine clearance was estimated using the Cockcroft formula (eq. 1.6 for men and eq. 1.7 for women), and if creatinine clearance was half of the values seen in healthy adults, then the dose should be reduced to half, assuming that there were no alterations in bioavailability and plasma concentration of the drug with age (eq. 1.8, where C_p is drug plasma concentration, CL is the apparent drug clearance and F is bioavailability) [37].

$$creatinine\ clearance\ (male) = \frac{1.23 \times (140 - age) \times body\ weight\ (kg)}{plasma\ creatinine\ (\mu/mol)} \tag{1.6}$$

$$creatinine\ clearance\ (female) = \frac{1.04 \times (140 - age) \times body\ weight\ (kg)}{plasma\ creatinine\ (\mu/mol)} \tag{1.7}$$

$$dose = Cp \times \frac{CL}{F} \tag{1.8}$$

Nowadays, pharmacokinetic and pharmacodynamic models (PK/PD) are also used to estimate the right dose for the elderly, being particularly useful when estimating the risk of concomitant treatments. For example, a recent model used to predict the risk of a drug-drug interaction (DDI) between codeine and a metabolic inhibitor of codeine, revealed that for a patient with normal renal function, the dual treatment would be rather safe, however, for patients with moderate/severe renal impairment, the model indicated a high risk of respiratory depression [33].

Lastly, similarly to what exists for the paediatric populations, a list of drugs with known adverse reactions in the elderly or which require a dose reduction has been compiled by geriatric expert committees based on years of clinical experience. In the USA, the most employed one is the Beers criteria [38], and in Europe the STOPP (Screening Tool of Older Persons Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment) protocols are the ones mostly in use [39; 40].

1.4 Understanding age-related differential drug responses

1.4.1 Physiological changes

As mentioned in section 1.1, there has been some attempts to understand some of the age-related differential drug responses reported in the literature. Most studies focused on physiological differences between differently aged populations, which could affect the absorption, distribution, metabolism, and clearance of drugs. Table 1.1 summarises the main findings from those studies.

Table 1.1: Known age-related physiology changes which affect drug disposition.

Variable influenced	**Physiological parameter**	**Changes in paediatrics**	**Changes in geriatrics**
Absorption	gastric pH	**Obs.:** Nearly neutral during neonatal period; Reaches adult levels at 2 years of age [41; 42]; **Effect:** Medications with weakly acidic compositions require higher doses to compensate for lower absorption (e.g. phenytoin, phenobarbital), and pH-sensitive drugs may require a lower dose (e.g. penicillin, erythromycin, and amoxicillin)	**Obs.:** Reduction in acid secretion, i.e. increase in gastric pH [19]; **Effect:** Possibly delay in absorption
	gastric emptying	**Obs.:** Delayed emptying, fully maturation only at 6 to 8 months old [42]; **Effect:** Delays the time a drug takes to get into the blood and exert its therapeutic action	**Obs.:** Decrease gastric emptying [19]; **Effect:** Possibly delay in absorption
	gastrointestinal motility	-	**Obs.:** Decreased [19]; **Effect:** Possibly delay in absorption

Continued on next page

Table 1.1 – *Continued from previous page*

Variable influenced	**Physiological parameter**	**Changes in paediatrics**	**Changes in geriatrics**
Distribution	Body water: body fat	**Obs.:** Increased in neonates and infants [41]; **Effect:** Increase volume of distribution for hydrophilic drugs (e.g. phenobarbital), thus higher dose required to achieve the same clinical effect, and intramuscular and subcutaneous administration tends to result in unpredictable plasma drug levels	**Obs.:** Decrease in the elderly, i.e. adipose tissue increases from about 18% to 36% in men and from 33% to 45% in women [19]; **Effect:** higher distribution of hydrophobic drugs (e.g. diazepam)
	Blood-brain barrier (BBB)	**Obs.:** Developed at 4 months old [43]; **Effect:** BBB more permeable, i.e. children are more susceptible to centrally acting drugs	-
Metabolism & Elimination	Hepatic blood flow	**Obs.:** Higher in children [44]; **Effect:** Increase clearance of drugs with high hepatic extraction ratios	**Obs.:** Hepatic blood flow is 35% lower in healthy people over 65 years of age [19]; **Effect:** Reduced clearance for drugs with high hepatic extraction ratios (e.g. propanolol, clomethiazole)

Continued on next page

Table 1.1 – *Continued from previous page*

Variable influenced	Physiological parameter	Changes in paediatrics	Changes in geriatrics
	Renal clearance	**Obs.:** Glomerular filtration in infants is only 25%-30% of the adult values, increasing to 50%-75% at 6 months, and reaching maturity at 2 or 3 years of age [41]; **Effect:** Drugs eliminated by renal excretion tend to accumulate (e.g. aminoglycosides, penicillins)	**Obs.:** May decrease with age [45]; **Effect:** Reduced clearance for drugs mainly excreted by the kidneys
	Kidney size	-	**Obs.:** Decreases with ageing [45]; **Effect:** Reduced clearance for drugs mainly excreted by the kidneys

Briefly, maturation processes in children are not linear, consequently, the required dose in certain age subsets may be lower, identical to, or higher than in adults, depending on the drugs metabolic pathway and the drug's physical properties. In addition, up to 12 to 18 months of age, the organs are still being developed, and therefore, the child pharmacokinetic parameters deviate significantly from the ones seen in adults. Although, in older children and teenagers, those parameters approach adult values and are thus easier to predict [16].

In geriatric populations, the system which suffers major alterations with ageing is the urinary system [46; 47; 48], with renal clearance being the factor which most dictates dose adjustments in this population. Even the alterations observed in parameters which may influence absorption, such as decrease in gastric pH, gastric emptying and gastrointestinal motility do not seem to delay drug absorption to the point where a dose adjustment is generally required [19]. The increase of adipose tissue, and consequently the decrease in body water to body fat ratio, is also a factor which plays a very important role in drug safety in this aged population. For example, given that lipid-soluble drugs, such as diazepam, are more widely distributed in the elderly, this population is thus exposed to a prolonged action and a "hangover" effect because of the longer elimination half-life [19].

1.4.2 Molecular changes

From the molecular point of view, however, not as many studies have been conducted, in order to understand how the ontogeny of pharmacologically relevant genes, or their downstream signals, may influence drug response [42; 49], and thus, potentially explain some of the age-related differential drug responses which have been reported in the literature. Table 1.2 summarises the most known molecular differences between differently aged populations, which could alter drug safety and efficacy.

Table 1.2: Known age-related molecular changes which affect drug disposition and response.

Variable influenced	**Molecular parameter**	**Changes in paediatrics**	**Changes in geriatrics**
Distribution	P-glycoprotein	**Obs.:** Expression levels in central nervous system lower than in adults; no information on developmental expression in the human intestine [42];	-
	Albumin	**Obs.:** Production of albumin is decreased in the first year of life [41]; **Effect:** Medicines highly protein bound (e.g. phenytoin and valproic acid) will have a higher unbound/total plasma concentrations than adults	**Obs.:** Plasma albumin decreases [50]; **Effect:** Increase free concentration of acidic drugs (e.g. penicillin, amiodarone, warfarin, digoxin)
	α-1-acid-glycoprotein	**Obs.:** Production of α-1-acid-glycoprotein is decreased in the first year of life [41]; **Effect:** Medicines highly protein bound (e.g. phenytoin and valproic acid) will have a higher unbound/total plasma concentrations than adults	**Obs.:** Plasma α-1-acid-glycoprotein increases [50]; **Effcet:** Decrease free concentration of alkaline drugs (e.g. macrolide antibiotics)

Continued on next page

Table 1.2 – *Continued from previous page*

Variable influenced	**Molecular parameter**	**Changes in paediatrics**	**Changes in geriatrics**
Metabolism	ADH1A	**Obs.:** Expression in fetal liver detected at 16 weeks; Predominant ADH1 enzyme in fetal liver, neglectabe expression values in adult liver [51];	-
	ADH1B	**Obs.:** Expression in fetal liver detected at 16 weeks; Predominant ADH1 enzyme in infant liver, comparable expression values in adult liver with ADH1C [51];	-
	ADH1C	**Obs.:** Expression in liver detected at birth; Comparable expression values in adult liver with ADH1B [51];	-
	CYP1A2	**Obs.:** Activity detected at 1 to 3 months of life; Infants older than 4 months old have activity levels equivalent to adults, and infants older than 6 months have activities higher than adult levels; Activity levels approach adult levels again during adolescence, faster in girls than in boys [42] ;	-

Continued on next page

Table 1.2 – *Continued from previous page*

Variable influenced	Molecular parameter	Changes in paediatrics	Changes in geriatrics
	CYP2C9	**Obs.:** Activity detected in the first week of life; Activity higher in prepubertal children than in adults [42]; **Effect:** Warfarin clearance is higher in prepubertal children	-
	CYP2C19	**Obs.:** Activity detected in the first week of life; Similar activity levels in children and in adults [41; 42] ;	-
	CYP2D6	**Obs.:** Activity detected within days of birth; Protein levels reaches 70% of adult levels at 4 weeks to 5 years of age; mRNA levels in adults correlate with protein levels, but not in infants [42; 51] ;	-
	CYP2E1	**Obs.:** Activity detected within hours of birth; Protein levels reach adult levels at 1 year of age [42; 51] ;	-

Continued on next page

Table 1.2 – *Continued from previous page*

Variable influenced	Molecular parameter	Changes in paediatrics	Changes in geriatrics
	CYP3A4	**Obs.:** Activity detected in the first week of life, reaches 20% of adult levels at 1 month, and 72% at 1 year; Between 2 and 12 years, activity is higher than in adults, decreasing to adult levels between 12 to 16 years [41; 42]; **Effect:** For carbamazepine, ethosuximide, and midazolam, infants require a lower dose, while children must have a higher weight-based dose compared to adults	-
	CYP3A7	**Obs.:** High activity in fetal liver, decreased protein levels after birth [51];	-
	FMO1	**Obs.:** High expression in fetal liver, decreased protein levels after birth [51];	-
	FMO3	**Obs.:** Low expression in fetal liver, increased protein levels after birth [51];	-
	PON1	**Obs.:** Expression levels increase after birth, and reach adult levels at 10 to 25 months of age [51];	-

Continued on next page

Table 1.2 – *Continued from previous page*

Variable influenced	Molecular parameter	Changes in paediatrics	Changes in geriatrics
	SULT2A1	**Obs.:** Activity detected in fetal liver; reach adult levels at 3 months of age [51];	-
	UGT1A1	**Obs.:** Expression levels reaches adult levels at 3 to 6 months of age [51];	-
	UGT1A6	**Obs.:** Decreased activity in newborns and young children in comparison with adolescents and adults, expression reaches 50% of adult levels at 6 months of age, but does not reach maturation until puberty [42; 51]; **Effect:** Glucuronidation of paracetamol is decreased in newborns and young children	-
	UGT2B7	**Obs.:** Detected activity in premature infants at 24 weeks, increasing considerably in the following weeks; Expression levels reach adult levels at 2 or 3 months of age [42; 51]; **Effect:** Increased doses of morphine needed to maintain effective analgesia	-

Continued on next page

Table 1.2 – *Continued from previous page*

Variable influenced	Molecular parameter	Changes in paediatrics	Changes in geriatrics
Efficacy	β-adrenergic receptors	-	**Obs.:** Decreased expression [45]; **Effect:** Decrease efficacy of β-adrenergic agonists and antagonists

Interestingly, it is clear that among the few studies conducted to explore molecular differences between differently aged populations, a lot of the effort has been in understanding the ontogeny of metabolic enzymes [51]. This effort has been particularly important to support the development of robust pharmacokinetic models that provide an improved means of predicting age-specific xenobiotic disposition.

Despite that, there are apparent contradictions in the literature regarding the ontogeny of some of the this metabolic enzymes. However, this may be explained by the fact that different studies measured different surrogate variables i.e. enzymatic activity, mRNA levels or protein levels, which may not always be linearly correlated. For example, a study conducted on patients between 5 months and 18 years of age, revealed CYP2C9 protein levels to be around 50% of the adult values, however, the metabolic ability of paediatric patients, regarding this cytochrome, is equivalent to adult levels after 1 year of age. This apparent discrepancy is likely due to the developmental changes in the liver size to body-weight ratio[1], which may compensate the lower protein levels [51]. Furthermore, levels of mRNA may be buffered by other cellular regulatory processes, and not always be positively correlated with the final protein levels.

Curiously, in the elderly, no differences in expression of metabolic enzymes have been reported, despite the fact that the hepatic metabolism of drugs eliminated by phase I enzymes is reduced by 30-50% in elderly patients. This observation is, however, attributed to changes in physiology with age and not molecular changes [12].

Regarding other important systems for drug disposition, up to now, it is unknown at what point the production of proteins crucial for renal excretion and reabsorption of molecules is fully matured [41]. In addition, the knowledge available on the development of the immune system is also rather scarce at the molecular level, with seldom information available at the cell-based level [52]. For example, it is known that B-cells in infants are only fully mature, and able to produce various antibodies at around two years of age [52], and in the elderly, this population of cells tend to decrease with age, which seems to explain why elderly patients respond to vaccination less favourably [50].

[1]The liver of a child is bigger than in adults in relative terms (1/18 vs. 1/36 of the entire body-weight) [44].

1.5 Aim of the project

Given the state of the field:

- no overview available on the most frequent age-related differential drug responses, nor the main therapeutic indications of the drugs belonging to that group;
- models and software available for dose extrapolation mainly built on data based on physiological differences between differently aged populations;
- and the lack of knowledge on the ontogeny of pharmacologically relevant genes, in particular, of the drug target machinery.

The obvious questions to address were: can we contribute to improve these dose extrapolation models by re-using gene expression data to understand pharmacodynamic changes of drug targets with age? Can we use that knowledge to minimise future differential drug responses when a drug is transferred from an adult to a child and/or to an elder? For example, if the gene expression of a drug target is higher in children than in adults, can we hypothesise that a higher dose may be required to achieve the minimum receptor occupancy necessary to obtain the desire clinical effect? Likewise, if the expression levels of a drug target are lower in old adults, can we expect that a reduction in the dose may be required to minimise over inhibition or over activation of these drug targets and/or other off-targets effects as illustrated in fig. 1.1?

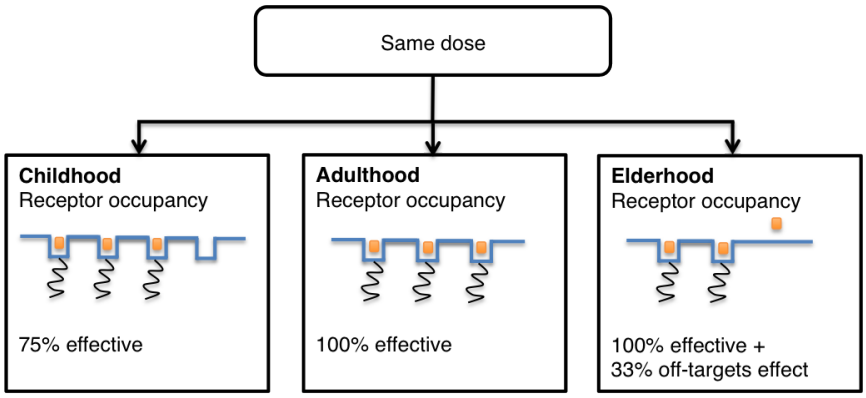


Figure 1.1: Diagram illustrating a working hypothesis of receptor occupancy in children, adults, and in the elderly, for drug targets, whose expression levels may decrease with age.

Based on this, the aim of my project was therefore:

1. to compile a list of drugs which have shown age-related differential drug responses (DDRs) from the literature;

2. to compile a list of pharmacologically relevant genes;

3. to study how those pharmacologically relevant genes responsible for the drug's safety and efficacy may change their expression with age;

4. to investigate if those differences in expression may explain some of the differential drug responses reported in the literature, i.e., higher incidence of side effects or unseen side effects, and/or lack of efficacy, in children and/or in the elderly;

5. and to explore how those differences in expression are translated to common animal models used in drug discovery, in order to identify potential issues when designing pre-clinical studies, and to identify the best age-driven animal model.

In the next chapters, these objectives will be addressed and thoroughly discussed. Briefly, in chapter 2, the process to compile a list of drugs which have

shown differential drug responses, will be presented, as well as, an overview of the main therapeutic indications prone for such events, and how the chemical properties of the drugs in the dataset vary per DDR category (safety *vs.* efficacy *vs.* control), and per aged population (paediatrics *vs.* geriatrics). Similarly, in chapter 3, an account will be given on the process to gather a list of pharmacologically relevant genes, i.e., drug targets, metabolic enzymes and transporters, along with, an analysis on its distribution and incidence per DDR category. In chapter 4, the expression of those genes with age in five primary tissues (brain, heart, lungs, liver and kidney) and in four species (human, macaque, rat and mouse) will be unveiled. Finally, in chapter 5, the story will come together when I attempt to rationalise age-related differential drug responses based on the expression changes detected in chapter 4, and draw the parallels with what was observed in animal models. This dissertation will then culminate with the conclusions and future plans section, where the main findings of this project will be presented, and new ideas and experiments will be suggested and discussed, which could be used to further develop this project.

Chapter 2

Compilation of age-related Differential Drug Responses

As mention in the introduction, to verify if molecular differences between differently aged human populations, could help rationalise why when certain drugs are transferred from adults to children, or from adults to the elderly, do not show the same efficacy or the same safety profile, is necessary to first gather a list of drugs that fall into those categories. Therefore, in this chapter, I describe the work conducted to compile a comprehensive and up to date list of drugs with reported age-related differential drug responses (DDRs). In addition, I present and discuss the main therapeutic indications prone for such events, as well as, how the chemical properties of the drugs in the dataset vary per DDR category (safety *vs.* efficacy *vs.* control), and per aged population (paediatrics *vs.* geriatrics).

2.1 Introduction

2.1.1 Conventional literature searches

The best sources of information to obtain a comprehensive list of drugs with reported DDRs are medical records and publicly accessible medical articles, as currently there is no specialised resource providing this type of information. Both of these have advantages and disadvantages. For example, medical records, would in principle be more accurate in terms of the information provided, as they are direct

annotations of the patients' drug responses. However, access to medical records is restricted to a subset of researchers, with collaborations with local paediatricians and geriatricians, besides the patient's consent, being required in order to gain access to those records. Additionally, as the majority of records are not provided in a standard digital format, querying them would also be quite time-consuming, and it would be more difficult to capture the overall incidence of DDRs at the population level.

Medical articles, on the other hand, are generally summaries of population studies and readably accessible through general medical literature databases such as PubMed/MEDLINE [53], Europe PMC [54], or Embase [55], being thus generally available in standard digital formats. The latter is a commercial database with a higher coverage of topics such as pharmacology, drug research and toxicology, while the first two databases are freely accessible and routinely synchronised. Querying these two last resources is generally enough to gain an overview of a medical or life-sciences related topic, given that when combined they provide access to over 23 million peer-reviewed scientific journal articles.

Keyword search methodologies, can be used to query these databases, and retrieve a list of articles containing pertinent information on the topic selected. These conventional methods can perform satisfactorily for well defined search criteria and topics with a manageable number of publications, however for wider topics, there are some limitations.

2.1.2 Refining conventional literature searches

When the user employs conventional key-word search methods to extract information on a wide topic, the amount of articles retrieved with these methods can be overwhelming. A prior knowledge of the field may be required to manage subsequent filters, like restrict the search criteria to certain publication years, or to certain subtopics. For example, to obtain information on drugs with reported DDRs for this project, I have used conventional key-word search methods to obtain an initial set of relevant articles. However, to reduce the number of articles retrieved to more manageable numbers, I lacked a good initial understanding of all drug classes that have shown age-related differential drug responses, to restrict the search criteria further. Therefore, I decided to combine a key-word search

approach with a text-based machine learning approach, to help prioritise articles that may contain relevant information for the project. This strategy has already proven to be successful in other fields such as toxicogenomics [56; 57].

2.2 Implementation

2.2.1 Machine learning application for literature search optimisation

The workflow for literature search optimisation, to retrieve and prioritise articles with information on drugs with reported DDRs, was adapted from the work done by Papadatos *et al* on a document classifier for medicinal chemistry publications trained on the ChEMBL corpus [58]. As depicted in figure 2.1, the protocol initiated with a series of key-word searches to retrieve articles from Europe PMC containing the word "drug", and age- and drug- related words (Appendix A). Subsequently, abstracts from articles with at least 5 age- and drug- related words were manually checked, and assigned a flag of "relevant", if they seemed to contain information about drugs with reported DDRs, and "non-relevant" if they didn't. This set was used to train and validate a Naïve Bayesian (NB) document classifier. A Naïve Bayesian classifier employs a probabilistic model to assign an unlabelled object into a class based on the simplifying assumption of conditional independence among features. Briefly, the NB classifier computes a set of discriminant functions (eq 2.1, where O is an unclassified vector of a attributes, f(O) is the discriminant function corresponding to the i$^{\text{th}}$ class, and vjk the value of attribute Aj in the object) to determine what is the probability of a certain object to belong to a certain class, given the presence or absence of attributes as seen in the training set, and assigns the object to the class for which the function is maximised.

$$f_i(O) = P(C_i) \prod_{j=1}^{a} P(A_j = v_{jk}|C_i) \tag{2.1}$$

In practice, a modified set of discriminant functions are employed, in order to correct for the fact that sometimes a given class and attribute value never occur

together in the training set, and thus to prevent all other probabilities to be wiped out by being multiplied by zero, a small-sample correction is incorporated into all probabilities, known as the Laplace correction [59]. Despite the fact that the "naive" assumption of independence among features is rarely true, the performance of the NB has been empirically found to be comparable to other more sophisticated machine learning methods, including many where clear attribute dependencies exist, and across a number of diverse domains, including text-mining applications [59]. The NB model thus scored the likelihood of an article to contain information about drugs that are not as effective or safe in paediatric or geriatric populations when compared with adult populations. This approach identified and prioritised approximately 1,400 articles out of a pool of 19,200. Articles freely available in PubMed central were selected for evaluation and data extraction. A more technical and detailed description of the protocol can be found in the Material and Methods section.

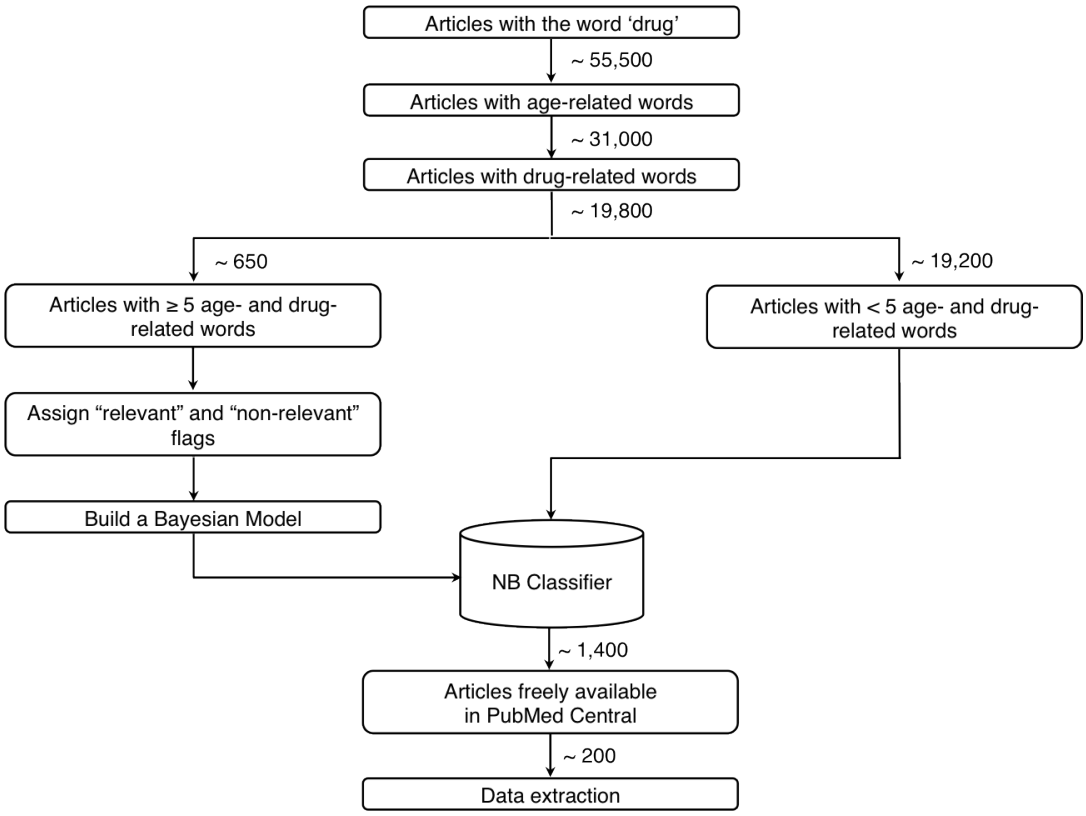


Figure 2.1: Document processing and classification workflow. NB - Naïve Bayesian.

From the 178 freely available articles in PubMed central, 21 contained relevant information for this project, resulting in the identification of 57 drugs with reported DDRs. Articles not freely available in PubMed central, but selected by the classifier were also manually checked if the score was higher than 4, and access was granted through the University of Cambridge library. This resulted in the examination of 14 additional articles. In total, 192 articles were assessed, resulting in the extraction of 68 drugs with reported DDRs. The NB classifier contributed, thus, to the identification of 45% of the drugs with reported DDRs in the final list. The remaining 55% ($n = 81$) of drugs, were identified by going through the 152 articles manually selected as relevant, during my first year, while performing conventional key-word searches in PuMed.

Despite its apparently modest performance, the classifier highlighted articles, which had not been previously identified by conventional literature search methods, hence contributing considerably to expand the list of drugs with reported DDRs. Furthermore, articles selected by the classifier, which contained relevant information, had on average a higher score than articles without it, indicating that the classifier had the power to distinguish the desired signal among the background noise in the data (fig. 2.2). From figure 2.2 it was also possible to conclude that the subset of articles selected by the classifier, which are freely available in PubMed central (Free), could be used as a representative of the full set (All), given the similarity of the distribution of the NB classifier score in the two sets. This trend suggested that the modest performance of the classifier was not related with a poor choice of articles for data extraction and evaluation.

Additionally, when comparing the most frequent words in articles selected by the classifier as potentially relevant, and articles identified by the classifier as potentially non-relevant, it was clear that these two sets were rather alike. For example, the word "age" was present in both word clouds, with only a slightly higher frequency in the first set. The similarity between this two sets seemed to be the main reason for the modest performance of the classifier.

Nevertheless, a closer look to the top 25 most important tokenised words influencing the NB classifier, revealed that meaningful words, contributing to a positive or a negative score, were being used to distinguish between articles with and without relevant information (Table 2.1). For instance, among positive tokens, were

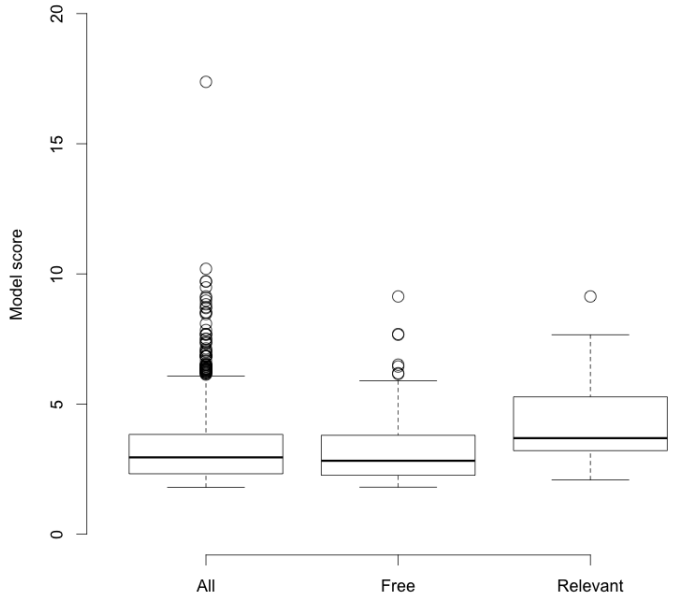


Figure 2.2: Box plots showing the distribution of the Naïve Bayesian document classifier score when different sets of articles were considered. All - 1,400 articles selected by the classifier has having potential relevant information; Free - 178 articles selected by the classifier has having potential relevant information, which are freely available in Europe PubMed Central; Relevant - 21 articles freely available in PubMed Central, selected by the classifier, which lead to the extraction of one or more drugs with reported DDRs. All *vs.* Relevant (t-test, p-value = 0.11); Free *vs.* Relevant (t-test, p-value = 0.090); All *vs.* Free (t-test, p-value = 0.44).

age-related words (or stems), such as, "maturation", "grow" and "neonatal", but also drug-related words from drug classes with suspected age-related differential drug responses, such as, "epilepsy", "seizur", and "depression". Among the negative tokens, on the other hand, were words not relevant for this project, such as "mice" and "mother", given that drug response studies in animal models or drug response studies during pregnancy were outside the scope of this project.

Table 2.1: Top 25 tokenised words (tokens) directing the Naïve Bayesian classifier.

Positive tokens	Weight	Negative tokens	Weight
secondary	0.63647	obtain	1.4568
maturation	0.58235	infect	1.3407
titrat	0.55123	direct	1.3407
epilepsy	0.55123	enhanced	1.2094
regular	0.55123	tool	1.2094
pharmacology	0.55123	process	1.2094
reuptake	0.55123	guid	1.2094
medlin	0.55123	randomis	1.2094
recurrent	0.55123	mother	1.2094
epileptic	0.55123	mice	1.0582
grow	0.55123	nevirapin	1.0582
awarenes	0.55123	concomitant	1.0582
ensur	0.55123	stage	1.0582
medication	0.53708	fibrosi	1.0582
neonatal	0.51869	transmission	1.0582
seizur	0.51869	medicin	1.0582
antifungal	0.50629	program	1.0582
unexpect	0.50629	linear	1.0582
compliance	0.50629	hiv	0.96178
arterial	0.50629	copy	0.88000
consistent	0.50629	load	0.88000
responder	0.50629	unchanged	0.88000
depression	0.50629	rna	0.88000
intermittent	0.50629	prophylaxi	0.88000
topiramat	0.50629	recruit	0.88000

2.3 Results and discussion

2.3.1 Overview of age-related differential drug responses

The effort of probing the medical literature for drugs with reported age-related differential drug responses resulted in a database with 149 drugs (Appendix A - Table 2), of which 147 are FDA approved[1]. The majority are small molecules

[1]Phenobarbital and nimesulide are only currently approved in some European countries.

(n = 146, 98%), and only 3 are biotherapeutics. The distribution of these drugs per DDR category and aged population can be found in Table 2.2. From this distribution it is clear that when transferring drugs from one aged population to another, safety should be the biggest concern, given the higher number of drugs in this category.

Table 2.2: Number of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric and/or geriatric populations when compared with adult populations, per differential drug responses categories and per aged population.

Population	DDR - safety	DDR - efficacy	DDR - control	Total Drugs
Paediatrics	40	31	31	90
Geriatrics	47	22	14	74
All	75	47	43	149

The 104 FDA approved drugs with reported DDRs, correspond to 6.5% of all FDA approved drugs[1]. It is an apparent small percentage, however the majority are indicated for the treatment of disorders, which are among the top 20 global leading disabling conditions, such as, depression, epilepsy, schizophrenia, anxiety and ischaemic heart disease [60], thus affecting a substantial number of paediatric and/or geriatric patients. For example, the lifetime prevalence rate of major depressive disorder (MDD) in adolescents has been estimated to range from 15% to 20%, which is comparable with the prevalence found in the adult population [61]. A look to the causes of the global disease burden in developed countries from the latest study (i.e. 2010), reinforces this idea, for instance, from figure 2.3 it is clear that mental and behavioural disorders explain a big portion of the disability-adjusted life years (DALYs) of teenagers (10-19 years), and even to a certain extend, of children (1-9 years), while cardiovascular and circulatory diseases are much more prominent among the elderly ($\geq$65 years).

Furthermore, the geriatric population is the biggest consumer of medicines,

[1]See chapter 3 for more details on the total number of FDA approved drugs.

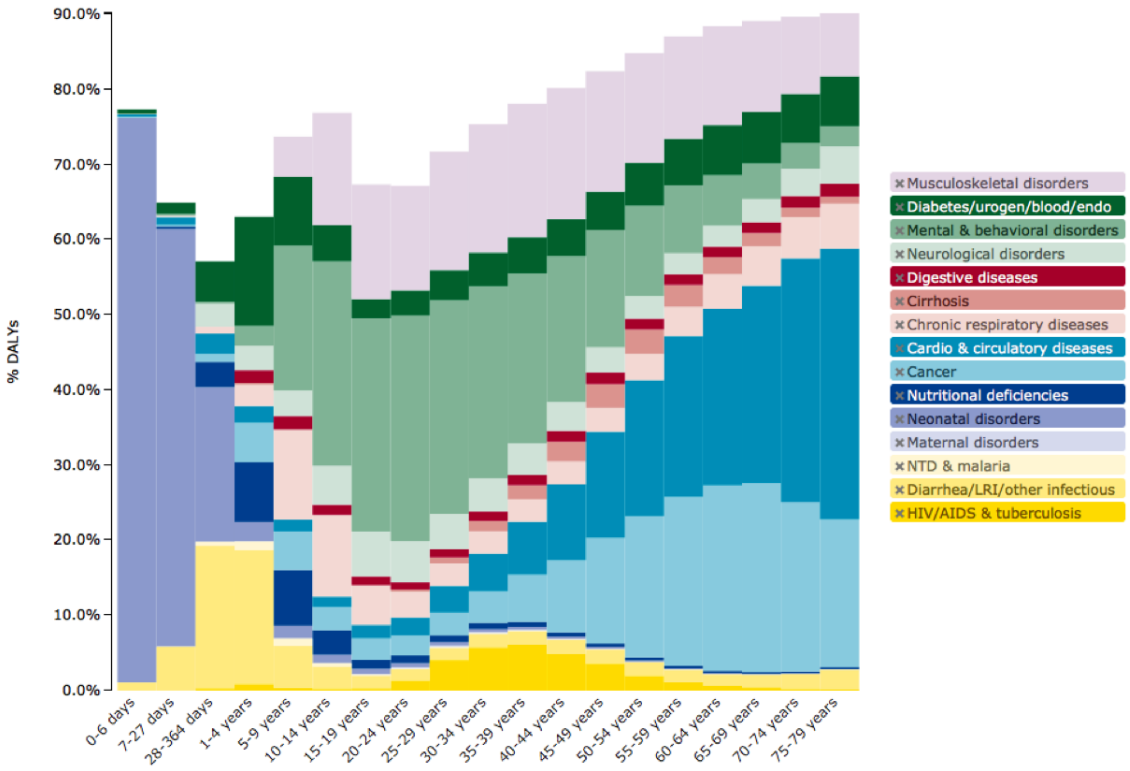


Figure 2.3: Global burden of disease (GBD) cause patterns from the 2010 study in developed countries considering both genders per age group. DALYs - Disability-Adjusted life years. Imaged adapted from the data visualisation tool provided by the Institute for Health Metrics and Evaluation (http://vizhub.healthdata.org/gbd-cause-patterns/).

accounting for 30-50% of all prescription drugs, and take on average two to nine different drugs per day [32]. Therefore, the lack of information on safety and efficacy of medicines in these populations, and how drug-drug interactions may change with age, poses an unnecessary risk to the patients. Hence, despite the small fraction of drugs which have shown age-related differential drug responses, its impact in the society is not to be ignored.

Curiously, newer drugs seem to be safer and more effective than older ones (fig. 2.4). This trend may be a result of the efforts from regulatory agencies, which have been pressing pharmaceutical companies to perform more paediatric clinical trials [3], and to include more geriatric patients in regular adult trials [34], thus helping to identify the right dose and safety profile in different aged populations.

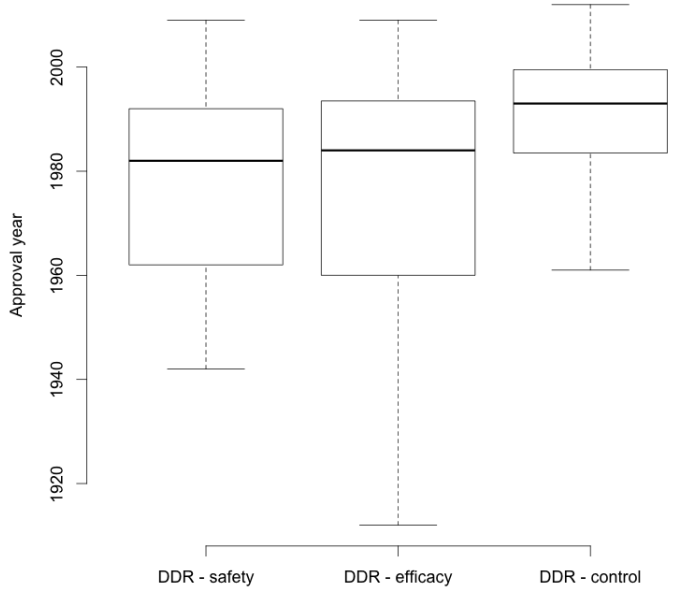


Figure 2.4: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control), independent of the aged population where the DDR was observed, per approval year. DDR - safety *vs.* DDR - control (t-test, p-value $= 2.8 \times 10^{-5}$); DDR - efficacy *vs.* DDR - control (t-test, p-value $= 3.4 \times 10^{-4}$); DDR - safety *vs.* DDR - efficacy (t-test, p-value $= 0.76$).

In terms of therapeutic indication, and chemical properties differences between the DDR categories, a throughout discussion will be presented per aged population in the next sections. Additionally, the WHO Anatomical Therapeutic Chemical Classification System code (ATC code)[62], is used as a way of obtaining a standard therapeutic indication for each drug. The ATC hierarchy consistently classifies drugs according to their pharmacological actions, therapeutic effects, chemical similarity and biological system of action.

2.3.1.1 In paediatric populations

Looking to the overall therapeutic indication of drugs with reported DDRs in paediatric populations per DDR category, it was clear that drugs modulating the

nervous system (ATC category N) were dominant, specially in the DDR-efficacy category, accounting for over 60% of the drugs in this group (fig. 2.5). Among the control group, the distribution of the therapeutic indications was evener, with drugs targeting the muscular-system (ATC category M) apparently standing out as being safe and effective when transferred from an adult to a child, like nimesulide and ibuprofen.

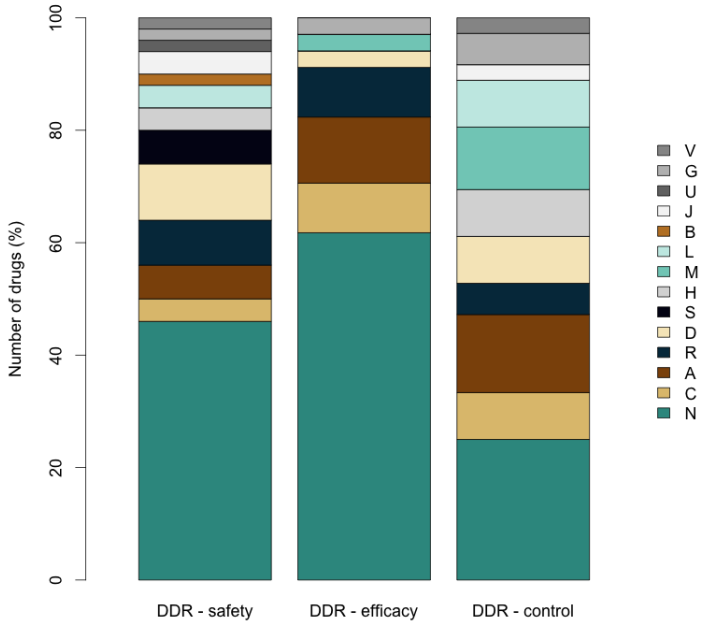


Figure 2.5: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per ATC category. A - Alimentary tract and metabolic system, B - Blood and blood forming organs, C - Cardiovascular system, D - Dermatologicals, G - Genito urinary system, H - Hormonal system, J - Anti-infectives for systemic use, L - Antineoplastic and immunomodulating agents, M - Musculo-skeletal system, N - Nervous system, P - Antiparasitic products, insecticides and repellents, R - respiratory system, S - Sensory organs, V - various, U - Unclassified. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.18); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.063); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.49).

A closer look to the data revealed that among drugs with safety issues, when transferred from the adult population to the paediatric population (DDR-safety), there were three main drug classes standing out: antidepressants ($n = 10$), antiepileptics ($n = 4$) and general anaesthetics ($n = 3$). The high number of antidepressants is mostly due to the presence of selective serotonin re-uptake inhibitors (SSRIs) in the database. This is due to the fact that FDA emitted a black box warning against the members of this class in 2004, due to the increased risk of suicides detected among the paediatric population when treated with this type of medicine [63]. Among the antiepileptic drugs, there was no common safety concern for their inclusion in the database. For example, in the case of gabapentin and levetiracetam, when given to children, there was an increased in hostility and aggression when compared with adults [64; 65], while in case of lamotrigine there was an increase in rash incidence [66], and for valproic acid an increase in hepatoxicity [42]. This trend was also true among the general anaesthetics in the list. For instance, when given sevoflurane to children there seemed to be an increase in seizures [64], but with deflurane there was an increase in adverse respiratory events [67]. Curiously, glucocorticoids, commonly used for the treatment of asthma, were under represented in this DDR category. This was due to the fact that only fluticasone propionate was routinely and explicitly mentioned in the literature as impairing children growth, despite some studies also alluding to similar safety issues with budesonide and beclomethasone dipropionate [68].

Among drugs with a lack of efficacy, when transferred from the adult population to the paediatric population (DDR-safety), the main drug classes were: antidepressants ($n = 8$), antiepileptics ($n = 5$) and drugs indicated for the treatment of migranes ($n = 4$). The high number of antidepressants was again mostly due to the presence of SSRIs in the database, with fluoxetine being the only member of this class widely accepted by the community has having shown some efficacy in the paediatric population [63]. However, data from another SSRI, escitalopram, have recently been reviewed, with a favourable outcome for its use in the paediatric population[1] [69]. Interestingly though, when looking to the chemical properties of

[1]Escitalopram is the S-enantiomer of citalopram, a racemic mixture of the active S-enantiomer and the clinically inactive R-enantiomer. Citalopram is not considered effective in paediatric populations. However, since the recommend dose for citalopram is the double of the one recommended for escitalopram, the same amount of S-enantiomer is expected to reach

the SSRIs, nor fluoxetine nor escitalopram appeared to stand out, suggesting other cause for their effectiveness. Although caution is necessary to avoid over interpretation of these observations, since antidepressants trials in children and teenagers generally shown a high rate of response to placebo, with even newer generations of antidepressants like nefazodone, venlafaxine and mirtazapine only showing a mild benefit, but never reaching significance [70]. In case of antiepileptics, there was a prevalence of drugs which target the GABA A receptor, or calcium channels. Finally, among drugs indicated for the treatment of migraines, there was a prevalence of drugs which target dopamine 2 receptors, and serotonin 1B and 1D receptors. Curiously, these drugs have only shown a lack of efficacy when administered orally, but seemed effective after nasal administration. This effect is probably a result of two factors, first the drug is not exposed to gut metabolism, and secondly, since migraines in children manifest as short and sharp peaks of pain, only drugs that can reach the neuropathic centre of pain quickly, can show some efficacy.

Among the minor drug classes, which have shown a lack of efficacy when transferred from the adult population to the paediatric population, was noteworthy the presence of antitussives and antihistaminics. Despite the low number of drugs from these classes in the database, these common over-the-counter medications used for the treatment of colds, are generally not considered effective in children by the medical community, and sometimes even harmful [71; 72].

Finally, among the control group, drugs which have not shown a difference in efficacy when comparing paediatric and adult populations (DDR-control), no particular drug class stood out. The drugs in this category are indicated for diverse disorders such as the treatment of hyperparathyroidism associated with chronic kidney disease (CKD), attention deficit hyperactivity disorder (ADHD), schizophrenia and/or autism, but also for the treatment of pain and fever.

In terms of the drugs' chemical properties, such as, molecular weight, logP, number of hydrogen bond donors, number of hydrogen bond acceptors, number of aromatic rings, and the polar surface area (PSA), when looking to its distribution per DDR category, only PSA had some discriminatory power. Specifically, drugs with safety issues (DDR-safety) had a lower PSA on average, when compared

the therapeutic target, consequently, it is not clear why one drug would be effective and not the other. Therefore, caution is necessary when interpreting this new reviewed data.

with drugs in the control group (DDR-control) (fig. 2.6). Interestingly, drugs with a lower PSA have a tendency to penetrate the blood-brain-barrier (BBB), and enter in the central nervous system (CNS) [73]. This trend thus seems to suggest that drugs in the DDR-safety category may have a higher propensity to lead to undesirable CNS-related side-effects.

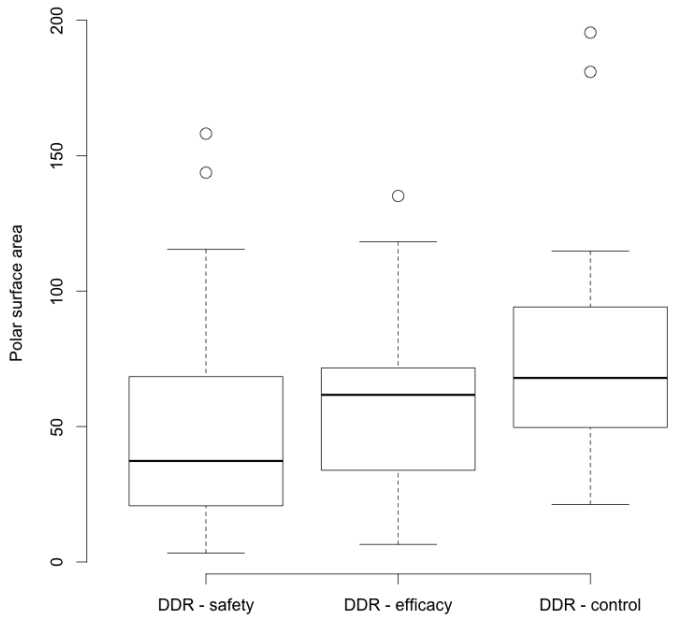


Figure 2.6: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per polar surface area (PSA). DDR - safety *vs.* DDR - control (t-test, p-value $= 6.9 \times 10^{-3}$); DDR - efficacy *vs.* DDR - control (t-test, p-value $= 0.022$); DDR - safety *vs.* DDR - efficacy (t-test, p-value $= 0.53$).

2.3.1.2 In geriatric populations

Contrary to what I observed in section 2.3.1.1, the overall therapeutic indication of drugs with reported DDRs in geriatric populations per DDR category, was dominated by drugs modulating the cardiovascular system (ATC category C), followed by the ones modulating the nervous system (ATC category N) and to

some extent by drugs modulating the alimentary tract and metabolic system (ATC category A) (fig. 2.7).

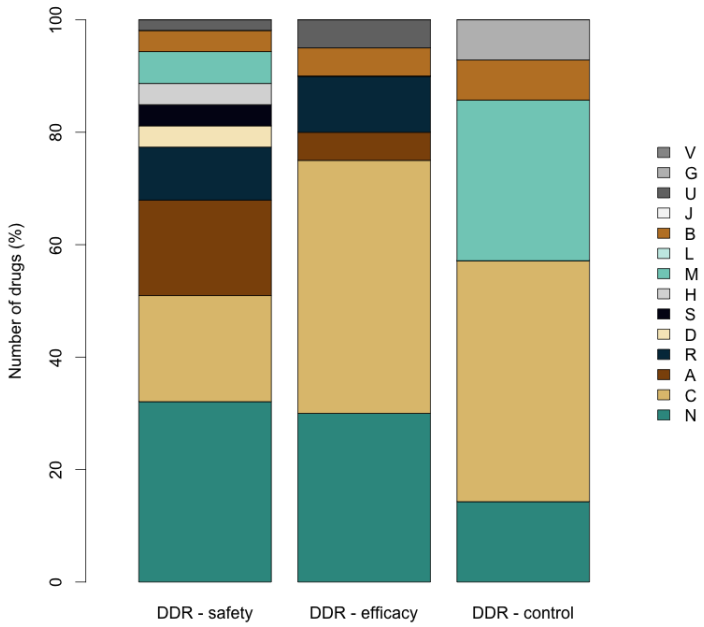


Figure 2.7: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per ATC category.. A - Alimentary tract and metabolic system, B - Blood and blood forming organs, C - Cardiovascular system, D - Dermatologicals, G - Genito urinary system, H - Hormonal system, J - Anti-infectives for systemic use, L - Antineoplastic and immunomodulating agents, M - Musculo-skeletal system, N - Nervous system, P - Antiparasitic products, insecticides and repellents, R - respiratory system, S - Sensory organs, V - various, U - Unclassified. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.037); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.077); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.48).

Exploring in more detail the therapeutic indications of drugs, which have shown severe or unseen side effects in geriatric populations when compared with adult populations (DDR-safety), it was clear that seven main drug classes stood out: antidepressants ($n = 4$), anxiolytics ($n = 4$), antihistamines ($n = 4$), sulfonylureas (n

= 4), anaesthetics ($n = 3$) or adjuncts to general anaesthesia ($n = 3$), diuretics ($n = 3$), and antiarrhythmics ($n = 2$) or antihypertensives ($n = 2$). There was a common pattern of side effects among the antidepressants, anxiolytics and anaesthetics in this category. In concrete, the benzodiazepines and tricyclic antidepressants in the database were routinely associated with an increase of fractures in the elderly after exposure to this type of medication [12], suggesting that a careful dosing scheme should be in place to avoid over inhibition of the GABA system, i.e. a lower dose may be required to achieve the same safety profile when compared with adults. Additionally, the use of benzodiazepines has also been associated with loss of memory [12]. In case of sulfonylureas, drugs indicated for the treatment of diabetes type II, there was a consensus in the literature that these medicines impose a risk of severe prolonged hypoglycaemia in the elderly, and thus should be avoided in this population [38; 40]. The first generation of antihistamines was also deemed unfit for the elderly by the medical community, given their high anticholinergic properties, which translates into a higher risk of confusion, dry mouth, constipation and other anticholinergic effects [38; 40]. In case of diuretics, the prevalent reported side effects in the elderly were lower levels of sodium and potassium than in adults receiving the same medication, which compromised the regular functions of the renal and the cardiovascular system [45]. Among the antiarrhythmics or antihypertensives in the database, such as nifedipine, the main reported side effects were a higher risk of hypotension in the elderly when compared with adults [38]. Finally, despite the fewer number of non-steroidal anti-inflammatory drugs (NSAID) in the database, there was a consensus in the literature to avoid its use in the elderly, due to a higher risk of stomach bleeding [12].

Regarding drugs that have shown a lack of efficacy when transferred from adults to the elderly (DDR-efficacy), there were three noticeable drug classes: antihypertensives ($n = 4$), diuretics ($n = 3$), and analgesics ($n = 3$). The high number of antihypertensives reflected the presence of β_2-adrenergic receptors antagonists in the database. The reason being, that in the literature, these medicines were no longer considered effective when given to lower the arterial pressure in the elderly, possibly due to a lower expression level of β_2-adrenergic receptor in this population, being nowadays commonly replaced by angiotensin-converting-enzyme (ACE) antagonists [13]. On the other hand, the lack of efficacy of diuretics, when pescribed

to lower the arterial pressure in the elderly, was described in the literature as being related with alterations in pharmacokinetics rather than pharmacodynamics with ageing [13]. In the case of analgesics, drugs targeting the μ opioid receptor were described as not been effective with current dosing schemes [13]. Finally, there was one more class often referred in the literature as having a lack of efficacy in the elderly: bisphosphonates for the prevention of hip fractures [12]; however, as the information was always referent to the class, and not to a specific drug, the database does not reflect this trend.

Among the drugs in the control group, which have not shown any difference in efficacy or safety when transferred from the adult population to the geriatric population (DDR-control), there were two main drug classes which stood out: angiotensin-converting enzyme (ACE) antagonists ($n = 4$) for the treatment of hypertension; and adjuncts to general anaesthesia to facilitate endotracheal intubation ($n = 4$), which modulate the muscle-type nicotinic acetylcholine receptor, such as, atracurium and pancurronium. Statins were also present in this group ($n = 2$), and despite the low numbers, were generally described in the literature as safe and effective in the elderly [74]. However, the total number of drugs in this DDR category is rather low, hence any trend needs to be interpreted with caution.

Finally, when observing the distribution of the drugs' chemical properties, it was clear that drugs in the control group (DDR-control) had on average a higher molecular weight (fig. 2.8), while drugs in the DDR-efficacy category had a slight lower logP (fig. 2.9), and like the drugs with reported DDRs in paediatrics, drugs in the control group also had a higher polar surface area (fig. 2.10). This trend, as mention before, reinforces the idea that drugs in the other DDR categories, specially in the DDR-safety category are more prone to cross the blood-brain-barrier, and enter in the central nervous system, and consequently lead to undesirable CNS-related side-effects. However, as mention before, the total number of drugs in the control group is rather low, therefore the trends observed may not necessarily be representative.

2.3.2 Limitations of the dataset

Saying that a medicine is not effective in a certain aged population, and/or more prone for side effects, depends on various factors, namely: the therapeutic indica-

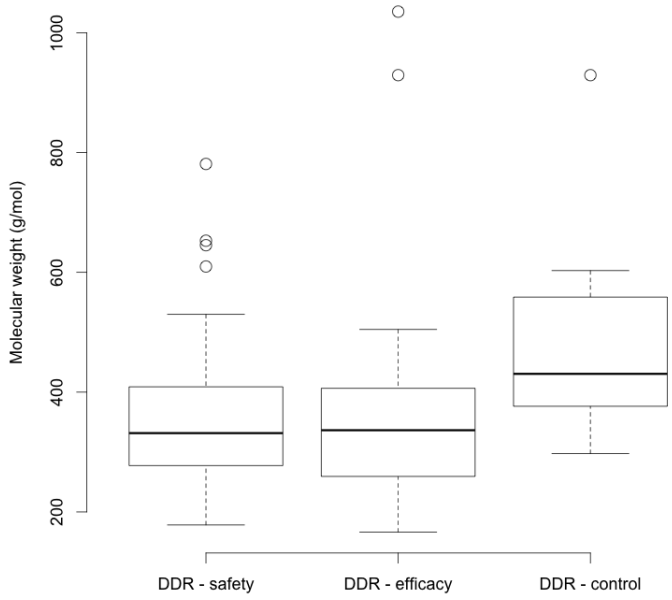


Figure 2.8: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per molecular weight (g/mol). DDR - safety *vs.* DDR - control (t-test, p-value = 0.016); DDR - efficacy *vs.* DDR - control (t-test, p-value = 0.12); DDR - safety *vs.* DDR - efficacy (t-test, p-value = 0.59).

tion being tested and how it is defined; the age range of the patients in the study, as well as the criteria used to include and exclude them, such as, the presence of other co-morbid disorders; the dose tested; and the endpoints of the clinical trial, for instance, the use of a specific scale to measure the efficacy of an antidepressant. However, the medical literature doesn't always provide this information explicitly or unambiguously, for example, the dose given can sometimes be described as "high" or "low", making it harder to compare studies. Consequently, my dataset of drugs with reported DDRs can only reflect the main trends in the field, and it is subjected to the amount of information available and its quality per drug class, which is not even. For example, the literature is rather conflicting when describing the effect of antiepileptics in the treatment of infantile spams [75], but seems to

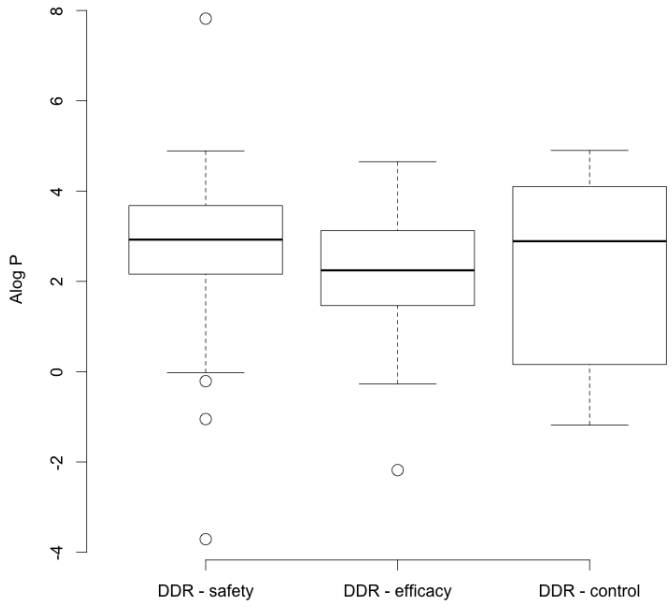


Figure 2.9: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per AlogP. DDR - safety *vs.* DDR - control (t-test, p-value = 0.57); DDR - efficacy *vs.* DDR - control (t-test, p-value = 0.78); DDR - safety *vs.* DDR - efficacy (t-test, p-value = 0.20).

have reached a consensus in terms of the effect of SSRIs in the paediatric population [63], therefore, not many antiepileptics were added to the DDR database, as it was unfeasible to verify which study better represented the overall response in paediatric populations. For other classes, despite some consensus in the literature, the number of patients used to draw conclusions were rather small. For instance, to assess the efficacy of antipsychotic agents for the treatment of irritability of autism in paediatric populations, the number of children in the studies ranged from 8 to 218, with an average value of 48 [76], which raises the question of how applicable and general is the outcome. In these cases, I gathered as much information from different sources as possible to help clarify and discern which drugs should be added to the final database. Additionally, information about drug response in

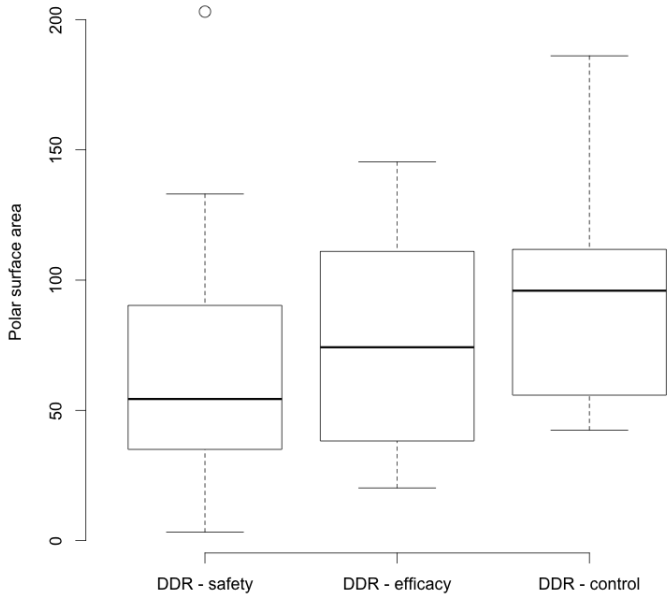


Figure 2.10: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per polar surface area (PSA). DDR - safety *vs.* DDR - control (t-test, p-value = 0.043); DDR - efficacy *vs.* DDR - control (t-test, p-value = 0.30); DDR - safety *vs.* DDR - efficacy (t-test, p-value = 0.25).

children, teenagers or in the elderly may change with time, for instance, SSRIs were though to be safe in paediatric populations until 2004 when FDA reviewed some clinical trials data, and decided to emit a black box warning for all members of this class [63]. Therefore, caution is always required when interpreting this data, as it just reflects the information we have available now, but may not be valid in the future, i.e. as more drugs get tested in paediatric or geriatric populations, unsafe features may be revealed or the efficacy diluted, for instance paroxetine was initially considered effective in children, until new data was added to the pool and reanalysed [77].

Given all this, I have tried to be as consistent with my annotations as possible, therefore a drug was not considered effective if the patients response to the

treatment was not significantly different from the response with placebo, or if the response was lower than 50%[1]. This means that a drug with a mild response of 30%, like levetiracetam when used for the treatment of intractable epilepsy in children and teenagers [65], was added to the DDR-efficacy group, and not considered effective in the paediatric population. Likewise, a drug was not deemed safe in an aged population, if the side effects reported, were more pronounced in the young or in the elderly than in the adult population, or if the side effects were unseen in the adult population.

Although, in spite of all of these limitations, the data reflects the accepted trends in the field, including famous historical cases, briefly described in the introduction of this thesis, such as, chloramphenicol and the grey baby syndrome, the safety profile of paracetamol in children [78], and the lack of efficacy of the β_2-adrenergic antagonists in the geriatric population [13].

2.4 Summary

By probing the medical literature, I identified 106 drugs with reported age-related differential drug responses and 43 drugs, which have not shown any difference in efficacy or safety when transferred from the adult population to the paediatric and/or geriatric population. Despite the limitations of the method used for data extraction, the final dataset reflects the most known cases of drugs with reported DDRs.

Overall, the majority of drugs in the database, have shown more safety issues when transferred from one aged population to another, than a lack of efficacy. In terms of therapeutic indications, among the drugs with reported DDRs when given to paediatric populations, the main drug classes were involved in the modulation of the nervous system, while drugs with reported DDRs when given to the elderly, were mostly interacting with the cardiovascular system. Lastly, drugs in the control group seem to be newer than drugs with reported DDRs.

In the next chapter, this dataset will be used to identify which pharmacologically relevant genes, responsible for the drugs' safety and efficacy, are pre-

[1]The true level of efficacy is class dependent, however per norm a response between 60 to 80% is desirable.

dominantly associated with reported safety issues and/or a lack of efficacy when comparing different aged populations. Later on, this information will be use to investigate how the expression of those genes changes with age, and if those changes could help explain the occurrence of some of the DDRs reported in the database.

2.5 Material and Methods

2.5.1 Machine learning application for literature search optimisation

The workflow for literature search optimisation was build in Pipeline Pilot [79]. A SOAP protocol was used to programmatically access the Europe PubMed Central database and retrieve articles containing the word "drug", independent of its localisation within the article. These articles were subsequently filtered, and only articles with age-related (Appendix A), and drug-related words (Appendix A) in the title or in the abstract, were kept. The total number of age- and drug- related words per article were counted, and articles with at least 5 of these words, were manually inspected. If the article seemed to contain relevant information about drugs with reported DDRs, a "relevant" flag was assigned, if not, a "non-relevant" flag would be attributed instead. Out of the 650 articles evaluated, 125 were assigned a "relevant" flag, and the remaining ones a "non-relevant" banner. In order to allow a fair representation of "relevant" and "non-relevant" articles, only the 125 relevant articles and 125 randomly selected non-relevant articles were used to train and validate the NB classifier (PP component: Learn good from bad). This set was further and equally divided into test set and training set. Both sets were submerge to a series of processing steps necessary to train and validate a bag-of-words (BoW) document classifier. This approach was implemented in a standard way by appropriately tokenizing titles and abstracts for the two classes of documents. The resulting terms were further submitted to a series of text mining preprocessing operations, such as punctuation removal, case normalisation, and removal of stop words (Appendix A). A document vector was then generated for each document, encoding the absence or presence of the remaining terms in a binary string. The vectors were then used as binary descriptors for an NB classifier.

Articles that were not used to train and validate the model, were then fed to the classifier, and information about the likelihood of an article to contain relevant information for the project was kept. A SOAP protocol was then used to verify which, of the 1,400 articles highlighted by the classifier as potential interesting, were freely accessible in PubMed central. Out of the 1,400 articles, 178 were freely accessible. These articles were downloaded and manually evaluated.

2.5.2 Validation of the NB classifier

The performance of the classifier was estimated based on its sensitivity, specificity, matthews correlation coefficient (MCC) and receiver operating characteristic (ROC) curve, calculating using the test set. Sensitivity was defined as the fraction of true positive predictions of the total positive documents (documents flagged as relevant): True Positives / (True Positives + False Negatives). Similarly, specificity was defined as the fraction of true negative predictions of the total negative documents (documents flagged as non-relevant): True Negatives / (True Negatives + False Positives). The MCC was calculated using eq. 2.2. The MCC parameter measures the quality of binary classifications and ranges from -1, i.e. complete disagreement between prediction and observation, 0, no better than a random prediction, and 1, complete agreement between prediction and observation.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)}} \tag{2.2}$$

The sensitivity of the classifier was 0.695, the specificity 0.729, the MCC 0.424, and the area under the curve (AUC) of the ROC curve 0.766 (fig. 2.11). Based on these values, the performance of the NB classifier was considered satisfactory.

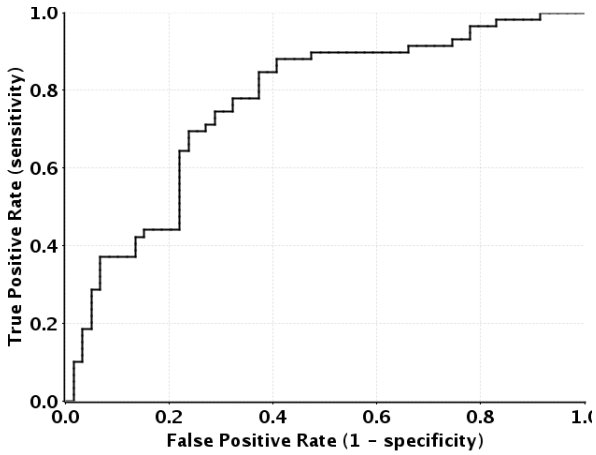


Figure 2.11: Receiver operator characteristic curve (ROC curve) for the NB classifier.

2.5.3 Chemical properties, approval year and ATC code of drugs in the DDR database

For each drug in the database, with reported age-related differential drug responses, the following chemical properties were retrieved from ChEMBL 17: molecular weight, log P, polar surface area, number of hydrogen bond acceptors and number of hydrogen bond donors. Box plots of the distributions of these chemical properties per DDR category were generated in R [80]. Additionally, the respective ATC code(s) was extracted, for each drug, from the ATC codes in ChEMBL 17, assigned to either its respective dosed ingredient (if applicable) or the parent ingredient itself. The approval year was also extracted from ChEMBL 17, and corresponds to the first known worldwide approval date.

Chapter 3

Compilation of Drug targets, Metabolic enzymes and Transporters

After having created a database of drugs with reported age-related differential drug responses (DDR), and having established the therapeutic indications prone for such events, the next objective of my work, presented in this chapter, was to identify the pharmacologically relevant human genes responsible for the drugs' efficacy and safety. These genes are the most likely candidates to influence drug response, and thus play a potentially determining role in age-related differential drug responses if their expression patterns changes with age. Therefore, in this chapter I describe the work conducted to compile an up-to-date and comprehensive list of drug targets, metabolic enzymes and transporters for each drug in the DDR database, and all other FDA approved drugs. Additionally, a thorough examination of the distribution of drug targets, metabolic enzymes and transporters families, in the context of age-related differential drug responses, will be discussed.

3.1 Introduction

3.1.1 Drug targets compendiums

3.1.1.1 Historical perspective and existing resources

Understanding the molecular basis of drug therapy has always been considered a valuable asset in drug discovery. For instance, such knowledge can be useful for the design of new medicines against precedented drug targets, but with safer profiles and/or new formulations to address currently untreated conditions. But also, to identify potential new drug targets, rationalise side-effects or, in a more recent context, rationalise how drug response may vary when the drug "hits" a product of a polymorphic gene. On that account, many compendiums of drug targets or "drug to drug targets" maps were compiled throughout the years. The first attempt dates from 1996, when Drews and Reiser, took on the task of estimating the number of human molecular efficacy targets for small-molecule approved drugs [81; 82]. After the completion of the human genome project in 2001, the first database of drugs and their efficacy targets, Therapeutic Targets Database (TTD) [83; 84], was created. Shortly after, Hopkins and Groom [85], and Golden [86] introduced the concept of a finite druggable genome, and made some estimates as to its size. Later, others decided to analyse and report in depth on the number and nature of drug targets [87; 88]. From this period, as well, is the creation of DrugBank [89; 90], now an established resource with information on drugs and their potential pharmacological targets. With the development of text-mining approaches, it wasn't long before the creation of SuperTarget [91], a database with semi-automatically inferred direct and indirect drug targets. More recently, Rask-Anderson *et al* [92] have provided an updated view on the status of current drugs and the human targets believed responsible for their primary efficacy, Munos [93] has highlighted the trends in drug and target innovation for the last decade, and Agarwal *et al* [94] have revealed the overlap and uniqueness in the drug targets that are currently being pursued by the industry.

Despite the extensive work around these compendiums or drug maps, only a minority of those has actually been made fully available to the research community, such as TTD, DrugBank, SuperTarget, and the supplementary material in Rask-

Anderson *et al.* However, in the context of this thesis, only TTD, DrugBank and the supplementary material in Rask-Anderson *et al* are of interest, as they are the only currently available resources able to provide an up-to-date map of marketed drugs and their direct drug targets. Even so, these resources don't come without limitations.

3.1.1.2 Issues with existing resources

When examining in detail the resources introduced in the previous section, it becomes clear that it is a challenge to retrieve, from each of them, a single list of approved drugs with their associated molecular efficacy targets, without making some compromises. For instance, TTD provides target annotation on Food and Drug Administration (FDA) approved drugs, clinical trial drugs and experimental drugs. However, if the drug is deemed non-selective, or the efficacy target is a protein complex, TTD only provides a target name, but no reference identifier to an external protein resource. This lack of information on subunit composition or family members, imposes some difficulties when one needs to further link drugs to human genes, as the user has to first try and figure out which Universal Protein Resource (Uniprot) IDs [95] are part of a certain protein complex, or a certain protein family.

DrugBank, on the other hand, always provides a link to the Uniprot for each target associated with the FDA approved drugs or experimental drugs in their database. However, as this resource does not discriminate between efficacy targets, targets responsible for side-effects, or targets responsible for secondary pharmacology, it is not possible at the moment to obtain a list of only efficacy targets[1] from DrugBank.

Finally, the map presented by Rask-Anderson *et al* addresses some of the target annotation issues that I have just described, mainly, they focus on efficacy targets, and they tend to provide the subunit composition of a protein complex target if the drug is non-selective. Despite that, as the scope of the study was limited to a subset of FDA-approved drugs, for 672 drugs, information about the mechanism of action is not available.

[1] targets directly responsible for the therapeutic effect of a drug

Giving all these limitations and inconsistencies between drug target sets, I have decided to create a new drug to drug target map.

3.1.2 Metabolic enzymes and transporters compendiums

3.1.2.1 Historical perspective and existing resources

Contrary to what we see in the "drug target world", existing compendiums of metabolic enzymes and transporters are harder to find. The information in the literature is organised by protein families, rather than combined to give an overview of all metabolic enzymes and transporters. The best attempt to overcome this barrier was led by the PharmaADME consortium [96], composed of individuals from academia and pharmaceutical and genomic technology industries, who recently sought to put together a consensus list of key protein/genes involved in the absorption, distribution, metabolism and excretion (ADME) of drugs.

Regarding drug to metabolic routes and potential transport system maps, there are just a couple of known resources with this information, however they tend to focus on a subgroup of related enzymes (i.e., CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, and CYP3A4 for phase I metabolism), such as Pharmacology weekly [97] or the popular P450-mediated drug interaction table from the Indiana University [98]. One notable exception is SuperCyp[1], a database which stores information on drug metabolism, by manually curating articles retrieved with text-mining tools [99]. Its main focus is phase I drug metabolism by members of the cytochrome P450 (CYP) family, although, it also provides information on phase II enzymes and transporters for a couple of drugs.

3.1.2.2 Issues with existing resources

As a starting point, the metabolic enzymes and transporters compendium from the PharmaADME consortium, is a good gene set, providing a comprehensive list of the most important proteins involved in the absorption, distribution, metabolism and excretion of drugs. However, as the set is not meant to be exhaustive, there

[1]DrugBank 3.0 also provides information on drug metabolism for well studied FDA approved drugs. However, much of the data is already covered in SuperCyp and/or the P450-table provided by the Indiana University.

are a few enzymes and transporters that are currently missing.

Likewise, SuperCyp was not designed to be a comprehensive resource with detailed information on all drug metabolism phases for all FDA approved drugs. However, given that information on drug metabolism is seldom reported for the majority of drugs, it would be impossible to manually extract the metabolic enzymes and transporters for all the FDA approved drugs, and create a better resource than SuperCyp.

Consequently, I decided to compile a new list of metabolic enzymes and transporter genes from the gold standard reviews in the field, instead on relying on the one provided by PharmaADME, but use the data in SuperCyp to get the most comprehensive and up-to-date maps of "drug to metabolic enzyme" and "drug to transporter" map.

3.2 Implementation

3.2.1 A new "drug to drug target" map

As mentioned in the introduction, given the limitations of available resources with drug to drug targets maps, I have decided to create a comprehensive drug map, and assign efficacy targets afresh from the primary literature and prescribing information. The 'prescribing information' or 'package insert' is an official document provided by the pharmaceutical company marketing the drug, with consistently presented and detailed information about the drug itself, including information on clinical pharmacology, such as the drug's mode of action, thus making it a valuable source of information regarding the drug's efficacy target[1].

Additionally, I have opted for a systematic approach, and whenever known, have manually extracted the efficacy targets for all FDA approved drugs and WHO antimalarials. Using this approach, I have covered and extracted the drug targets of all potential relevant drugs for this thesis.

A task of this magnitude is, however, quite time-consuming, as extracting the relevant information for one single drug can take up to one hour. On that

[1]In the European Union, the corresponding document is the Summary of Product Characteristics (SPC).

account, Ramesh Donati, Patricia Bento, Anne Hersey and Anna Gaulton from the ChEMBL[1] group, joined me in a effort to speed up the curation process. The data was later fully integrated in the ChEMBL schema, and it was released with ChEMBL 17 [100].

Currently, the drug to drug target map has 1273 small molecules and 224 biotherapeutics drugs with an efficacy target assigned. As mentioned in chapter 2, we have mapped each drug to the WHO Anatomical Therapeutic Chemical Classification System code (ATC code)[62], as a way of obtaining a standard therapeutic indication for each of them. The ATC hierarchy consistently classifies drugs according to their pharmacological actions, therapeutic effects, chemical similarity and biological system of action. From the distribution of drugs per ATC code category in figure 3.1, it is clear that the majority of drugs in our data set are approved to treat conditions that affect the nervous system (category N of the ATC code), followed by drugs that act on the cardiovascular system (category C of the ATC code), and drugs indicated for the treatment of systemic infections (category J of the ATC code).

From the targets perspective, we have flagged 612 human proteins[2], and 127 pathogen proteins as drug targets[3]. These numbers do not actually reflect unique efficacy targets, as a big proportion of those are proteins subunits or isoforms assigned as drug targets to non-selective drugs. Consequently, if grouped into reasonable molecular entities, such as protein complexes or protein families, the sum reduces to 315 human proteins, and 52 pathogen proteins (see section 3.5.2 for a detailed explanation).

Other biomolecules, such as DNA, or cellular complexes like the ribosome, have also been flagged as efficacy targets. They are responsible for the therapeutic effect of a small, but significant fraction of drugs in our set (11.8%).

Finally, a closer look to the familial patterns of the human proteins flagged

[1]ChEMBL is relational database of bioactive drug-like small molecules, it contains 2-D structures, calculated properties (e.g. logP, Molecular Weight, Lipinski Parameters, etc.) and abstracted bioactivities (e.g. binding constants, pharmacology and ADMET data). It is currently maintained by the service team at EMBL-EBI under John P. Overington supervision.

[2]The complete list can be found in Appendix B in Table 3.

[3]The number of pathogen protein targets does not include all potential orthologues. Please refer to section 3.2.1.1 for a thorough explanation.

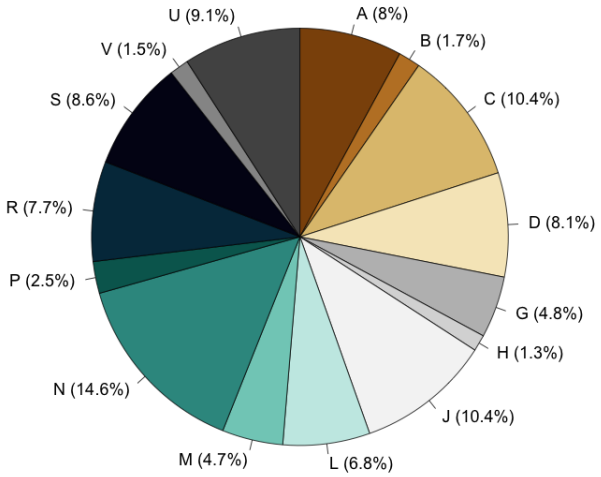


Figure 3.1: Percentage of FDA approved drugs and WHO antimalarials per ATC category. A - Alimentary tract and metabolic system, B - Blood and blood forming organs, C - Cardiovascular system, D - Dermatologicals, G - Genito-urinary system, H - Hormonal system, J - Anti-infectives for systemic use, L - Antineoplastic and immunomodulating agents, M - Musculo-skeletal system, N - Nervous system, P - Antiparasitic products, insecticides and repellents, R - respiratory system, S - Sensory organs, V - various, U - Unclassified.

as drug targets in our new map, shows that enzymes are the largest familial-class with 267 members, corresponding to 41.8% of total human protein drug targets (fig. 3.2a), followed by ion channels (21.8%) and membrane receptors (17.4%). However, when looking to the size of the family class in terms of how many drugs exert their therapeutic effect by modulating members of that family, membrane receptors supersede enzymes by 10% (fig. 3.2b). Additionally, despite the small number of unique transcription factors in our target list (2.8%), this familial-class accounts for the therapeutic action of almost as many drugs as the familial-class of ion channels.

3.2.1.1 Problematic assignments

Assigning a clear and unambiguous therapeutic target to a drug is sometimes a more daunting task than one might initially think, as there is often disagreement or inconsistency between existing primary sources. In particular, the primary li-

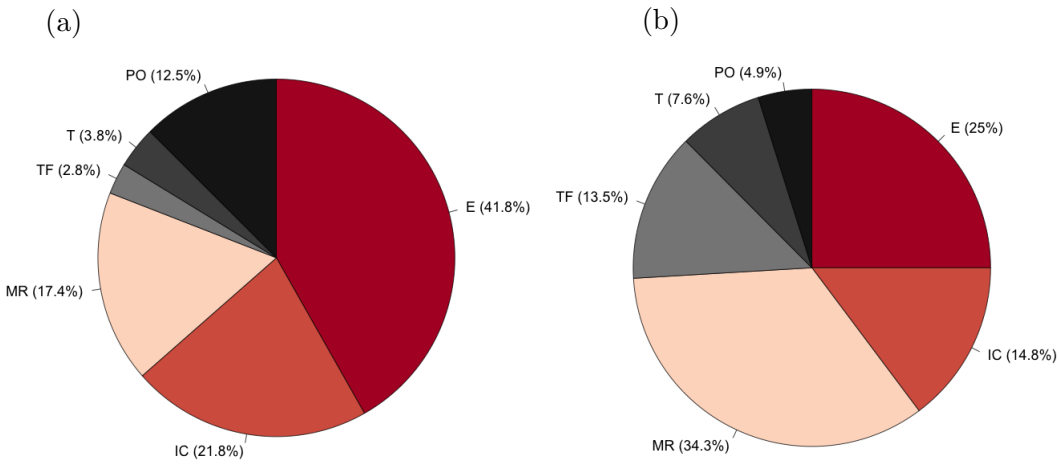


Figure 3.2: a) Percentage of human protein efficacy targets per drug target class category. b) Percentage of FDA approved drugs whose drug target is a human protein, per drug target class category. E - Enzyme, IC - Ion channel, MR - Membrane receptor, TF - Transcription factor, T - Transporter, PO - Protein other.

terature can sometimes be contradictory, and the prescribing information doesn't always provide all the clinical pharmacology that one may expect on the molecular interactions of the drug. For instance, how do we assign efficacy targets to a compound when the prescribing information states that the drug is a muscarinic antagonist, or a voltage-gated potassium channel blocker, or a broad-spectrum β-lactam antibiotic? One possible solution would be to list the 5 muscarinic receptors for the first case, the 40 α-subunits that may form voltage-gated potassium channels for the second, and for the latter, pick all the penicillin-binding proteins (PBPs) from all the species against which the drug is effective. However, in the case of the human targets, a more restricted subset, based on selectivity data or expression data could also be chosen, while for the pathogen targets, a representative pathogen species could be chosen and only the biomolecules or cellular components of that species could be assigned as drug targets.

For our assignments, I have identified such subsets among the human targets where sufficient information was available to do so, but have otherwise listed all possible proteins where there was inadequate evidence to determine which subunits or family members play a key role. For example, all antimuscarinic agents indicated

to treat bronchospasms had muscarinic receptor 3 (M3) assigned as their efficacy target, since this is the muscarinic receptor with the highest expression levels in the airways. In the case of pathogen targets, for instance, if a drug was deemed effective against a variety of bacteria species, *Escherichia coli* was chosen as the representative species, thus biomolecules or cellular components from *Escherichia coli* were assigned as drug targets to those drugs.

These rules are part of our attempt to define a consistent set of guidelines to help us assign therapeutic targets, which are fully described in section 3.5.1. It is important to emphasise though, that we have not attempted to assign targets purely on the basis of reported biochemical and pharmacology data. While there may be evidence for drugs binding with moderate, or even high, affinity to multiple additional targets, we have not considered these as efficacy targets unless there is also compelling evidence for their role in the therapeutic effect of the drug.

Nevertheless, even using these guidelines, the final assignment is subjective. Therefore, for a few FDA-approved drugs, there will always be disagreement between curators when assigning efficacy targets. For example, for dronedarone, an antiarrhythmic approved in 2009, the only information in the package insert is that it has antiarrhythmic properties belonging to all four Vaughan-Williams classes[1]. This implies that at a molecular level, it has the capacity to inhibit sodium channels (Class I), beta adrenergic receptors (Class II), voltage-gated potassium channels (Class III) and L-type calcium channels (Class IV) [101]. In this case, we have decided to list all of the alpha ion channel subunits[2] and beta adrenergic receptors as dronedarone efficacy targets, as we were unable to identify, from the primary literature, the most relevant set of targets for its antiarrhythmic action. Others have however, decided on a different list. For instance, Rask-Anderson *et al* assigned the voltage-gated potassium channel HERG, due to its importance in the control of the QT interval[3], and two adrenergic receptors, while TTD doesn't list

[1]System developed in the 70s to distinguish the action of cardiovascular drugs based on how they altered inward and outward currents. Molecular studies in the 90s made it possible to link those classes to the action of specific ion channels.

[2]For all ion channels, the alpha subunits are the pore-forming subunits, while the beta subunits are considered to provide structural support. A drug that interferes in the opening and closing of the channel, is deemed more likely to interact with the pore-forming subunits than with the structural subunits, therefore, we have only annotated alpha subunits as drug targets for all ion channels.

[3]The QT interval refers to the time between the start of the Q wave and the end of the T

any individual protein target for dronedarone, but mentions potassium channels as the efficacy target. Finally, DrugBank lists a total of 18 proteins composed of some adrenergic receptors, some sodium and potassium channels and some L-type calcium channels, but all flagged with pharmacologic action unknown, since for the curator their therapeutic role is uncertain.

Fortunately, the majority of drugs are easier to annotate, such as vastatins [1], used in the treatment of hypercholestemia, or loop diuretics, among others.

3.2.1.2 Comparison with existing resources

To evaluate how our drug-target assignments compare with previously published sets, we started by identifying which drugs, whose efficacy target is a human protein, were also present in TTD, DrugBank and in the supplementary material of Rask-Anderson *et al.* We have restricted the comparison to drugs whose efficacy target is a human protein, since it is more reliable to use stable Uniprot IDs to establish the overlap of efficacy targets between resources, rather than rely on the matching of target names. Also, Rask-Anderson *et al* only provides information for drugs with a human target, on that account, it wasn't possible to include drugs assigned to pathogen targets in the comparison.

The drugs were matched based on their database name, as not all the resources provided access to a chemical representation of the drug. Matching drugs based on names, however, is not always trivial, as different databases may follow distinct naming conventions, such as the International Nonproprietary Name (INN) from the WHO, or the United States Adopted Name (USAN) from the American Medical Association[2]. Furthermore, some databases may choose not to strip the salt from the drug name, or, in case of prodrugs, may decide to use the name of the active ingredient instead. Also, when considering biotherapeutics, the naming conventions are not as strict, thus contributing to potential name mismatches. I have tried to account for this issues where sufficient information was available to do so, however despite the effort, not every drug in our list matched a drug in TTD, DrugBank and/or Rask-And. These mismatches may be due to the fact

wave in the heart's electrical cycle.

[1] Usually referred to as statins, despite the fact that HMGCoA reductase inhibitors were originally known as vastatins [102].

[2] For example, Acetaminophen is the USAN name for paracetamol (INN name)

that the drug is really not listed as a drug in the other resources, or simply, the synonyms in ChEMBL are not sufficient to find a match. The Venn diagram in fig. 3.3 provides an illustration to this problem by giving an overview of how many drugs in our map are present in other resources, and how many overlap with each other.

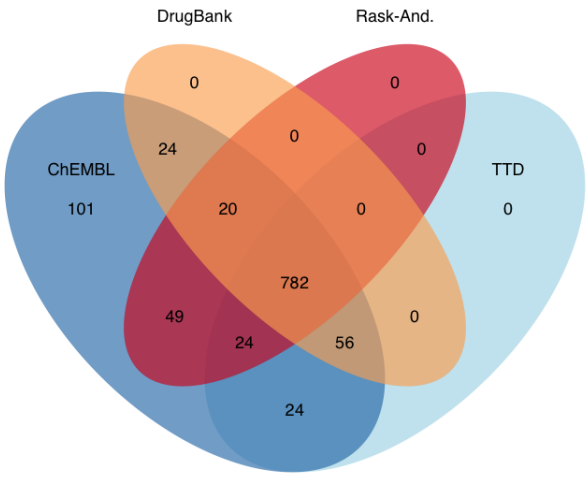


Figure 3.3: Number of drugs with human protein targets in ChEMBL, which are also present in TTD, and/or DrugBank, and/or Rask-Anderson *et al.*

Focusing now on the targets of the 782 drugs present in all the resources, one can notice that 41% of the human proteins flagged as efficacy targets in our set are considered drug targets by all the resources, while 69% are considered drug targets by both DrugBank and Rask-And., and 22% are only considered drug targets by us (fig. 3.4). This additional 83 proteins are a product of our annotation strategy to list all family members or protein subunits when we do not have information on which subset, based on expression or selectivity, is responsible for the therapeutic effect of drugs. Unsurprisingly, TTD is the resource with the lowest number of human proteins flagged as drug targets, since for non-selective drugs, whose efficacy target is a protein complex or ion channel, they do not provide a list of Uniprot IDs for specific subunits or family members.

When looking at individual assignments, and to the overlap between the effi-

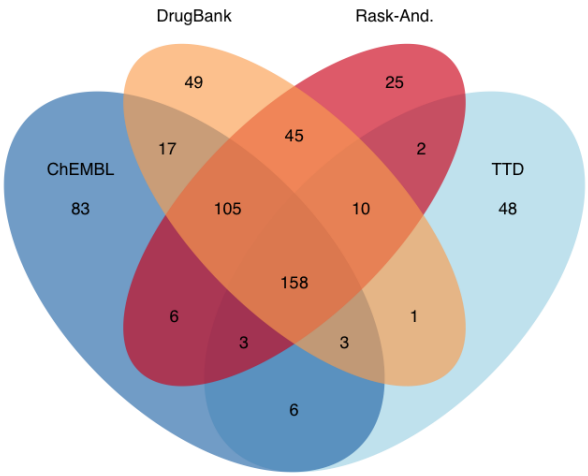


Figure 3.4: Number of human protein targets in ChEMBL, and/or in TTD, and/or in DrugBank and/or in Rask-Anderson *et al*, assigned to the 782 drugs present in all the resources.

cacy targets we have assigned for a common drug with those assigned by others, we observed that for 46% of drugs, all of our targets are the same as the ones provided by DrugBank if we exclude the targets with unknown pharmacological action, 42% of drugs have the same target as the ones flagged by Rask-And., and only 33% of our drugs have the exact same targets as the ones indicated by TTD (fig. 3.5). On the other side of the spectrum, for 11%, 18% and 43% of drugs we completely disagree with the targets assigned by DrugBank, Rask-And. and TTD respectively. Once more, the agreement is particular low for TTD because for non selective drugs, the resource does not provide any protein accessions that we can use for comparison. Additionally, as we have listed all of the subunits or isoforms as efficacy targets in cases where we couldn't identify which was the most important for the therapeutic action of the drug, the many drugs that target ion channels, protein complexes (such as the nicotinic acetylcholine receptors), or protein families (such as carbonic anhydrases), all contribute to lower the agreement scores.

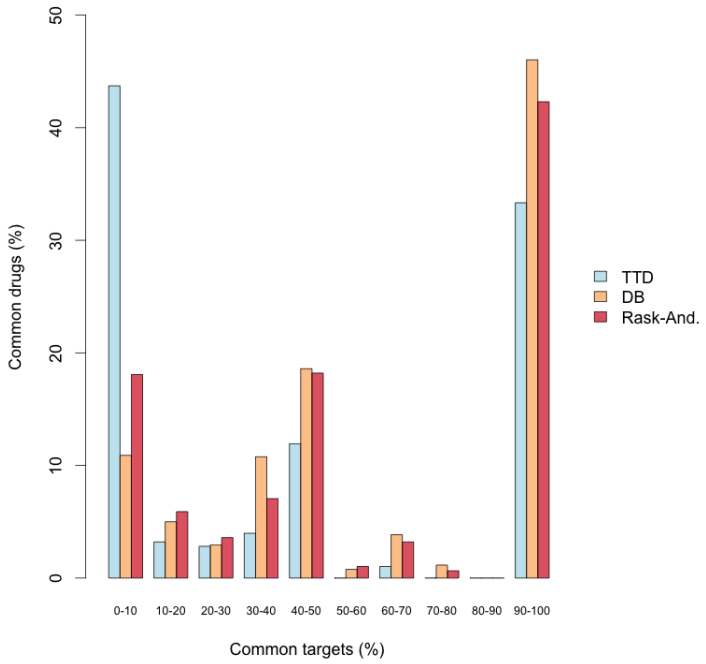


Figure 3.5: Comparison of drug-target associations. Percentage of common drugs between our dataset, DrugBank, Therapeutic target database (TTD) and Rask-Anderson *et al* dataset (Rask-And.), as a function of the percentage of common assigned human protein targets.

3.2.2 A new list of Metabolic and Transporters genes

In order to get a full coverage of all possible genes involved in the absorption, distribution, metabolism and excretion of drugs, as stated in the introduction, I have decided to compile a new list of metabolic enzymes and transporters genes from the gold standard reviews in the field [103; 104; 105; 106; 107; 108]. My effort resulted in a list composed of 253 metabolic enzymes and 76 transporters[1].

The metabolic enzymes are distributed over 73 family classes (fig. 3.6). Unsurprisingly, the largest family is the superfamily of monooxygenases known as cytochrome P450 (CYP), followed by aldehyde dehydrogenases (ALDH), Glutathione transferases (GST) and UDP-glucoronosyltransferases (UGT).

In comparison with the pharmaADME core and extended metabolic enzymes

[1]The complete list of metabolic enzymes can be found in Appendix B in Table 4, and the complete list of transporters in Table 5.

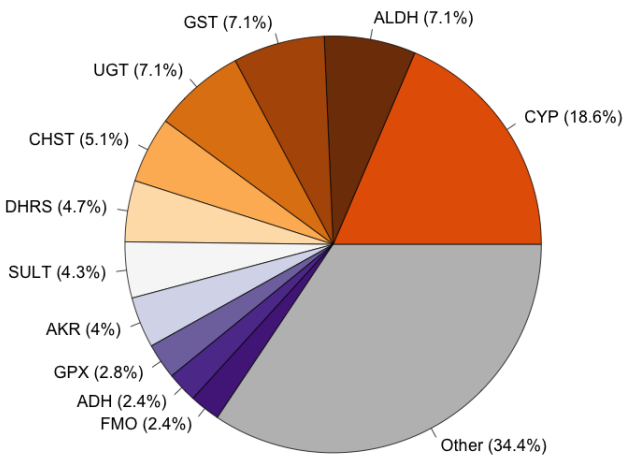


Figure 3.6: Distribution of metabolic enzymes per family classes. CYP - Cytochrome P450; ALDH - Aldehyde dehydrogenase; GST - Glutathione transferase; UGT - UDP-glucoronosyltransferase; CHST - Carbohydrate sulfotransferase; DHRS - Dehydrogenase/reductase SDR family; SULT - Sulfotransferase; AKR - Aldo-keto reductase; GPX - Glutathione peroxidase; ADH - Alcohol dehydrogenase; FMO - Flavin monoxygenase.

gene set, we have flagged 41 additional genes as relevant for drug metabolism. They are mainly members of the Aldo-keto reductase family (AKR), Alcohol dehydrogenase family (ADH), Sulfotransferase family (SULT), and a few secondary members of less prominent families. Nonetheless, I observed an overlap of 212 genes (84%), and when only the core data set is considered, the overlap increases to 98%. Interestingly, pharmaADME has flagged another 10 genes as relevant for metabolism that have been excluded from our list, since they are mostly pseudogenes, such as, CYP2D7P1 and LOC728667.

Contrary to what one can see with the metabolic enzymes, transporters can be grouped in just a handful of family classes (fig. 3.7), being the largest the solute carrier superfamily (SLC) and the ATP-binding cassette sub-family C, also known as the ABC transporter subfamily C. The agreement between my transporters gene set and the pharmaADME gene set is also higher than when considering metabolic enzymes, since all of the transporters in our list are flagged as transporters in the pharmaADME gene set. The only exception is the ABCC13 gene, which is only present in the pharmaADME set, since there was no explicit reference to it in the literature that I have consulted.

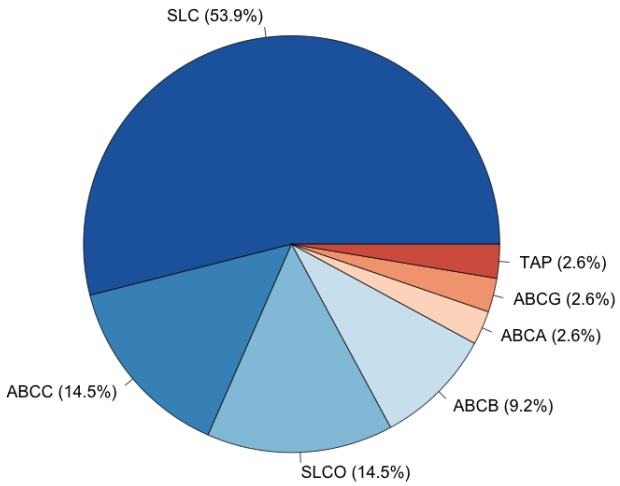


Figure 3.7: Distribution of transporters per family classes. SLC - Solute carrier superfamily; ABCC - ATP-binding cassette sub-family C; SLCO - Solute carrier organic anion transporter family; ABCB - ATP-binding cassette sub-family B; ABCA - ATP-binding cassette sub-family A; ABCG - ATP-binding cassette sub-family G; TAP - Antigen peptide transporter.

3.2.3 Linking drugs to the metabolic and the transporter system

As previously mentioned, given the overall lack of reported data on drug metabolism, and consequently the impossibility to outperform existing resources by manually extracting the metabolic enzymes and transporters for all FDA approved drugs, I have decided to use the information provided by SuperCyp to link drugs to their metabolic and transporter systems. On that account, I have mapped each drug in our set to a compound in SuperCyp, and I have extracted the relevant metabolic links by matching their ATC codes. In this way, I avoided some of the issues discussed in section 3.2.1.2 regarding matching drugs based on their names, as I benefitted from the individual efforts to map internally as many compounds as possible to the ATC code classification.

From the 1378 drugs in our list (1203 small molecules and 175 biotherapeutics), which have at least one ATC code assigned to it, only 418 (30%) match a compound in SuperCyp, despite their 1170 curated drug-entries. There are several

factors contributing to this modest overlap. For instance, not all the compounds in SuperCyp are FDA approved drugs, such as Nitrazepam (a benzodiazepine approved in the UK, Netherlands, Sweden, etc) and barbital (a barbiturate approved in the UK). Also, not all the compounds in SuperCyp are considered drugs by the ChEMBL standards, like glucose, folic acid, citric acid, etc. Most importantly though, for the majority of drugs, especially old ones, there are almost no reported studies about their metabolism, thus a complete map of "drug to metabolic enzymes" and "drug to transporters" is not attainable.

Looking to the metabolic enzymes and transporters assigned to those common 418 drugs, it is clear that the depth of annotation is different between family classes. For example, for members of the cytochrome P450 (CYP) or the solute carrier (SLC) families, it is easy to know for a specific member, which drugs are metabolised or transported by it respectively. However, for members of the Sulfotransferase (SULT) or the Glutathione transferase (GST) families there is no such detailed information, just an indication that the family is involved in the metabolism of that drug. The breakdown of the number of genes per family class, for which there is specific information available, can be found in Table 3.1.

Table 3.1: Prominent families of metabolic enzymes and transporters, their size and the number of genes within the family, for which SuperCyp has information on drug metabolism.

Family classes	**Members**	**Members assigned to drugs**
CYP	47	23
ALDH	18	0
GST	18	All family
UGT	18	All family
CHST	13	0
DHRS	12	0
SULT	11	All family
AKR	10	0
GPX	7	0
ADH	6	0
FMO	6	0
SLC	41	21
ABCC	11	8
SLCO	11	9
ABCB	7	3
ABCA	2	0
ABCG	2	2
TAP	2	0

3.3 Results and discussion

3.3.1 Distribution of drug target classes in the context of age-related differential drug responses

With the new drug map, whose implementation I have just described, I have then checked which are the most common drugs targets associated with drugs, which have shown age-related differential drug responses (DDRs)[1].

3.3.1.1 In paediatric populations

Starting by analysing the familial-class of drug targets responsible for the therapeutic action of drugs that have shown either a lack of efficacy, or more severe side effects, or no difference in response (control group) in paediatric populations when compared with adult populations, I have observed, as illustrated in figure 3.8, that among drugs with safety issues, "transporter" was the class with the highest number of drugs in that category (11), followed by "ion channel" (9), "membrane receptor" (9) and "enzyme" (7). While among the drugs with efficacy issues, the dominant familial-class was membrane receptor (14), followed by transporter (8) and ion channel (7). Lastly, in the control group, the top familial-classes were "membrane receptor" (9) and "enzyme" (7).

Additionally, a pathway enrichment analysis with skypainter, provided by the Reactome database [109; 110], which performs a hypergeometric test and calculates the probability of n or more genes to be involved in a given event purely by chance, showed that drug targets responsible for the therapeutic action of drugs with safety issues in the paediatric population, tend to have a role in developmental and neuronal systems (Table 3.2). While in the efficacy group, only the neuronal system was represented with a certain significance, and in the control group, only the developmental system was significantly enriched.

Furthermore, a closer look to the top drug targets in each DDR category (Table 3.3), revealed that among ion channels, there was a fair representation of

[1]The full list of drugs with DDRs and their drug targets, metabolic enzymes and transporters can be found in Appendix A.

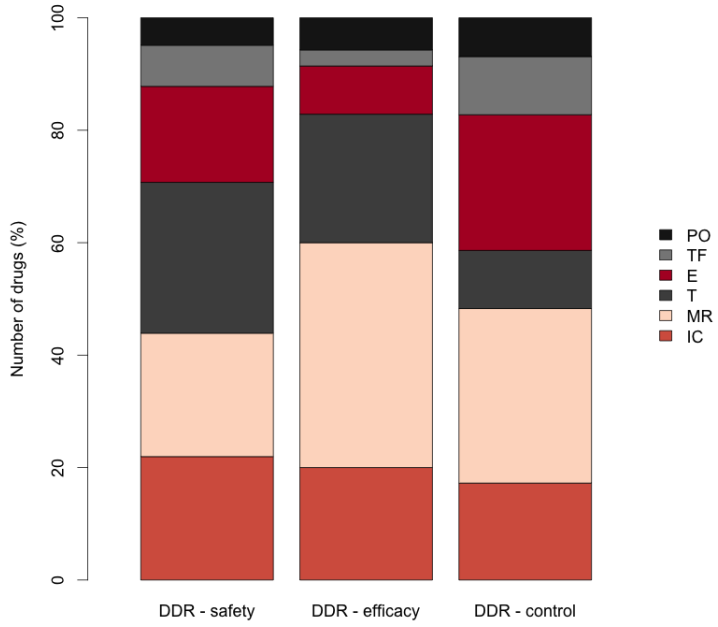


Figure 3.8: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per differential drug responses categories and per drug target familial-classes. IC - Ion channel, MR - Membrane receptor, T - Transporter, E - Enzyme, TF - Transcription factor, PO - Protein other. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.57); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.35); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.58).

GABA A receptors associated with drugs with reported safety and efficacy issues in the paediatric populations, mainly due to the presence of general anaesthetics and antiepileptics in those categories. While among membrane receptors, there was a prevalence of serotonin and dopamine receptors in all three categories. This pattern is due to the presence of antidepressants, antiemetics and drugs used in the treatment of migranes.

In case of transporters, I have also observed a high number of drugs associated with serotonin-related targets, in particular the serotonin transporter (SLC6A4), especially in the safety and efficacy issues groups. In the control group, however,

Table 3.2: Pathway enrichment analysis of the drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations. n number of drug targets (individual proteins) in that pathway; FDR - p-value from the hypergeometric test corrected for multiple testing arising from evaluating the submitted list of genes against every event for the given species currently annotated in Reactome.

	DDR - safety		**DDR - efficacy**		**DDR - control**	
Pathway	**FDR**	***n***	**FDR**	***n***	**FDR**	***n***
Cell Cycle	-	0	-	0	0.12	5
Developmental Biology	0.0031	12	0.19	10	0.035	3
Disease	0.30	6	-	0	0.12	3
Gene Expression	0.85	3	0.70	1	0.85	1
Hemostasis	0.55	5	0.42	3	0.55	1
Immune System	0.37	4	-	0	0.30	2
Metabolism	0.64	9	0.47	54	0.13	7
Metabolism of proteins	0.64	1	-	0	0.13	12
Neuronal System	0.0044	27	0.017	22	0.38	8
Signal Transduction	0.25	32	0.57	22	0.30	17
Transmembrane transport of small molecule	0.57	16	0.42	14	0.57	4

Note: The high number of proteins associated with the Metabolism pathway, among drugs with efficacy issues in paediatric populations, is due to the fact that mitochondrial complex I (NADH dehydrogenase), composed of 50 individual proteins, is the drug target of metformin, a drug used to treat diabetes, which is apparently not effective in children.

the weight of this transporter decreased substantially. This trend can be explained by looking to the literature as already mentioned in chapter 2. Selective serotonin reuptake inhibitors (SSRIs) are one of most studied antidepressant classes in the context of age-related differential responses, due to the fact that recent reviews of clinical trials indicate a higher risk of suicide among children and teenagers when taking SSRIs, and an overall lack of efficacy, being fluoxetine one of the only accepted exceptions.

Regarding the remaining transporters, I have observed that the norepinephrine transporter (SLC2A2) closely followed the tendency of the serotonin transporter, whilst the dopamine transporter (SLC6A3) was more predominant in the control group.

Curiously, among enzymes, no particular drug target was over represented in any category, suggesting that the reason why the drugs show age-related differential drug responses, has more to do with their chemical properties, rather than the type of target they interact with. However, it is important to keep in mind, that the number of drugs in our DDR database, which exert their therapeutic action through enzymes, is rather small, hence any results need to be interpreted with caution. The same is valid for transcription factors, which could just indicate a lack of reported cases in the literature, rather than a true absence of age-related differential drug responses associated with drugs targeting this class. For example, drugs targeting the glucocorticoid receptor have been reported to affect children growth, despite the fact that only two drugs were annotated in the DDR database.

3.3.1.2 In geriatric populations

Performing similar analyses with the drug targets responsible for the therapeutic effect of drugs, which have shown a higher incidence of side effects or lack of efficacy in the elderly, or no difference in response (control group), when compared with an average adult, the first thing I have observed was that the distribution of the familial-classes among the DDR-safety category differs substantially from what I had observed in section 3.3.1.1 (fig. 3.9). In particular, "transporter" was no longer the leading familial-class, having been replaced by "ion channel" (21), and "membrane receptor" (14). Among the drugs with a lack of efficacy, though, the same pattern as seen with the drug targets associated with drugs lacking efficacy in the paediatric populations could be detected, meaning, "membrane receptor" (13) was the key target familial-class, followed by "ion channel" (5) and "transporter" (4). Curiously, the biggest drug target familial-class associated with the control group was "enzyme" (6), followed by "ion channel" (4).

Secondly, a pathway enrichment analysis did not reveal any specific event as a common denominator. With the exception of "hemostasis" in the control group, which is likely to be a false positive, due to the low numbers of drug targets associated with drugs in this category.

Finally, a look to the top drug targets associated with drugs which have shown

Table 3.3: Top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations. *n* - number of drugs associated with that particular drug target.

Familial-class	**DDR - safety**		**DDR - efficacy**		**DDR - control**	
	Gene Symbol	***n***	**Gene Symbol**	***n***	**Gene Symbol**	***n***
Ion channel	GABR*	4	GABR*	2	GRIN*	2
	GLR*	2	CACN*	1	HTR3A	1
	KCNK*	2	CACNA1B	1	KCNH2	1
	other	5	other	3	SCN*	1
Membrane receptor	DRD2	2	HTR1D	3	DRD*	3
	HTR2A	2	DRD2	2	HTR2A	3
	HTR2C	2	HTR1B	2	CHRM3	2
	other	7	other	10	other	8
Transporter	SLC6A4	10	SLC6A4	7	SLC6A2	3
	SLC6A2	5	SLC6A2	3	SLC6A3	1
	SLC6A3	1	SLC12A3	1	SLC6A4	1
	-	-	-	-	SLC18A2	1
Enzyme	ALDH5A1	1	ABAT	1	PTGS2	2
	CSF1R	1	MTND*	1	EGFR	1
	EGFR	1	PTGS2	1	ERBB2	1
	other	5	-	-	other	4
Transcription factor	NR3C1	2	NR1H4	1	VDR	3
	PGR	1	-	-	-	-
	RARB	1	-	-	-	-

GABR* - GABA A receptor; KCNK* - Potassium channel subfamily K members; GLR* - Glycine receptor; CACN* - Voltage-gated calcium channel; MT-ND* - Mitochondrial complex I (NADH dehydrogenase); GRIN* - Glutamate [NMDA] receptor; SCN* - Sodium channel; DRD* - Dopamine type II receptors.

differential drug responses when comparing geriatric and adult populations (Table 3.5), revealed a few resemblances with the analysis in section 3.3.1.1, but also some remarkable differences. Specifically, GABA A receptor was still a dominant drug target among the drugs with safety issues, but no longer present in the DDR-efficacy category. This pattern is due to the over representation of benzodiazepines in the safety category, for which the elderly are known to be more sensitive. Additionally, and contrary to what I had observed before, voltage-gated L-type calcium

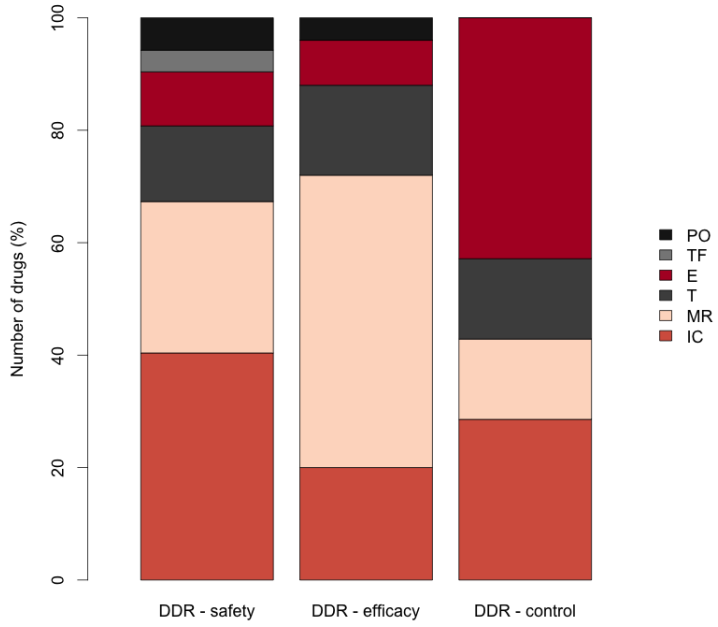


Figure 3.9: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per differential drug responses categories and per drug target familial-classes. IC - Ion channel, MR - Membrane receptor, T - Transporter, E - Enzyme, TF - Transcription factor, PO - Protein other. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.13); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.036); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.29).

channels, and the sulfonylureas receptor were prevalent in the DDR-safety category. The presence of these two drug targets reflect the fact that in the literature both sulfonylureas, drugs used for the treatment of diabetes type II, and drugs to treat angina and/or arrhythmias, are described as not usually safe in the elderly (see chapter 2 for more information).

Among the membrane receptors, compared with the top ones associated with drugs with safety issues in paediatric populations, the prevalence of serotonin receptors was comparably lower, having been replaced by histamine receptors, opioid receptors and muscarinic receptors. Likewise, in the DDR-efficacy group, the pre-

Table 3.4: Pathway enrichment analysis of the drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations. n number of drug targets (individual proteins) in that pathway; FDR - p-value from the hypergeometric test corrected for multiple testing arising from evaluating the submitted list of genes against every event for the given species currently annotated in Reactome.

	DDR - safety		**DDR - efficacy**		**DDR - control**	
Pathway	**FDR**	***n***	**FDR**	***n***	**FDR**	***n***
Developmental Biology	0.43	4	0.37	3	-	0
Gene Expression	0.43	2	-	0	-	0
Hemostasis	0.48	5	0.36	4	0.034	1
Immune System	0.86	1	-	0	-	0
Metabolism	0.56	7	0.39	5	0.034	2
Metabolism of proteins	0.56	1	-	0	0.034	2
Neuronal System	0.43	18	0.37	5	0.23	4
Signal Transduction	0.43	21	0.48	18	0.23	4
Transmembrane transport of small molecule	0.43	26	0.37	7	0.37	1

vious predominant drug target was replaced by the β-adrenergic receptor, since β-adrenergic receptor antagonists are commonly considered ineffective in the elderly, when prescribed to lower the arterial pressure, hence being nowadays replaced by angiotensin-converting-enzyme inhibitors. Consequently, it was not surprising to find that ACE was the leader enzyme in the control group. Lastly, among transporters, there was still a fair representation of serotonin, and norepinephrine transporters, but with a lower number of associated drugs.

3.3.2 Distribution of metabolic enzymes and transporters classes in the context of age-related differential drug responses

In a similar fashion, with the new metabolic enzymes and transporters drug map, whose implementation was described in section 3.2.3, I have also examined which were the most common metabolic enzymes and transporters associated with drugs,

Table 3.5: Top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations. n - number of drugs associated with that particular drug target.

Familial-class	DDR - safety		DDR - efficacy		DDR - control	
	Gene Symbol	n	Gene Symbol	n	Gene Symbol	n
Ion channel	GABR*	8	CHRN*	2	CHRN*	4
	CACNA1*	4	CACNA1*	1	-	-
	ABCC8-KCNJ11	3	HTR3A	1	-	-
	other	6	SCNN1A	1	-	-
Membrane receptor	HRH1	4	ADRB*	4	ADRB3	1
	OPRM1	4	OPRM1	4	HTR2A	1
	CHRM3	2	DRD2	2	HTR2C	1
	other	5	other	5	-	-
Transporter	SLC6A2	3	SLC12A1	2	SLC6A4	2
	SLC6A4	3	SLC6A2	1	SLC6A2	1
	SLC12A1	2	SLC6A4	1	-	-
	ATP1A*	1	SLC29A1	1	-	-
Enzyme	PTGS*	2	ACHE	1	ACE	4
	ATP1A*	1	PDE5A	1	F2	1
	PDE3*	1	-	-	HMGCR	1
	other	2	-	-	-	
Transcription factor	NR3C1	1	-	-	-	-
	PPARG	1	-	-	-	-

GABR* - GABA A receptor; CACNA1* - Voltage-gated L-type calcium channel; ATP1A* - Sodium/potassium-transporting ATPase; PTGS* - Cyclooxygenases; PDE3* - Phosphodiesterase 3; CHRN* - Muscle-type nicotinic acetylcholine receptor; ADRB* - Adrenergic receptor beta.

which have shown age-related differential drug responses (DDRs)[1].

3.3.2.1 In pediatric populations

Once more, starting by analysing the family class distribution of the metabolic enzymes responsible for the metabolism of drugs which have shown either a lack

[1]The full list of drugs with DDRs and their drug targets, metabolic enzymes and transporters can be found in Appendix A

of efficacy, or more severe side effects, in paediatric populations when compared with adult populations, I have observed that cytochrome P450s were the metabolic leaders in every category (fig. 3.10), and that UDP-glucoronosyltransferase (UGTs) were only responsible for the metabolism of drugs with safety or efficacy issues. This last familial-class is normally reported in the literature has having a delayed ontogeny, and this delay is responsible for one of the most known examples of age-related differential drug response, chloramphenicol and the grey baby syndrome (see chapter 2). Hence it is not surprising to see this family represented in these two categories. However, the number of drugs in the DDR database with annotated metabolic information is rather low, therefore, this results need to be interpreted with caution, as small variations in the data can significantly affect the final results.

Additionally, a closer look to the individual cytochrome P450s in each DDR category, revealed CYP3A4 leading all the three DDR categories (Table 3.6). Unfortunately, an equivalent analysis among the UGTs enzymes can not be performed, since, as already mentioned in section 3.2.3, with the exception of cytochrome P450s, SuperCyp does not provide in depth familial annotation. This lack of information reflects what the literature provides, and is a clear indication on how difficult it is to identify the contribution of individual family members in the metabolism of a specific drug.

Table 3.6: Top cytochrome P450s associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations. n - number of drugs associated with that particular metabolic enzyme.

Familial-class	**DDR - safety**		**DDR - efficacy**		**DDR - control**	
	Gene Symbol	n	**Gene Symbol**	n	**Gene Symbol**	n
CYP	CYP3A4	12	CYP3A4	9	CYP3A4	5
	CYP1A2	8	CYP2D6	8	CYP2C9	3
	CYP2D6	8	CYP3A7	5	CYP1A2	2
	other	11	other	8	other	5

Furthermore, focusing on the family distribution of transporters responsible

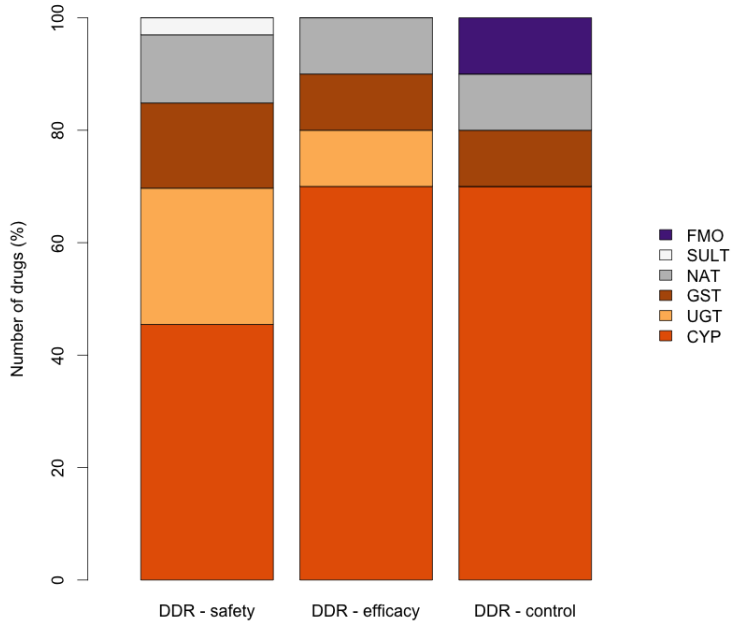


Figure 3.10: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per differential drug responses categories and per metabolic enzyme familial-class. CYP - Cytochrome P450; UGT - UDP-glucoronosyltransferase; GST - Glutathione transferase; NAT - N-acetyltransferase; SULT - Sulfotransferase; FMO - Flavin monoxygenase. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.21); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.79); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.49).

for transport across cells and tissues of drugs which have shown either a lack of efficacy, or more severe side effects, in paediatric populations when compared with adult populations, I observed that solute carrier transporters (SLC) and ABC type C transporters (ABCC) were present in both safety and efficacy issues groups, but absence from the control group (fig. 3.11). However, the numbers behind this plot are small, hence caution is required when interpreting these results, as small variations in the data can significantly alter the final results.

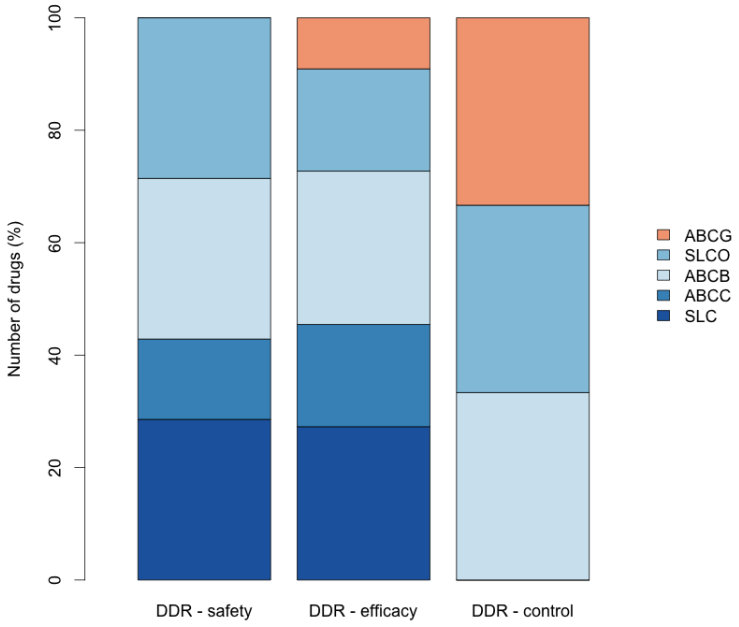


Figure 3.11: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per differential drug responses categories and per transporters familial-class. SLC - Solute carrier superfamily; ABC - ATP-binding cassette family. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.85); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.90); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 1).

3.3.2.2 In geriatric populations

Repeating the same type analysis, but now examining the distribution patterns of metabolic enzymes responsible for the metabolism of drugs which have shown either a lack of efficacy, or more severe side effects, in the elderly, I noticed a similar landscape (fig. 3.12) as before. Once more, cytochrome P450s dominated the three DDR categories. However, UGTs were no longer heavily represented.

Looking to the cytochome P450s distribution in more detail, (Table 3.7), the results were rather similar with what I had observed when analysing the top CYPs

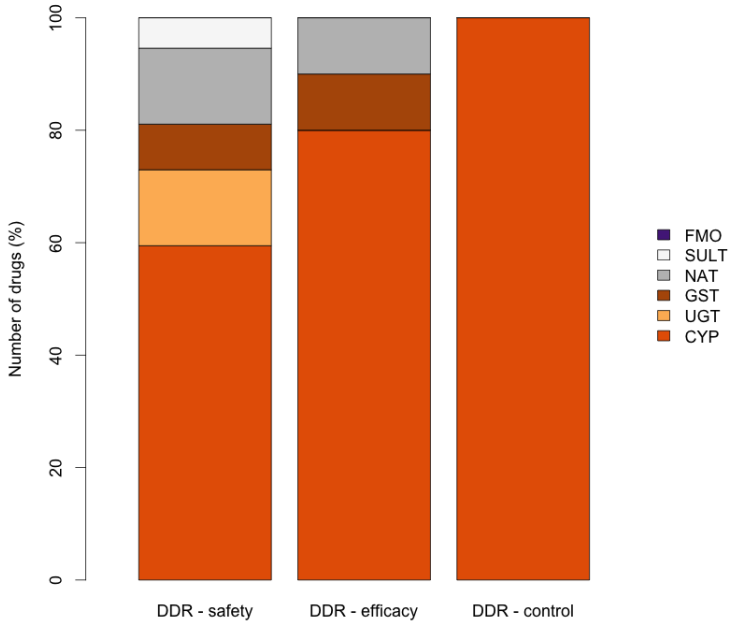


Figure 3.12: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per differential drug responses categories and per metabolic enzyme familial-class. CYP - Cytochrome P450; UGT - UDP-glucoronosyltransferase; GST - Glutathione transferase; NAT - N-acetyltransferase; SULT - Sulfotransferase; FMO - Flavin monoxygenase. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 1); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 1); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.80).

associated with drugs with differential drug response when comparing paediatric and adult populations.

Finally, the family distribution of transporters responsible for the transport of drugs, across cells and tissues, which have shown either a lack of efficacy, or more severe side effects, in the elderly, followed the same pattern as the one seen before (fig. 3.13). Although, in this case the presence of the SLCs and ABCCs in the first two categories was more pronounced.

Table 3.7: Top cytochrome P450s associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations. n - number of drugs associated with that particular metabolic enzyme.

Familial-class	DDR - safety		DDR - efficacy		DDR - control	
	Gene Symbol	n	Gene Symbol	n	Gene Symbol	n
CYP	CYP3A4	13	CYP3A4	5	CYP1A2	1
	CYP2C9	9	CYP1A2	4	CYP3A4	1
	CYP2D6	9	CYP2D6	4	-	-
	other	163	other	6	-	-

3.4 Summary

In conclusion, I have presented an up-to-date and comprehensive list of pharmacologically relevant human genes responsible for the drugs' efficacy and safety, and an up-to-date map of drug targets, metabolic enzymes and transporters against FDA approved drugs. With this information, I explored the distribution of those drug targets, metabolic enzymes and transporters families, in the context of age-related differential drug responses. The overall trend, when focusing on drugs which have shown a lack of efficacy and/or an increased incidence of side effects in children, is a higher incidence of serotonin related targets, and GABA A receptor among drug targets, and UGTs among metabolic enzymes. On the other hand, for drugs which have shown a lack of efficacy and/or an increased incidence of side effects in the elderly, the prevalent drug targets are GABA A receptors, sulfonylureas receptors, voltage-gated L-type calcium channels, histamine receptors, opioid receptors and β-adrenergic receptors, with no noteworthy pattern among metabolic enzymes and transporters. These proteins and protein complexes are thus the ones whose gene expression as a function of age is worth to investigate, in order to verify if based on those gene expression patterns it is possible to rationalise why drugs modulating these targets, should be transferred with caution from one aged population to another.

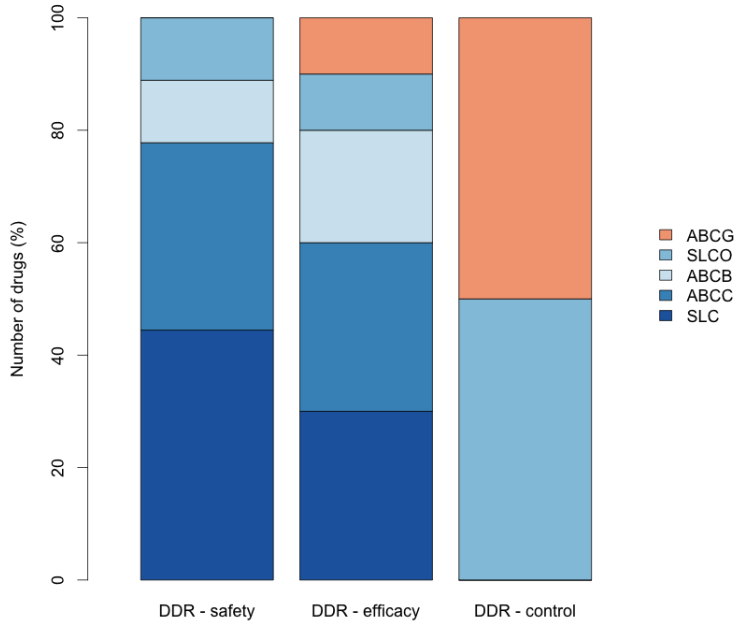


Figure 3.13: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatrics populations when compared with adult populations, per differential drug responses categories and per transporters familial-class. SLC - Solute carrier superfamily; ABC - ATP-binding cassette family. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.11); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.32); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 1).

3.5 Material and Methods

3.5.1 Drug list and target assignments

A list of all FDA-approved drugs (small molecules and biotherapeutics) and WHO antimalarials was compiled based on the data content of ChEMBL 17. Specifically, all the compounds with max phase 4, therapeutic flag 1, and dosed ingredient 1 were selected for this analysis. The list was further divided into small molecules, if the molecule type was equal to Small molecule, and to biotherapeutics if the compound was classified as something else. In total, the list comprises 1,749 dis-

tinct compounds, and after salt stripping, the number of compounds decreases to 1,368 unique small-molecule drugs and 232 unique biotherapeutic agents. For each of these drugs, the efficacy target was extracted from the current version of the prescribing information (available through DailyMed or Drugs@FDA), and complemented with the scientific literature in the cases where, either the prescribing information wasn't available, or the mechanism of action was not reported. The following rules were used as guidelines when assigning the efficacy targets:

1. Identify the target from the Mechanism of Action description in the prescribing information. If it is available assign the therapeutic target(s) to the compound.

2. If the information in the prescribing information is ambiguous complement this with a literature search for publications relating to the mechanism of action of the drug.

3. For conflicting cases, look at review articles, either about the biology of the disease or the pharmacology of other chemical related drugs, and determine the most plausible mode of action.

4. For the cases where several subunits or isoforms match the information described in the prescribing information, evaluate which is more likely to be the real therapeutic target by looking to its expression patterns in relevant tissues for the disease.

5. If the specific subunit(s) or isoform(s) cant be identified assign all of them as targets.

6. For pathogen targets, if the prescribing information lists several microorganisms against which the drug is effective, pick a representative one rather than assigning all as targets.

7. If the mechanism is still not clear or unknown, and the literature did not provided any conclusive information on the molecular target, do not assign any target to the compound.

Out of 1,368 small molecules and 232 biotherapeutics, 1,273 and 224 have now an assigned efficacy target, respectively.

For each drug, the respective ATC code(s) was extracted from the ATC codes in ChEMBL 17, assigned to either its respective dosed ingredient (if applicable) or the parent ingredient itself. For the 1600 drugs, 1378 could be mapped to ATC codes and the remaining 222 were labeled Unclassified.

The protein target classification was retrieved from the existing classification in ChEMBL 17. Both level 1 and 2 were attributed to each of the human and pathogen efficacy targets. For simplicity, all the entries with level 1 labeled as Cytosolic other, Secreted, Structural, and Surface antigen were all renamed to Protein other.

3.5.2 Grouping of efficacy targets

To estimate the actual number of targets flagged as drug targets in the new drug-map, without reflecting a bias in protein subunits or isoforms, efficacy targets were grouped into reasonable molecular entities, such as protein complexes or protein families, when appropriate. The following sets of rules were used for grouping. In our grouping method, a protein target can refer to an individual protein, a protein complex, a protein family or a group of orthologous proteins.

- Human protein targets:

    1. All alpha sodium channels subunits (10) count as one protein target.
    2. All L-type calcium channel subunits (4) and T-type calcium channel subunits (3) count as two protein targets. Other additional calcium channel subunits count as individual targets, since some drugs specifically bind to those subunits. In total, we have assigned 5 protein calcium channel targets, reflecting 5 distinct mechanisms of action.
    3. All alpha potassium channels subunits (40) count as one protein target. However, Kv7.2, Kv7.3, Kv7.4, Kv7.5 and HERG (Kv11.1) are also counted as five individual protein targets, since ezogabine specially acts through the first four channels, and HERG is deemed important for the therapeutic effects of class III antiarrhythmic drugs. Other potassium

channels that are counted as individual drug targets are: the inwardly rectifying potassium channels Kir6.1, due to its specific role in diabetes; the intermediate conductance calcium-activated potassium channel protein 4 (KCa3.1); and five family members of two-P potassium channels (K2P2.1, K2P3.1, K2P9.1, K2P10.1, K2P18.1) for interacting and potentially contributing to the therapeutic action of some general anaesthetics. In total, we have assigned 12 protein potassium channel targets.

4. Other ion channels subunits such as those which compose GABA A receptors (16), glutamate receptors (16), nicotinic acetylcholine receptors (11), and glycine receptors (4) were also clustered into protein complex targets, one for GABA mediated (including all binding sites), three for glutamate mediated, two for acetylcholine mediated and one for glycine mediated channels.
5. All the carbonic anhydrases isoforms (12) count as one single protein family target, as there is no known current drug, which binds selectively to a specific isoform.
6. All the histone deacetylases (10) count as one single protein family target, as there is also no known current drug, which selectively binds to a specific isoform.
7. All the subunits in the mitochondrial respiratory complex I (50) count as one single protein complex target, as we do not know which protein component(s) the drug binds to.
8. The phosphodiesterases are counted by family type, as, for instance, among the phosphodiesterase family 3 we don't reliably know which isoform is the most relevant, therefore the 23 phosphodiesterases are distributed in 11 distinct protein targets.
9. All beta tubulins (8) and alpha tubulins (2), count as two protein family targets, as we don't know of any drug, which acts selectively to a specific isoform.
10. Finally, I have decided to count some kinases as individual targets if the drug has shown some selectivity towards it in detriment of the other

subfamily members as in the case of Ephrin type-A receptor 2, but if I don't know which member of a certain subfamily is the most important for the therapeutic effect of the drug, then the family, rather than the individual proteins, count as the main efficacy target. For instance, Janus kinases (JAK), which function as dimers (e.g. JAK1/JAK2 or JAK1/JAK3), are counted not by their individual protein members: JAK1, JAK2, JAK3 and TYK2, but as a family as one protein target.

- Pathogen protein targets:

  1. For antibiotics, antithelmintics and antifungals, I identified plausible redundant isoforms within the same species, and its respective orthologues, and grouped them into one single higher target class. For instance, β-lactam antibiotics, independent of their activity spectrum, have one single mode of action, to inhibit the enzyme responsible for synthesising the cross-links in the peptidoglycan, known as D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4) or penicillin-binding protein (PBP). In this case, I have grouped all penicillin binding proteins in our list (Escherichia coli, for example, has 7 distinct PBP enzymes), independently of their taxonomy ID, into one single species-independent protein family representing "Bacterial penicillin-binding proteins".

  2. For drugs with more specific indications against pathogens responsible for tuberculosis, malaria or head lice, I grouped efficacy targets into protein targets only if the mechanism has already being described for other non-pathogen-specific drugs, and if the pathogen belongs to the same taxonomy class, as is the case of enoyl-[acyl-carrier-protein] reductase, an efficacy target for ethionamide, an antibiotic indicated to treat tuberculosis, but also for triclosan, a broad-spectrum antibacterial agent.

  3. For drugs targeting viruses such HIV, Herpes, Hepatitis C or influenza, efficacy targets from different strains were grouped.

- Other biomolecules as targets:

1. There are a small number of drugs where the efficacy target is a biomolecule other than a protein. These include cytotoxic drugs, which target human DNA and/or RNA and are indicated for the treatment of neoplastic diseases. These biomolecules count as two targets.
2. In the case of antibiotics that inhibit protein synthesis by targeting the ribosome, the pathogen biomolecules such as DNAs and ribosomal RNAs were grouped into single targets, independently of the species with which they were initially associated. Specifically, in case of antibiotics targeting the 23S or the 16S ribosomal RNA of the ribosome (and potentially some ribosomal proteins), ribosome was assigned as the target and counted as one single target.
3. In case of drugs that act by targeting cell-specific compounds like artemether, which binds to ferriprotoporphyrin IX in plasmodium cells and generates toxic free-radicals ultimately leading to cell death, each cell-specific compound counts as one target.
4. Finally, there is a small subset of drugs that are metal chelators, such as penicillamine or succimer, indicated to treat hypercupremia and lead poisoning, respectively, however, metals have been excluded from the counting system.

3.5.3 Comparison of drug-target assigments

Each of the drug-target associations in the new drug-map were compared with the drug-target associations present in DrugBank 3.0, TTD, and those reported by Rask-Anderson *et al*, if the drug was common between the data sets and the efficacy target was a human protein. Due to the closely related methodology here compared to the analysis of Overington *et al*, few differences were found to this set, excluding new drugs, and where clarification of mechanism of action has appeared since that time. All the data from the external datasets were retrieved in April 2012. The mapping of drugs between the datasets was based on the drug names provided by each dataset and the parent drug name in ChEMBL 17. To improve the mapping, synonyms associated with each parent drug name were also taken into consideration, as well as active drug names in case of prodrugs. The mapping

of targets between the datasets was purely based on UniProt IDs. Out of the 1135 drugs (small-molecules and biotherapeutics) with assigned human protein targets, 880 were mapped to a compound in DrugBank 3.0, 884 to a drug in TTD, and 873 to a drug in the drug list provided by Rask-And.. The percentage of agreement between data sets was calculated per drug by dividing the number of common human protein targets by the number of total unique human protein targets assigned to that drug in both datasets.

3.5.4 Extracting data from SuperCyp

Information on drugs and their respective metabolic and transporter systems was programmatically extracted from SuperCyp. Out of the 1170 compounds available in SuperCyp, structured information was obtained for 886 of them. For each of those compounds their ATC was recorded, as well as, a list of the cytochrome P450s responsible for the compound metabolism, the phase II enzymes that play a role in the compound metabolism, and the list of transporters that play a role in the distribution and excretion of the compound. Subsequently, each compound was mapped to a drug in ChEMBL 17 if their ATC codes were an exact replica. Out of the 886 compounds extracted from SuperCyp only 418 matched a drug in ChEMBL 17.

Chapter 4

Analysis of gene expression datasets as a function of age

Having compiled a list of drugs with reported age-related differential drug responses (DDR) in chapter 2, and the respective list of pharmacologically relevant human genes in chapter 3, the next aim of my project was to gather and analyse suitable gene expression datasets available in public repositories. Only then, one can investigate the ontogeny of genes directly involved in the drugs' efficacy and safety, and thus later attempt to rationalise age-related differential drug responses. Therefore, in this chapter, I describe the work conducted to gather such a list of potential useful human gene expression datasets. Additionally, giving the importance of equally understanding how the ontogeny of pharmacologically relevant genes translates into animal models commonly used in drug discovery, I also describe the work conducted to compile a list of gene expression datasets from macaque, rat and mice. Furthermore, I report on the experimental design of a mouse gene expression dataset generated at EMBL-Monterotondo, designed to complement the data from the public domain. In the end, an overview of the gene expression changes of pharmacologically relevant genes with age, in five primary tissues, and in four species, will be presented and discussed.

4.1 Introduction

To study the ontogeny of pharmacologically relevant genes one needs access to experimental data reporting either, on the levels of the gene's messenger RNA (mRNA), or/and on the levels of the functional gene product (eg. proteins), across development stages, for several disease-relevant tissues. Generating such a complete data set though, would be prohibitively expensive. Fortunately, a wealth of functional data is already available from public archives, which can be reused to analyse gene ontogeny.

Genome-wide transcriptional profiling studies rely on high throughput approaches to measure mRNA levels, which have been around since the mid-90s, such as, serial analysis of gene expression (SAGE) libraries, DNA microarrays, and more recently, RNA high-throughput sequencing (RNA-seq). Conversely, proteomic profiling studies depend on high-throughput protein quantification, which is still in its infancy. Two-dimensional protein gels can, for now, only truly quantify highly abundant proteins, and protein microarrays have yet to reach widespread use due to its inherent difficulty in maintaining the structure and function of immobilised proteins [111]. Additionally, mass spectrometry (MS) based proteomics, despite being quickly shifting from being used to enumerate the protein constituents of a biological sample, to also provide quantitative measurements at a high-throughput scale [112], it is still at a stage where data standards are still being defined for optimal data sharing [113]. As a consequence, the bulk of data available in the public domain belongs to the first category of studies. However, ideally one would have access to both study types, as overall there is only a week positive correlation between mRNA and protein levels [114; 115], thus what we see at the transcription level, may not completely reflect what is happening at the translational level. Nevertheless, genome-wide transcriptional profiling studies have been able to provide insights into a variety of biological processes.

The two main public repositories providing access to such experiments are ArrayExpress [116], and the Gene Expression Omnibus (GEO) [117]. Both of these resources provide access to a plethora of genome-wide transcriptional profiling studies and a few proteomic profiling studies. They are routinely synchronised to insure that the user has always access to all publicly available data, independently

of the initial deposition point.

Besides these two primary public archives, there are a few other added-value databases, which also provide access to gene expression data, such as the MGI-Mouse Gene Expression Database (GXD), with information on embryonic mouse expression, or the 4DXpress database, which focus in cross species comparisons [118]. Although, these added-value resources rarely provide access to data, which have not been deposited in the primary archives already. Therefore, mining the primary archives is usually sufficient to identify potential useful genomic experiments, such as the ones that can be used to investigate the ontogeny of pharmacologically relevant genes.

4.1.1 Data requirements

To study how pharmacologically relevant genes may change their expression patterns with age, the data needs to fulfil certain criteria. Specifically, the samples need to originate from single healthy subjects (not from pooled references), or wild type specimens in case of animal studies, belonging to one of these organisms: *Homo sapiens*, *Macaca mulatta*, *Rattus norvegius* and *Mus musculus*. In this way, a base-line gene expression of pharmacologically relevant genes can be established. Additionally, given that the majority of drugs with reported DDRs when transferred from one aged population to another, exert their therapeutic action by modulating the nervous system, the cardiovascular system and the respiratory system, samples need to be from one of the following tissues: brain, heart or lungs; or from liver and kidney, since these organs are highly responsible for drug metabolism and clearance. Moreover, age and age units should be clearly specified for each sample, and not provided in ranges. Also, the experiment should contain samples which cover at least two developmental stages, such as, childhood and adulthood, or adulthood and old age, or all. Finally, an experiment should contain enough samples per age point, to ensure enough statistical power for differential gene expression analysis.

In the next section, I'll describe how I have queried public repositories, and how I have selected the experiments which fulfilled the above criteria.

4.2 Implementation

4.2.1 Datasets from the public domain

In order to identify gene expression data from public archives to unravel the ontogeny of drug targets, transporters and metabolic enzymes, I have programatically queried ArrayExpress [116] (see section 4.5.1 for a technical description). Since ArrayExpress and GEO [117] are routinely synchronised, it is sufficient to query only one to access the best annotated datasets.

The first query was conducted on the 4th of July 2012. From the 19,508 transcriptomic experiments (microarray or RNA-seq) with human, macaque, rat or mouse samples available in ArrayExpress at the time, 2,308 experiments with information on age and gender were identified as potential relevant for the project (fig. 4.1)[1]. However, after manually verifying the total number of samples, the developmental stages covered, and the disease state of each experiment, the final number of relevant experiments was significantly reduced, with only 96 meeting all the requirements (4.1%).

As every year new data gets deposited in ArrayExpress, it is important to perform periodic queries, in order to identify new and potentially relevant experiments. Therefore, on the 19th of July 2013, I conducted another search in ArrayExpress. This query resulted in the identification of 29 additional relevant experiments (fig. 4.2). Additionally, when comparing this last query with the search conducted in 2012, it was clear that there was no increase in the number of experiments being deposited with information on age and gender, despite the appeal from the community to provide as many annotations as possible [118].

From the 125 experiments selected as relevant, from both queries, only 79 experiments have samples from the five primary organs, namely brain, heart, lungs, liver and kidney. These tissues are the most relevant organs to analyse, in order to study the ontogeny of drug targets, transporters and metabolic enzymes, respon-

[1] I have also queried for gene expression data containing samples from other animal models relevant for drug discovery, such as, *Pan troglodytes* and *Canis lupus*, however there were no datasets that met the requirements described in section 4.1.1

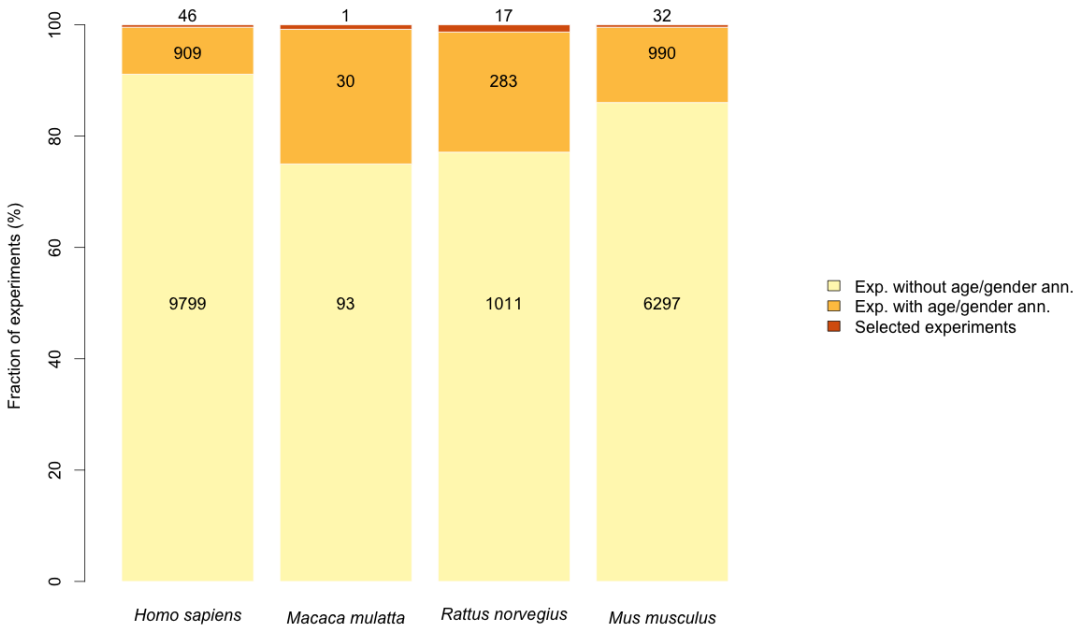


Figure 4.1: Number of transcriptomic experiments (microarray or RNA-seq) from *Homo sapiens*, *Macaca mulatta*, *Rattus norvegicus* and *Mus musculus* with (Exp. with age/gender ann.) and without (Exp. without age/gender ann.) age and gender annotation available in ArrayExpress on the 4th of July 2012. Experiments preliminary selected for gene expression analysis are indicated in orange (Selected experiments).

sible for the therapeutic effect of the majority of drugs with reported age-related differential drug responses, as briefly mentioned in section 4.1.1. Consequently, only those 79 experiments were considered for differential gene expression analysis as a function of age. Table 4.1 provides the number of the final selected experiments broken down per species and tissue. The majority of the experiments containing human data are from brain and liver, and in case of animal models, mouse data is the most complete set, with a high number of experiments with samples also from brain and liver. It is important to mention, however, that one experiment may contain more than one gene expression dataset, for example experiment E-GEOD-25219 contains samples from nine different brain regions, and experiment E-MEXP-1504 has samples from liver, kidney, and lungs. Therefore,

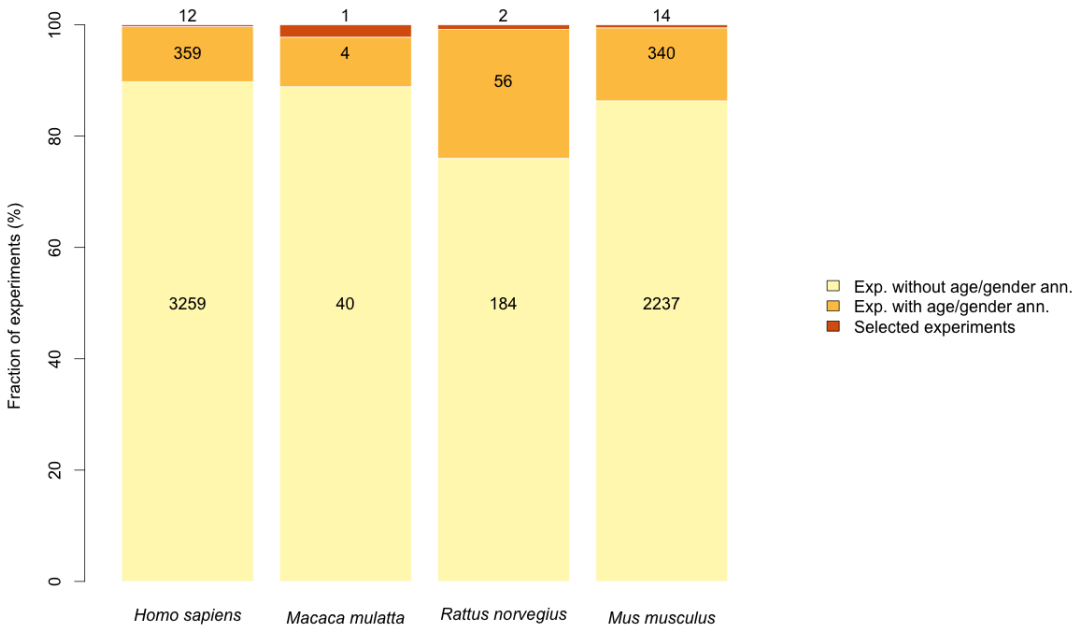


Figure 4.2: Number of new transcriptomic experiments (microarray or RNA-seq) from *Homo sapiens*, *Macaca mulatta*, *Rattus norvegicus* and *Mus musculus* with (Exp. with age/gender ann.) and without (Exp. without age/gender ann.) age and gender annotation available in ArrayExpress on the 19th of July 2013. Experiments preliminary selected for gene expression analysis are indicated in orange (Selected experiments).

the total number of relevant gene expression datasets analysed were actually 91. Additional information on each individual experiment, and expression dataset is provided in Appendix C.

A closer look to those 91 relevant gene expression datasets revealed some interesting patterns, such as, a lack of animal models data covering all developmental stages, i.e. the full life cycle of the organism (fig. 4.3), and a gender bias (fig. 4.4), with a clear over representation of male subjects, specially among experiments with animal models.

Additionally, among the microarray experiments selected for analysis, the majority of samples were hybridised into a platform, which covered at least 60% of all the pharmacologically relevant genes, previously compiled in chapter 3 (fig. 4.5).

Table 4.1: Number of relevant gene expression experiments analysed for differential gene expression as a function of age per species and per tissue.

	Homo sapiens	*Macaca mulatta*	*Rattus norvegius*	*Mus musculus*
Brain	14	1	7	11
Liver	6	1	5	9
Kidney	4	-	3	2
Heart	2	-	1	2
Lungs	8	-	-	3
Blood	14	-	1	1
Other	8	-	2	20

This relatively high coverage, allowed me to explore, and compare between experiments, the ontogeny of the majority of genes involved in the efficacy and safety of FDA approved drugs.

4.2.1.1 Pitfalls

There are, however, a couple of pitfalls with re-using gene expression data from public repositories. The first being the obvious dependency on the experiments available, and the information deposited by the user, such as age, gender and other phenotypic annotations. Specifically, the number of experiments without those two attributes is remarkably high among human experiments (91-89%), and to a certain extend in animal studies too ($\geq$ 70%), which substantially decrease the number of potentially useful experiments to start with. This issue has been recently discussed by Rung and Brazma [118], and hopefully will encourage researchers to deposit data with more phenotypic information in the future.

The second main pitfall with this approach, has a technical origin, and lies in the structure of the public archives. Specifically, the data deposited in these archives, being deposited by the user, does not always obey a rigid structure, nor uses standardised terms to describe the data. This fact has certain implications when one tries to programatically query experiments based on the annotation flags that the user provide. For example, some users have a gender flag entitled 'gender', and others use the word 'sex' instead. Others, only report about the strain or the disease status of the samples deposited in the description of the experiment, but not

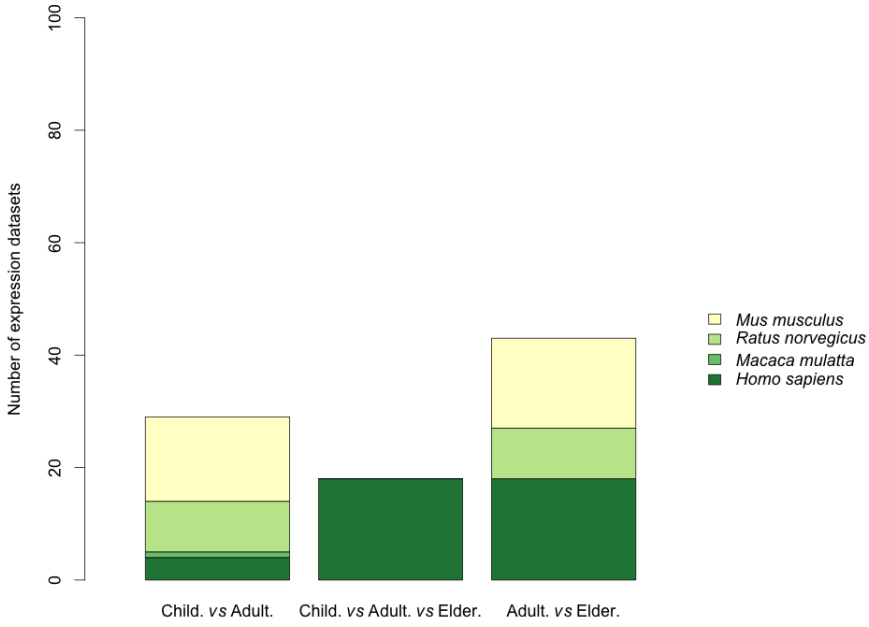


Figure 4.3: Number of relevant gene expression datasets per species covering the following developmental stages: childhood and adulthood (Child. *vs.* Adult.); childhood, adulthood and old age (Child. *vs.* Adult. *vs.* Elder.); and adulthood and old age (Adult. *vs.* Elder.). Human development: childhood from 0 to 18 years; adulthood from 18 to 65 years; and old age 65 years onwards. Macaque development: childhood from 0 to 4 years; adulthood from 4 to 20 years; and old age from 20 years onwards. Rat development: childhood from 0 to 2,5 months; adulthood from 2,5 to 20 months; and old age from 20 months onwards. Mouse development: childhood from 0 to 2 months; adulthood from 2 to 12 months; and old age from 12 months onwards.

in the attributes of the sample itself, thus making it impossible to programatically extract that information. Sometimes, the user decides to split the experiment in two, and deposit the data separately. For example, after manually checking the description of experiment E-GEOD-21716 and E-GEOD-21224, I realised that they were actually deposited by the same user, and that the samples were hybridised at the same time, hence could be combined for analysis. In this case, instead of having two mouse liver expression datasets to analyse, just covering the beginning and the end of the mouse life, when these two experiments are combined, one can actually

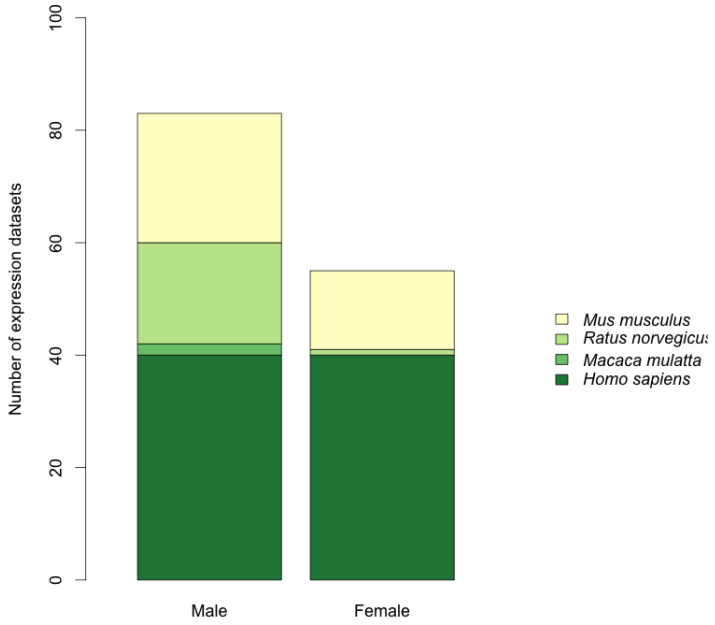


Figure 4.4: Number of relevant gene expression datasets per gender for each species.

get a male liver mouse expression dataset covering the full life cycle of the animal model. However, the link between these two experiments, and potentially others, is not obvious when one is querying for gene expression datasets to re-use for their own large-scale analysis, raising the question to how many more experiments have been deposited under the same circumstances, and thus, how many potential useful experiments may be missed when there is no room for detailed manual inspection of the selected experiments.

Nevertheless, there has been an effort from the teams maintaining these repositories, to address this technical matters, and tackle some of the issues that I've mentioned above. For example, the development of deposition tools such as the MAGE-TAB ArrayExpress submission platform [119], and the development of the experimental factor ontology (EFO) [120], has helped and will continue to improve the standardisation among deposited experiments. Hopefully, this effort will result, in the future, in a reduction of time spent on manually inspecting individual

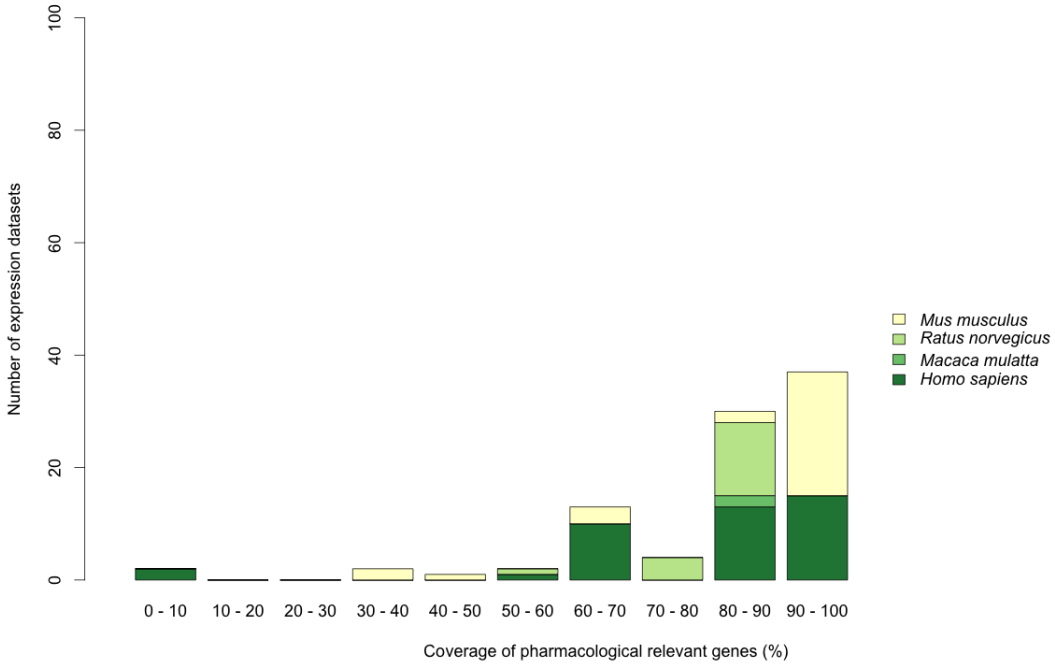


Figure 4.5: Number of relevant gene expression datasets per array coverage of pharmacologically relevant genes for each species.

experiments, when attempting to compile a list of useful expression datasets.

4.2.2 Dataset generated in-house

From the analysis on the developmental stages covered by public gene expression datasets, available to study the ontogeny of pharmacologically relevant genes in human, macaque, rat and mouse, it was clear that there is a lack of expression datasets from animal studies, covering the full life cycle (i.e. childhood, adulthood and old age) (fig. 4.3). Consequently, we decided to generate one ourselves, in order to complement the data from the public domain. Additionally, given that the majority of experiments with animal models in the public domain is done with male specimens (fig. 4.4), we decided to generate an expression set with female specimens instead. Mice was the chosen animal model to perform the experiments, given the fact that it is the easiest to breed, handle, house, and we had access to the appropriated facilities to perform the experiments, in accordance

with the Institutional Animal Care and Use Committee (IACUC) of the European Molecular Biology Laboratory (EMBL, Heidelberg, Germany).

In the next section I will describe the experimental design of this new mouse expression dataset, which I generated at EMBL-Monterotondo, under the supervision of Dr. Cornelius Gross.

4.2.2.1 Experimental design

To design an expression dataset, which covers all the developmental stages of the life cycle of the mouse, it is necessary to understand how the mouse matures. From manuals on mouse husbandry [121], in general, mouse breeders seem to agree that the average life expectancy of a mouse is about 2 years of age, and that mice reach adulthood at around 6 to 8 weeks of age, being then fit for breeding. Additionally, from their experience, a mouse develops 150 times faster than a human being from birth to one month of age; 45 times faster from one to six months; and from six months on, mice age 25 times quicker than humans do. Based on this information, and in order to capture potential transcriptomic switches as the mice progresses from newborn to infant, from infant to teenager, from teenager to adult, and finally from adult to old-adult, we decided to sample six wild-type C57BL/6JCrl female mice at ten different ages, as enumerated in Table 4.2.

Table 4.2: Experimental design of the mouse gene expression dataset generated at EMBL-Monterotondo. The human years were estimated based on the mice maturation rates from figures 20-23 in reference [121].

Mouse age	Human age	Origin
P1 - 1 day old	?	EMBL-Monterotondo
1W - 1 week old	3 years	EMBL-Monterotondo
2W - 2 weeks old	6 years	EMBL-Monterotondo
3W - 3 weeks old	9 years	EMBL-Monterotondo
1M - 1 month old	12.5 years	EMBL-Monterotondo
2M - 2 months old	16 years	EMBL-Monterotondo
3M - 3 months old	20 years	Harlan Laboratories
6M - 6 months old	30 years	Harlan Laboratories
1Y - 1 year old	42 years	Harlan Laboratories
2Y - 2 years old	70 years	Harlan Laboratories

Animals assigned to the young age groups (P1, 1W, 2W, 3W, 1M and 2M) orig-

inated from the in-house colonies and were bred at EMBL-Monterotondo, while animals assigned to old age groups (3M, 6M, 1Y and 2Y), in the interest of time, were purchased from Harlan. Consequently, to minimise effects in gene expression due to potential different housing conditions between the two animal providers (i.e. food, temperature, humidity, etc), all external animals, prior to sacrifice, were allowed to acclimatised to the new environment for one month. The five primary tissues, i.e. brain, heart, lungs, liver and kidney, as defined in section 4.1.1, were then collected and stored at -80°C for each of the 60 females. Due to the high costs associated with the extraction of RNA and hybridisation of all of these samples, only one tissue was selected for further processing at the time. Samples from liver were then selected and processed for RNA extraction, while the remaining tissues were kept frozen. This decision was made based on the fact that, given that liver is a rather homogenous tissue, in comparison, for instance, with brain, the impact of subregion differences in expression patterns as a function of age could be minimised. In addition, as some human metabolic enzymes had been reported as changing their expression with age [51], I was interested in seeing how those switches would be translated in mice. Furthermore, as the metabolic machinery relevant for pharmacology in rodents has considerably expanded through gene duplications, in comparison with the one in humans (see Appendix B - Table 4), I was also curious to investigate if there was any difference in the expression patterns of paralogues as a function of age. The samples were then hybridised into Affymetrix GeneChip Mouse Gene 1.0 ST Array at the gene core facilities in EMBL-Heidleberg, given that, at the time, it was the array with the best coverage of pharmacologically relevant genes (97%). A detailed and technical description of the experimental procedure is available in section 4.5.

It is important to clarify though, that when this experiment was designed, and liver was selected as the first tissue to study the ontogeny of pharmacologically relevant genes in mice, I wasn't yet aware of the fact that experiments E-GEOD-21716 and E-GEOD-21224 could be combined into one single male liver mouse expression set, as already discussed in section 4.2.1.1. Nonetheless, these two sets are complementary, and allow me to compare differences in gender in the context of gene expression as a function of age.

Furthermore, despite the value of the dataset, there is one confounder factor,

which was identified after the completion of the experiment: the stage of the oestrus cycle prior to sacrifice was not annotated. The oestrus cycle is known to influence gene expression [122; 123], for example, Yip *et al* [123] identified a series of genes which are differently expressed in estrus versus proestrus stages, in the mouse uterine transcriptome, including members of the cytochrome P450 family. Therefore, some of the variation in the female mouse liver transcriptome reported in the next sections may be due to differences in the oestrus cycle. It is important to notice, however, that for the majority of pharmacologically relevant genes analysed, the variation in gene expression per age group was minimal. Moreover, females belonging to the older age groups were kept in the same cage, which promotes cycle synchronisation, and consequently, minimises the influence of the oestrus cycle in the overall transcriptome.

Finally, we have recently received more funds to process another batch of samples, and given the high incidence of drugs with reported DDRs which modulate the nervous system, brain was the tissue selected. Therefore, brain samples are currently being processed for RNA extraction and hybridisation at EMBL-Heidelberg. Gene ontogeny analysis from brain samples will, consequently, be later reported in a manuscript, which is under preparation. The gene ontogeny analysis from the liver samples will be presented and discussed in the section bellow.

4.3 Results and discussion

A total of 91 public relevant gene expression datasets were selected for analysis in the context of age-related differential drug responses. In addition, one expression dataset was generated in-house to complement the data in the public domain. All experiments were analysed in the same way: each expression dataset was analysed for differential gene expression as a function of age, in order to study the ontogeny of pharmacologically relevant genes, such as drug targets, transporters and metabolic enzymes, and later attempt to rationalise age-related differential drug responses. Briefly, samples were evaluated for quality, and samples which failed more than one test, were marked as outliers and removed from the downstream analysis. A principal component analysis was then performed to ascertain if samples would naturally cluster per age group, followed by differentially gene

expression analysis as a function of age. A pharmacologically relevant gene was considered to be significantly differentially expressed as a function of age, if the overall absolute fold change was ≥ 1.25 between different age groups, and the corrected p-value for multiple testing was ≤ 0.05. In the end, a list of genes which passed this criteria would be stored per experiment, and used later to compare if signals from different experiments would be reproduced in an independent, but equivalent experiment (i.e. same species, tissue, and age span). A detailed protocol can be found in section 4.5.4. In the next subsections, an overview of the most interesting and reocurring gene expression changes with age, of those pharmacologically relevant genes, with an emphasis on genes associated with drugs with reported DDRs, in brain, heart, lungs, liver, and kidney will be presented per species, together, with a comparison on how those changes translate between species. The implications of these results for drug response in aged populations will be further examined in the general discussion of the thesis.

4.3.1 Analysis of human expression data

Starting with the analysis of human data, 40 expression datasets were analysed, as described in the methods section. Briefly, each experiment was subjected to quality assessment, principal component analysis, gene differential expression as a function of age, and clustering of gene expression profiles as mentioned previsouly, before the data could be collectively analysed. To illustrate this pipeline, and aid in the final interpretation of the results, the main steps will be exemplified using the data from experiment E-GEOD-11512. This experiment consists of 44 human prefrontal cortex samples fairly distributed across the life span of a human being (fig. 4.6).

As already mentioned, the first step of the analysis pipeline was to verify the quality of the samples, i.e. it was important to evaluate if when looking for the distance between arrays, and the distribution of arrays intensities among these 44 samples, no sample was considered an outlier. After this standard quality control, a principal component analysis was then performed, to verify if the samples would naturally separate and cluster by age group. A principal component analysis

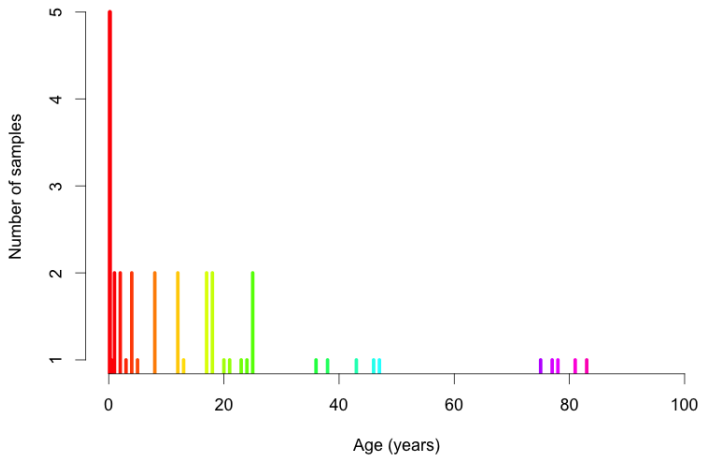


Figure 4.6: Distribution of human prefrontal cortex samples from experiment E-GEOD-11512 per age group.

(PCA) is a mathematical technique that reduces the dimensionality of the data while retaining most of the variation from the original data set [124]. The principle behind PCA is based on the fact that variables which are correlated with each other are, to some extend, measuring the same thing, therefore, by finding components of the original data which are uncorrelated, it is possible to remove some redundancy in the data. Each principal component, is consequently, a linear combination of all the original variables being given a certain weighting. PCA achieves most when the original variables are highly correlated. In case of gene expression data, if the data set is normalised, such that each gene is centred to zero average expression level, the principal components are normalised eigenvectors of the covariance matrix of the genes, and ordered according to how much of the variation present in the data they contain. Each component can then be interpreted as the direction, uncorrelated to previous components, which maximises the variance of the samples when projected onto the component. Samples can then be plotted, making it possible to visually assess similarities and differences between samples and determine whether samples can be grouped [124]. In this particular experiment, principal component 1 (PC1) was found to clearly separate the samples by development stage as can ben seen in figure 4.7.

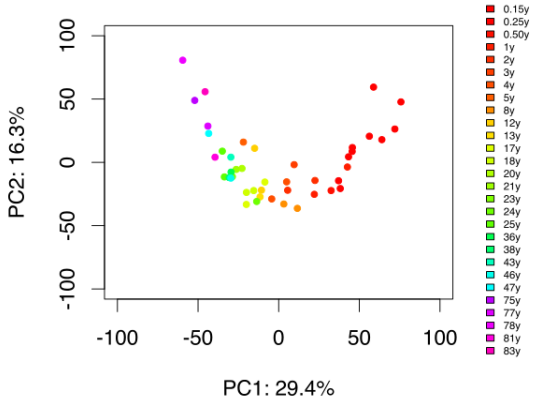
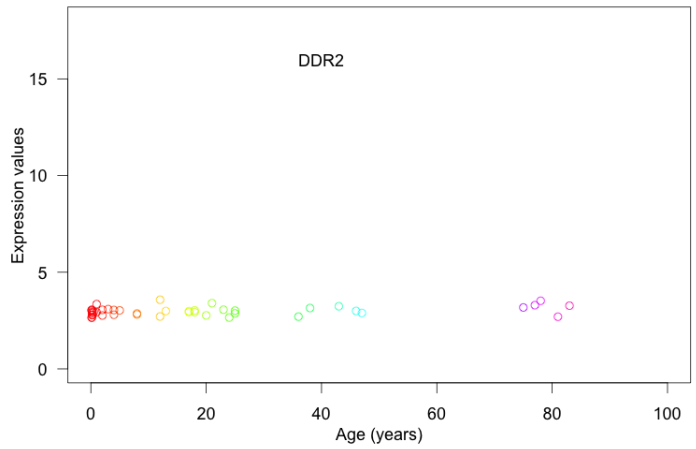


Figure 4.7: Principal component analysis of the human prefrontal cortex samples from experiment E-GEOD-11512. Each dot represents one of the 44 human prefrontal cortex samples in a multidimensional gene expression space projected on the principal plane formed by the first and second principal axes. The dots are coloured according to the age group the sample belongs to as indicated in the legend.

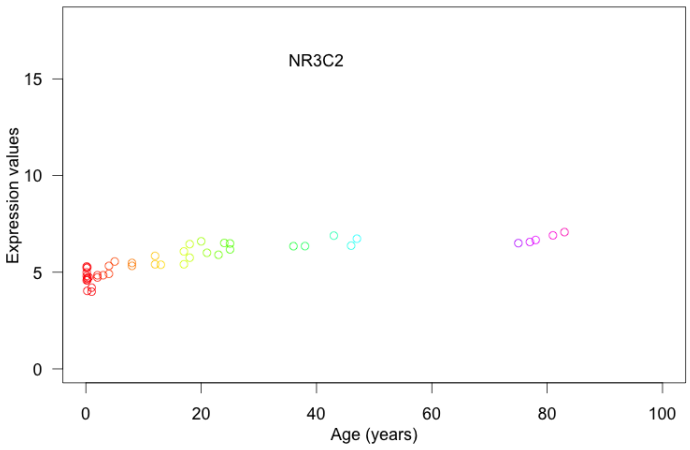
The next step was to investigate if and how the 961 human pharmacologically relevant genes (Drug target (DT) - 654, Metabolic enzyme (ME) - 254, and Transporter (TP) - 89) would alter their expression with age, by fitting a linear model to each gene profile. This linear model was then evaluated for significance, by performing a modified t-test, in order to evaluate if the slope was significantly different from zero. Overall, only a minority of those pharmacologically relevant genes were found to change their expression with age (15%), and the ones which did change, were predominantly drug targets (74%). The individual gene profiles were then finally clustered using the fuzzy c-means algorithm [125] (see section 4.5.4 for more details), which revealed the presence of three main expression patterns: a) no changes in expression with age, 797 genes felt into this category (82.9%); b) the expression increases with age, 85 genes in this category (8.8%); and c) the expression decreases with age, 60 genes in this category (6.2%). A representative gene profile of each of these three categories is depicted in figure 4.8.

With all the pharmacologically relevant genes classified into this three expression categories (increase, decrease, or no changes in expression with age) per experiment, it was then possible to create a map of the expression of those genes in several tissues. Additionally, further linking those gene expression changes to the

(a)



(b)



drugs, via the respective drug targets or metabolic enzymes, or transporters, it was possible to link the expression changes to the ATC code[1]. Thus visualise, and investigate if the anatomical systems where more genes tended to change their expression with age, were indeed connected to more drugs with reported age-related

[1]As already mentioned in chapter 2 and in chapter 3 the Anatomical Therapeutic Chemical Classification System (ATC code) from the WHO [62] was used as a way of obtaining a standard therapeutic indication for each drug.

(c)

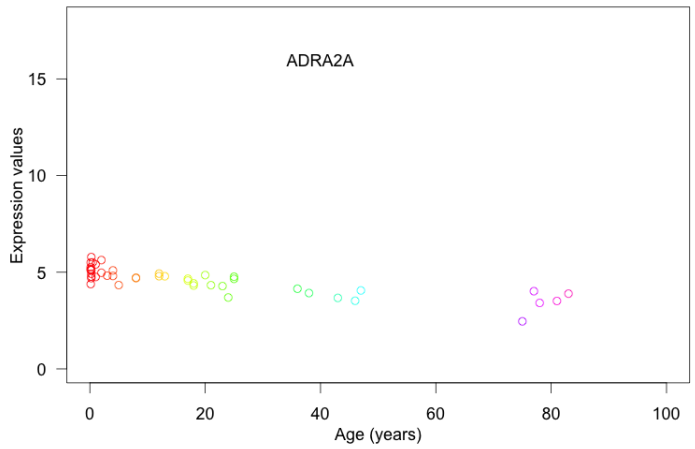


Figure 4.8: Overview of the main expression profiles as a function of age, among the pharmacologically relevant genes in human prefrontal cortex. a) Expression profile of dopamine D2 receptor (DRD2); b) Expression profile of mineralocorticoid receptor (NR3C2) ; c) Expression profile of α-adrenergic receptor 2A (ADRA2A).

differential drug responses (DDR). For Human brain, for example, if one ordered the experiments per brain region, and per age range, and grouped the expression patterns of drug targets per ATC level 4, one could then study how the majority of drug targets alter their expression with age, for a certain group of pharmacological related drugs, which act in a certain anatomical system. In this way, it would then be possible to verify the reproducibility of that signal by checking if the same behaviour was present again in an independent experiment, with different subjects, but in the same subregion, and age range. Figure 4.9 is the end result of establishing those connections. Specifically, the plot indicates whether drug targets increase (in green) or decrease (in red) their expression with age in human brain, and it provides a quick visualisation tool to investigate if those expression patterns are associated: a) with drugs, which have shown either severe and unseen side effects in paediatric or geriatric populations when compared with adult populations (DDR - safety) - the ATC label is depicted in dark blue; b) with drugs, which have shown a lack of efficacy in paediatric or geriatric populations when

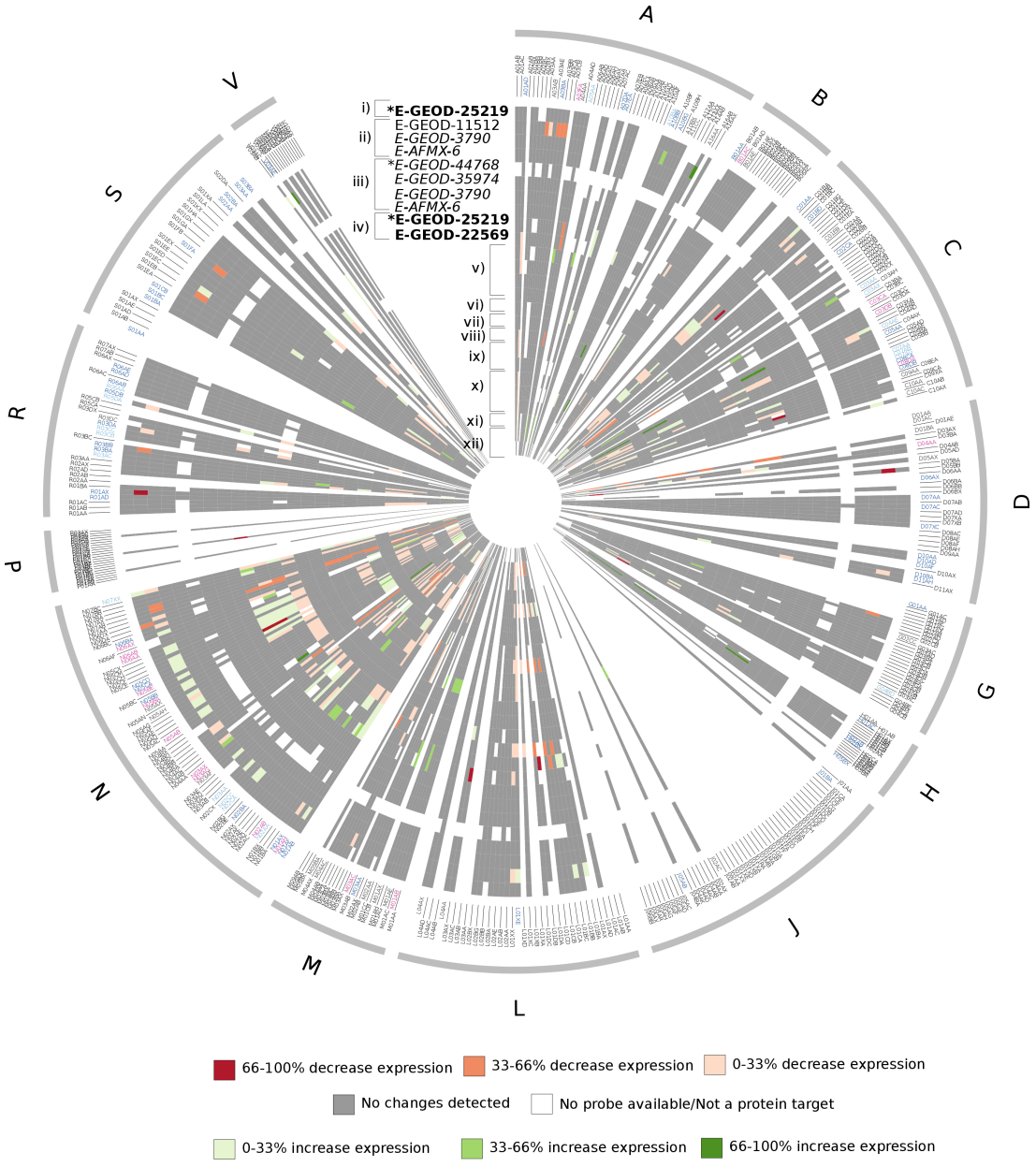
compared with adult populations (DDR - efficacy) - the ATC label is depicted in light blue; c) with drugs, which have reported both safety and efficacy issues when transferred from one aged population to another - the ATC label is depicted in pink; or d) with drugs from the control group, which can be safely and effectively transferred from one aged population to another (DDR - control) - the ATC label is depicted in grey.

A closer look to the plot, specially in category N of the ATC code (i.e. drugs which modulate the nervous system), revealed some interesting signals. For instance, it was clear that among the drug targets in human prefrontal cortex, which change their expression with age, the majority actually tend to decrease their expression with age, as α-adrenergic receptor 2A (ADRA2A) (fig. 4.8c). However, a few drug targets, like the monoamine oxidase B (MAOB), a known drug target for Parkinson's disease [101], linked to drugs in ATC level 4 N04BD, were found to follow the opposite trend, increasing their expression with age in the same subregion. For this particular target, MAOB, this behaviour is in agreement with the literature, where the levels of this protein target have been found to increase with age, suggesting a role in natural age-related cognitive decline [126]. Another clear and strong signal, is the decrease in expression with age, of serotonin 1a receptor (HTR1A), drug target of buspirone (N05BE, an anxiolytic agent, which is not effective in children), in two experiments in cerebral cortex, both covering all pertinent development stages (childhood, adulthood, and old age). The fact that a signal can be reproduced in equivalent experiments, increases the confidence in what we observe.

In terms of brain subregions, there were no significative alterations in the gene expression of drug targets as a function of age in cerebellum, when comparing samples from adults with samples from the elderly.

The same plot was recreated with the expression patterns of metabolic enzymes (fig. 1 in Appendix C), and transporters (fig. 2 in Appendix C). However, since brain is not a tissue specialised in metabolism, it was not surprising that only a minority of metabolic enzymes, were detected to alter their expression with age, like some members of the glutathione S-transferase family (GSTM5, GSTM4, GSTT1), whose expression seemed to increase with age, and sulfatase 1 (SULF1), whose expression seemed to decrease with age in human prefrontal cortex. In terms

(a)


A
B
C
D
G
H
I
L
M
N
P
R
S
V
i) *E-GEOD-25219
ii) E-GEOD-11512
E-GEOD-3790
E-AFMX-6
iii) *E-GEOD-44768
E-GEOD-35974
E-GEOD-3790
E-AFMX-6
iv) *E-GEOD-25219
E-GEOD-22569
v)
vi)
vii)
viii)
ix)
x)
xi)
xii)
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

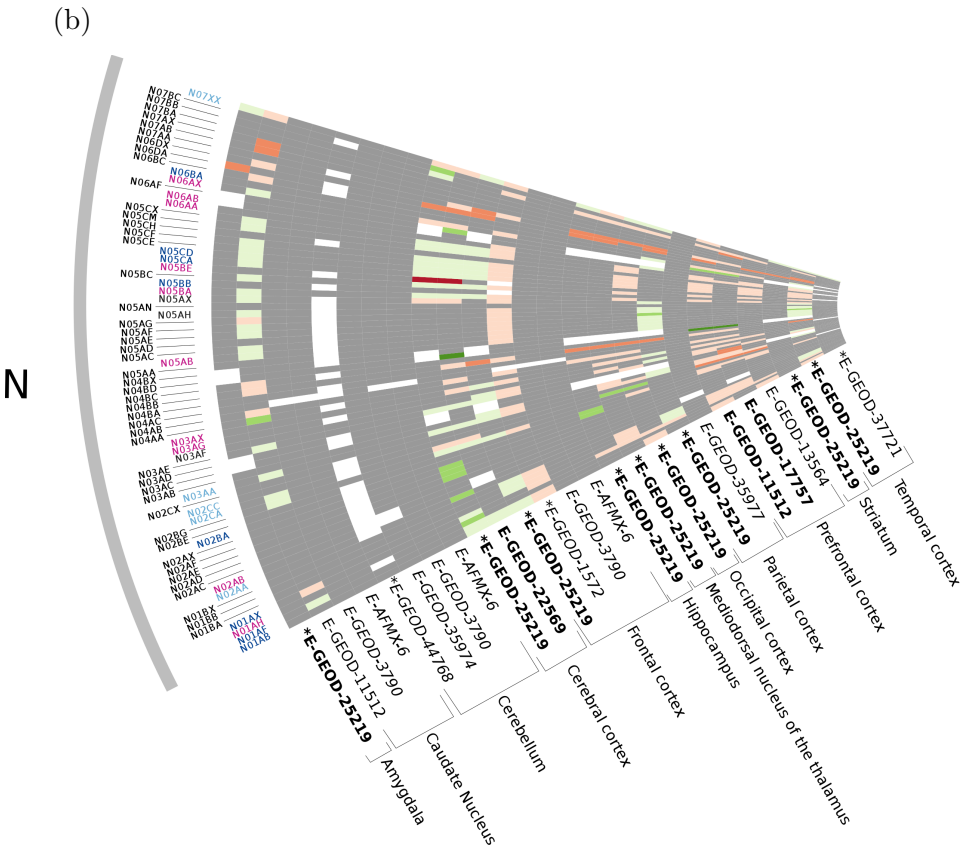


Figure 4.9: Overview of gene expression changes as a function of age of drug targets in (a) human brain per ATC level 4, and (b) in the nervous system category (ATC code 'N'). The outer shell corresponds to the level 1 of the ATC code, scaled to the number of drugs in that category. The inner semi-rings represent the level 4 of the ATC code, coloured by the gene expression variations detected in each individual experiment, and reflect the number of drug targets associated with that ATC level 4, which have decrease its expression with age (depicted in red), or have increased with age (depicted in green). The labels of semi-rings code are coloured in dark blue, if at least one drug in that class has shown either severe and unseen side effects in paediatric or geriatric populations when compared with adult populations (DDR - safety), light blue if at least one drug in that class has shown a lack of efficacy in paediatric or geriatric populations when compared with adult populations (DDR - efficacy), pink if at least one drug has shown both reported safety and efficacy issues when transferred from one aged population to another, and grey if the drug can be safely and effectively transferred from one aged population to another (DDR - control). The experiment accession is depicted in bold if the developmental stages covered by the samples range from childhood to old age, italic if from adulthood to old age, and normal if it only covers from childhood to adulthood. Asterisks (*) indicate expression datasets without raw data. i) Amygdala, ii) Caudate nucleus, iii) Cerebellum, iv) Cerebral cortex, v) Frontal cortex, vi) Hippocampus, vii) Mediodorsal nucleus of the thalamus, viii) Occipital cortex, ix) Parietal cortex, x) Prefrontal cortex, xi) Striatum, xii) Temporal cortex.

of transporters, only solute carrier family 16, member 1 (SLC16A1) was detected to increase with age in human brain.

For the other primary tissues, the same type of analysis and visualisation was employed. Curiously, though, when looking to the expression pattern of pharmacologically relevant genes in human liver, especially metabolic enzymes, no expression as a function of age was actually detected in all the five expression sets available (fig. 4.10, and fig. 3 and fig. 4 in Appendix C). This observation initially seemed to contrast with some reports in the literature. For example, UDP-Glucuronosyltransferase-2B7 (UGT2B7), the enzyme responsible for the metabolism of chloramphenicol, was recently reported to have a delay in ontogeny [6]. Also, the expression of CYP3A4, one of the main metabolic enzymes in adulthood, was reported to be diminished at birth [41]. However, both of these enzymes actually reach adult levels by the age of 6 months to 1 year. Therefore, giving that the switch happens in such a small window, the reason why I was unable to observe these switches may be because, simply, there are not enough samples in the experiments in the public domain at those stages, for one to be able to detect them.

In spite of that, if one looks to the activity levels of some metabolic enzymes, some alterations in activity as a function of age can be detected. For instance, in the experiment from the Human Liver cohort (SAGE) [127; 128], where activity levels were also measured, if one looks to the expression and the activity levels of the CYP2C9 in human liver, one can observe the activity levels increasing with age, despite no significant alterations in the gene expression of that enzyme (fig. 4.11). It is important to note, though, that despite the activity levels measured, the metabolic ability of paediatric patients, regarding this cytochrome, is equivalent to adult levels after 1 year of age, due to the developmental changes in the liver size to body-weight ratio [51].

Likewise, in human kidney, there were almost no changes detected in the expression of pharmacologically relevant genes as a function of age associated with drugs which modulate the genito-urinary system (ATC category G), and among the few which were detected, the signals were not reproduced in equivalent experiments (fig. 5 in Appendix C). This observation seems to be in agreement with what is reported in the literature, i.e. the fact that diuretics are not deemed effec-

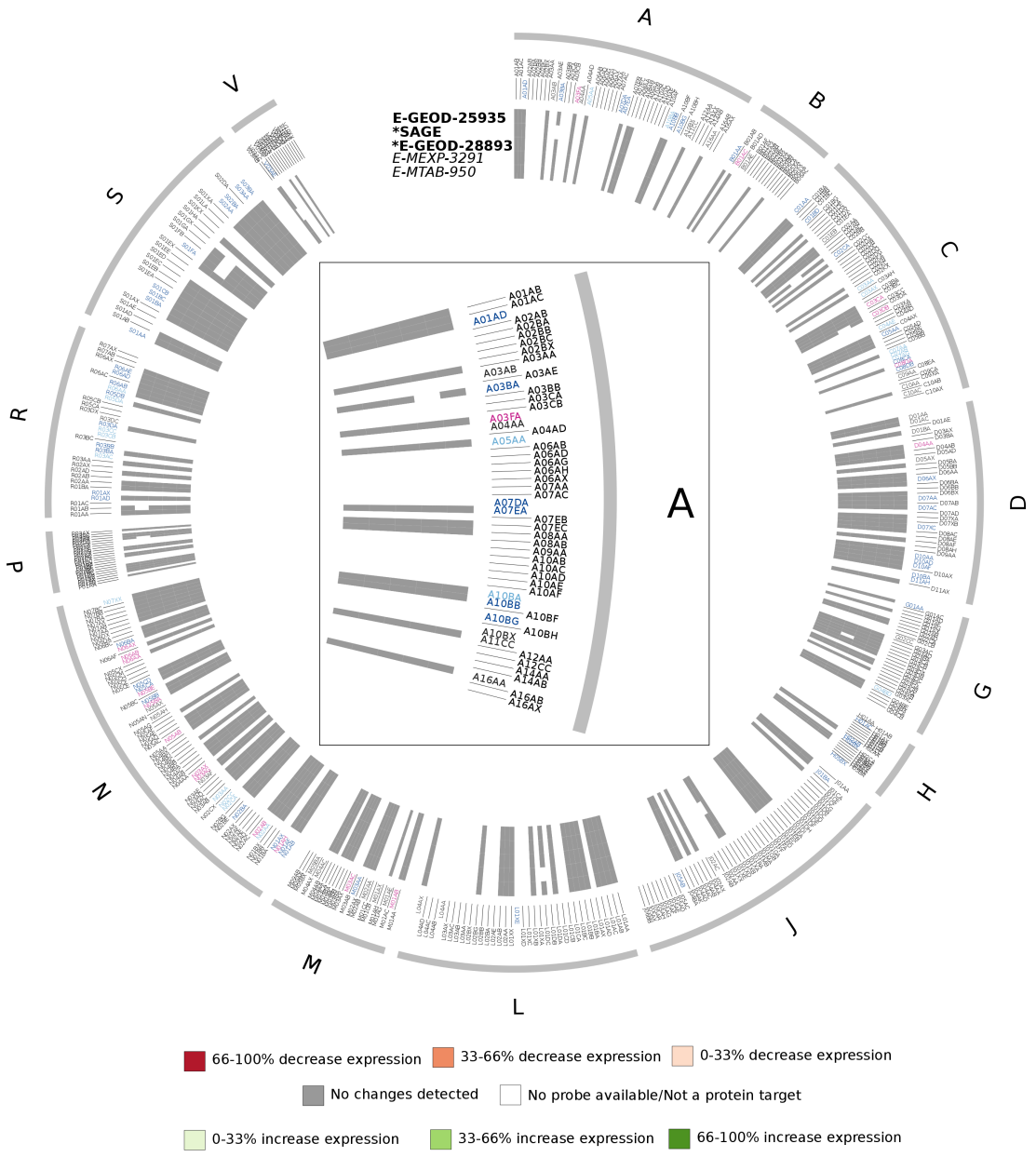
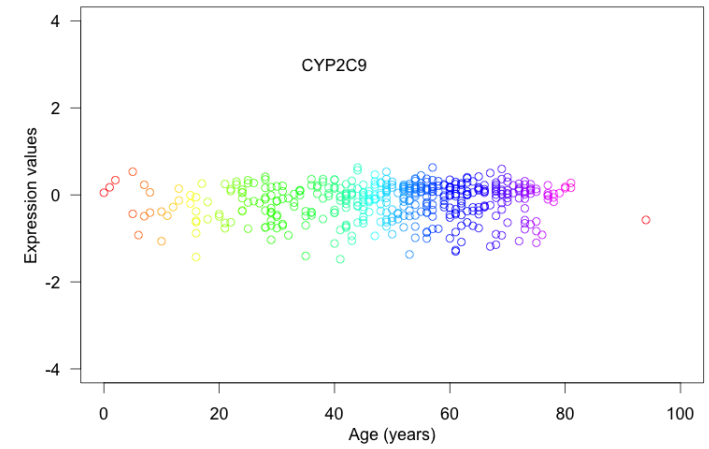


Figure 4.10: Overview of gene expression changes as a function of age of metabolic enzymes in human liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 4.9 for further details.

tive in the elderly for instance, is not because of pharmacodynamic changes in the kidney, but due to the fact that the size of the kidney decreases with age, altering the pharmacokinetics of the drugs [13].

(a)



(b)

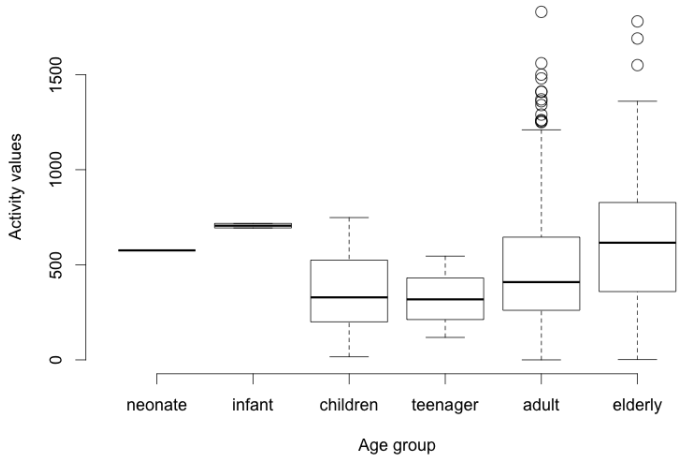


Figure 4.11: Expression and activity values of cytochrome P450 2C9 in human liver (CYP2C9) as measured in the human liver cohort experiment (SAGE). a) Expression of cytochrome P450 2C9 in human liver (CYP2C9); b) Activity of cytochrome P450 2C9 in human liver.

Similarly, in human heart, no alterations in the gene expression of pharmacologically relevant genes were identified (fig. 6 in Appendix C), despite the abundance of drugs with reported DDRs, which target the cardiovascular system (see chap-

ter 2 for more details). For example, β-2 adrenergic receptor inhibitors, which are not considered effective in the elderly for the treatment of hypertension, as mentioned in chapter 2, according with the literature, loose their efficacy because, β-2 adrenergic receptor supposedly decreases its expression with age [13]. This pattern was, however, not detected in any of the human heart experiments available in the public domain, and it is not clear the reason for this apparent discrepancy, besides, the possibility of influence on gene expression by confounder factors, such as non-reported medical comorbidities, or simply differences in study design (i.e. number of subjects per age group).

Finally, in human lungs, some changes in the expression of pharmacologically relevant genes as a function of age were observed, however, they were just detected in one single experiment, which only contained pre-processed data, and were not reproduced in the others (fig. 7 in Appendix C). Consequently the confidence in the signals detected was not very high.

Overall, when looking to the expression changes with age of the top drug targets associated with drugs with reported DDRs in paediatric populations, as defined in chapter 3 (Table 3.3), the trend observed was an overall decrease in expression with age among ion channels, and an increase in expression with age among serotonin receptors in human relevant tissues[1] (Table 4.3).

The full list of expression patterns for these drug targets, their respective subunits, and the frequency of the signal detected is available in Table 11 in Appendix C. It is noteworthy to mention however, that in case of ion channels, and other protein complexes, indicating their overall expression with age is not straightforward. For example, the GABA A receptor (GABR*) is composed of five subunits: two alphas, two betas and one gamma. There are six different alpha subunits, three different beta subunits, and three different gamma subunits to pick from. Alternatively to this configuration, the receptor may be composed of the delta, or theta subunit instead of the gamma subunit. When looking to

[1]The expression of drug targets associated with drugs in the N category of the ATC code was extracted from the human brain experiments; the expression of drug targets associated with drugs in the C category of the ATC code was extracted from the human heart experiments; the expression of drug targets associated with drugs in the R category of the ATC code was extracted from the human lungs experiments; and the expression of drug targets associated with drugs in the G category of the ATC code was extracted from the human kidney experiments.

Table 4.3: Predominant expression patterns of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, in the respective relevant human tissues. See Table 11 in Appendix C for the full detailed expression patterns per subunit and tissue region.

	DDR - safety		DDR - efficacy		DDR - control	
Family class	**Gene Symbol**	**Exp.**	**Gene Symbol**	**Exp.**	**Gene Symbol**	**Exp.**
Ion channel	GABR*	↓	GABR*	↓	GRIN*	↓
	GLR*	↑	CACN*	↓	HTR3A	-
	KCNK*	↓	CACNA1B	-	KCNH2	-
	other		other		SCN*	↑ & ↓
Membrane receptor	DRD2	-	HTR1D	-	DRD*	-
	HTR2A	↑	DRD2	-	HTR2A	↑
	HTR2C	↑	HTR1B	↑	CHRM3	-
	other		other		other	
Transporter	SLC6A4	-	SLC6A4	-	SLC6A2	-
	SLC6A2	-	SLC6A2	-	SLC6A3	-
	SLC6A3	-	SLC12A3	-	SLC6A4	-
					SLC18A2	-
Enzyme	ALDH5A1	-	ABAT	-	PTGS2	-
	CSF1R	?	MTND*	-	EGFR	?
	EGFR	?	PTGS2	?	ERBB2	?
	other				other	
Transcription factor	NR3C1	↓	NR1H4	-	VDR	-
	PGR	↑				
	RARB					

GABR* - GABA A receptor; KCNK* - Potassium channel subfamily K members; GLR* - Glycine receptor; CACN* - Voltage-gated calcium channel; MT-ND* - Mitochondrial complex I (NADH dehydrogenase); GRIN* - Glutamate [NMDA] receptor; SCN* - Sodium channel; DRD* - Dopamine type II receptors; ↑ - Increase expression with age; ↓ - Decrease expression with age; '-' - No change in expression with age detected; ? - No relevant tissue to analyse gene expression.

their individual expression with age, though, it was clear that both GABRA5 and GABRG1 decrease with age in more than one experiment and in more than one brain region, but curiously GABRG2 increased with age, for instance. Overall,

there seemed to be more subunits whose expression decreased with age, and thus, I decided to consider that the GABA A receptor in human brain decreased its expression with age. However, it is important to reflect that drugs with different affinities to the different GABA A receptor subunits (and different GABA A receptor compositions) may have slightly different responses with age.

The sodium channel (SCN*) is also a very interesting case, as well as the Glutamate [NMDA] receptor (GRIN*). The first one, is composed of three subunits, one alpha, the pore subunit, and the subunit where drugs bind, and two structural beta subunits. There are ten different alpha subunits which can form the channel, and when looking to the individual expression profile of each one, it was interesting to observe that SCN1A increased its expression with age, in almost as many experiments, as SCN3A decreased its expression with age, while the other eight subunits did not significantly alter their expression with age (Table 11 in Appendix C). In case of the glutamate receptor, GRIN2A and GRIN3A, clearly also seemed to decrease their expression with age in several experiments and regions, but curiously, the drugs which they are associated with in the context of this project, belong to the DDR-control group, where no difference in drug response was observed between different aged populations. Hence, it would be interesting to study the affinity of these drugs, which modulate this receptor, per subunit, and investigate if the drugs have a preference for the subunits which do not alter their expression with age.

Moving from the ion channels to the membrane receptor family class, it was interesting to observe some serotonin receptors increasing their expression with age, but no alterations detected among the dopamine receptors. In terms of transporters, there was not a single alteration detected in their expression as a function of age. This included the serotonin transporter (SLC6A4), target for the selective serotonin receptor inhibitors (SSRIs), which seems to suggest that the reason why SSRIs are in the list of drugs with reported safety and efficacy issues when transferred from adult to paediatric populations is not related with pharmacodynamic changes of their drug target. In fact, as already discussed in chapter 2, one of the working hypothesis for the reason why SSRIs don't seem to be effective in teenagers, is actually because, in clinical trials for the treatment of depression, the placebo response is high [70], making it very difficult to get any statistically

significant improvement when a drug is given, despite the trends reported in that direction.

Among enzymes, there are a few drug targets associated with drugs with reported DDRs for the treatment of cancer, for which there was no relevant tissue to analyse their gene expression as a function of age. Finally, among transcription factors, it was interesting to observe the glucocorticoid receptor (NR3C1) decreasing its expression with age in lungs, however, the implications of this are not yet clear. In particular, drugs for the treatment of asthma in the "DDR - safety" category, whose drug target is the glucocorticoid receptor, are in this category because they impair children growth. This side effect, however, seems to be intrinsically linked to the receptor itself, given that agonists of glucocorticoid receptor promote antiangiogenic effects [129], i.e., there is a decrease in the growth of blood vessels in the surrounding tissues, which may impair organ maturation, and consequently affect children growth. Paradoxically, though, glucocorticoids are also prescribed to premature babies in order to promote lung maturation [130]. Therefore, hypothesising about the effect of this receptor decreasing its expression with age in the context of this project is not straightforward.

When focusing on the expression patterns of the top drug targets of drugs with reported DDRs in geriatric populations, it was interesting to observe less drug targets changing their expression as a function of age, with GABA A receptor being one of the few exceptions (Table 4.4). This behaviour seems to potentially explain why benzodiazepines, when given to the elderly for the treatment of anxiety, if not dosed correctly may end up over inhibiting the GABA system and causing several side effects. In addition, it was also interesting to observe Histamine H1 receptor decreasing its expression with age, given the reported increase in sensitivity in the elderly with this type of medication. However, the side effects are normally attributed to high anticholinergic properties of the first generation antihistamines [38; 40], and not directly linked to the interaction with the histamine receptor itself.

In terms of metabolic enzymes, as already mentioned when describing the overall expression patterns of pharmacologically relevant genes in human liver, no significant changes as a function of age were detected for the top metabolic enzymes (see Table 3.6 and Table 3.7 in chapter 3) associated with drugs in all three DDR

categories in paediatric and in geriatric populations.

Table 4.4: Predominant expression patterns of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, in the respective relevant human tissues. See Table 12 in Appendix C for the full detailed expression patterns per subunit and tissue region.

	DDR - safety		DDR - efficacy		DDR - control	
Family class	**Gene Symbol**	**Exp.**	**Gene Symbol**	**Exp.**	**Gene Symbol**	**Exp.**
Ion channel	GABR*	↓	CHRN*	?	CHRN*	?
	CACNA1*	-	CACNA1*	-		
	ABCC8-KCNJ11	-	HTR3A	-		
	other		SCNN1A	-		
Membrane receptor	HRH1	↓	ADRB*	-	ADRB3	-
	OPRM1	-	OPRM1	-	HTR2A	↑
	CHRM3	↑	DRD2	-	HTR2C	↑
	other		other			
Transporter	SLC6A2	-	SLC12A1	-	SLC6A4	-
	SLC6A4	-	SLC6A2	-	SLC6A2	-
	SLC12A1	-	SLC6A4	-		
	ATP1A*	-	SLC29A1	-		
Enzyme	PTGS*	↑	ACHE	?	ACE	-
	ATP1A*	-	PDE5A	-	F2	-
	PDE3*	↑			HMGCR	-
	other					
Transcription factor	NR3C1	-				
	PPARG	-				

GABR* - GABA A receptor; CACNA1* - Voltage-gated L-type calcium channel; ATP1A* - Sodium/potassium-transporting ATPase; PTGS* - Cyclooxygenases; PDE3* - Phosphodiesterase 3; CHRN* - Muscle-type nicotinic acetylcholine receptor; ADRB* - Adrenergic receptor beta; ↑ - Increase expression with age; ↓ - Decrease expression with age; '-' - No change in expression with age detected; ? - No relevant tissue to analyse gene expression.

4.3.2 Analysis of macaque expression data

Repeating a similar analysis with the primate data, would in principle allow me to verify how the changes in expression as a function of age of pharmacologically relevant genes, previously identified in the human data, translate to one of the animal models closer to humans in evolutionary terms. As mention in the implementation, though, there were only two expression datasets from *Macaca mulatta* available in the public archives, which met the necessary requirements to be analysed in the context of age-related differential drug response. One expression set contained samples from hippocampus, and the other samples from liver, and both contain data ranging from childhood to adulthood only.

In the macaque hippocampus, no single alteration in the expression of drug targets, metabolic enzymes and transporters as a function of age was detected (fig. 8 in Appendix C). However, as it was the result of the analysis of just one experiment, from one single brain region, it is not possible to conclude if the changes in gene expression detected in the human brain as a function age are actually present or absent in macaque brain. It is also important to note, that in the only human expression dataset from hippocampus, not many alterations were identified either (fig. 4.9).

In the liver expression set, though, there were some pharmacologically relevant genes increasing their expression with age, mostly metabolic enzymes (fig. 9 in Appendix C), but none of them were orthologues of the top metabolic enzymes associated with drugs with reported DDRs (see Table 3.6 and Table 3.7 in chapter 3). In fact, among those, it was only possible to study the expression profile as a function of age in macaque liver of the orthologue of cytochrome P450 1A2 (CYP1A2), given that none of the others (cytochrome P450 3A4 (CYP3A4), cytochrome P450 2D6 (CYP2D6), and cytochrome P450 2C9 (CYP2C9)) had probes in the microarray used, i.e. it wasn't possible to study the expression of those genes with the Affymetrix GeneChip Rhesus Macaque Genome Array.

4.3.3 Analysis of rat expression data

Repeating the same approach to analyse the expression data from rodents, is also useful, as it provides insight into how certain gene expression of pharmacologically

relevant genes translate to more common animal models, which are also much easier to handle, and to use in a pre-clinical testing setting (i.e. require less compound for testing). Starting with rat data, there were 18 expression datasets identified as containing data with information on age, which met the necessary requirements for this type of analysis. From the five primary tissues, only data from lungs is not available. Nevertheless, it is a set of useful experiments to compare with the human data.

Overall, in rat brain, just a few expression datasets showed some alterations of gene expression of pharmacologically relevant genes with age (fig. 10 in Appendix C). In particular, one expression dataset from rat cerebral cortex seemed to have genes whose expression significantly increased or decreased with age. However, the data from this experiment (E-GEOD-34272) was not generated with microarray technology, but with RNA high-throughput sequencing (RNA-seq), and from my experience in analysing this type of data, it is much more sensitive and subjective to variations in the expression levels, and thus it is not surprising that stronger signals were identified. However, caution is necessary when interpreting these signals, specially in comparison with signals from microarray experiments, which are, in my experience, much more robust to experimental noise and artifacts.

In rat liver, on the other hand, there were also significative changes detected in the expression of drug targets, metabolic enzymes and transporters, specially in metabolic enzymes, which tended to decrease with age (fig. 11 in Appendix C).

In rat kidney, not many alterations with age were detected with pharmacologically relevant genes associated with drugs in the category G of the ATC code, with the exception of a few metabolic enzymes, which tended to increase with age (fig. 12 in Appendix C). It is important to keep in mind though, that the urinary system is intrinsically linked with the regulation of the cardiovascular system. For example, diuretics, whose target is the amiloride-sensitive sodium channel (SCNN1A) in the kidney, are drugs prescribed to decrease the arterial pressure. Thus, for some drug targets associated with drugs in the category C of the ATC code, the expression in the kidneys was also evaluated, and significative changes were detected, which will be further discussed in the next paragraphs.

Finally, in rat heart, only a few drug targets were detected to change their expression with age, but none associated with drugs with reported DDRs (fig. 13

in Appendix C).

Overall, there were not many strong signals detected compared with the human data, i.e., signals which were reproduced in more than one rat experiment from the same tissue and subregion. One factor, which may be contributing to this lack of reproducibility is the fact that different strains and gender of rats were used in this 18 expression datasets. Additionally, the age range was also not always comparable, even among the experiments which have samples ranging from childhood to adulthood.

When looking to the expression of orthologues of the top drug targets associated with drugs with reported DDRs in paediatrics, in relevant rat tissues, the main difference when comparing with the human expression patterns (see Table 4.3 above), was the fact that GABA A receptor in rat brain has a tendency to increase, rather than decrease its expression as a function of age. Curiously, the sodium channel in rat brain has a predominant number of subunits whose expression decreases with age (Table 13 in Appendix C). Additionally, some of the subunits of the mitochondrial complex I were also detected to decrease their expression with age, but only in one of the five liver experiments, therefore, the signal wasn't considered very significant. Finally, it was also interesting to observed, that the majority of transcription factors decreased their expression with age in rats, which is in contrast with the absence of alterations in humans.

Doing the same comparison with the orthologues of the top drug targets, associated with drugs with reported DDRs in geriatric populations, in relevant rat tissues, the main difference detected when comparing with the human expression patterns (see Table 4.4 above), was that in rat, the sulfonylureas target (ABCC8-KCNJ) was found to decrease its expression with age in rat liver, while in human liver no alteration in expression was detected (Table 14 in Appendix C). In addition, the amiloride-sensitive sodium channel (SCNN1A), drug target for diuretics, which are no longer deemed effective to lower the arterial pressure in the elderly, was found to increase its expression with age in rat kidney, contrary to what was observed in the human data. Finally, the drug target of statins, 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR), seemed to decrease its expression in rat liver.

4.3.4 Analysis of mouse expression data

Moving on to the analysis of the mouse expression data, 31 expression datasets were identified from the public archives, which met the data requirements. Among the data from animal models, it was the species with the most complete data set, having experiments from all five primary tissues, thus making a full comparison with the human data possible. Before presenting the overall expression of pharmacologically relevant genes in mouse tissues, I'll present in detail the main results of the female liver expression data set which was generated at EMBL-Monterotondo.

First, it was interesting to observe the significant increase in the weight of liver between 1 month and 2 months of age, indicating clearly the switch during puberty (fig. 4.12).

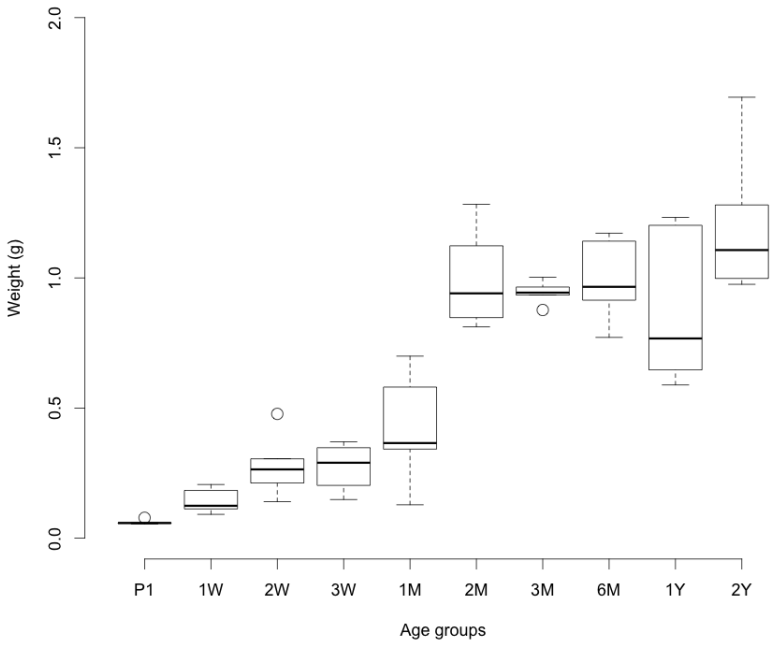


Figure 4.12: Distribution of the liver weight per age group as enumerated in Table 4.2.

Second, similarly to the analysis pipeline which I've employed to analyse the public gene expression datasets, I performed a principal component analysis, to

see if the samples would naturally cluster per age groups, after checking that all the samples passed the necessary quality controls. From figure 4.13, it was clear that principal component 1 (PC1) separated the samples per age, specially in the beginning of life, in a similar fashion to that observed in the experiment with human prefrontal cortex samples presented above. However, once the mice reaches adulthood ($\geq$ 2M), it becomes hard to distinguish it from older mice. It was also interesting to observe that the black dot from a two year old female, which does not cluster with its own age group, belongs to a female, which at the time of sacrifice and tissue extraction, had a tumour in the spleen and a discolouration in liver, thus being not in the most desirable healthy state.

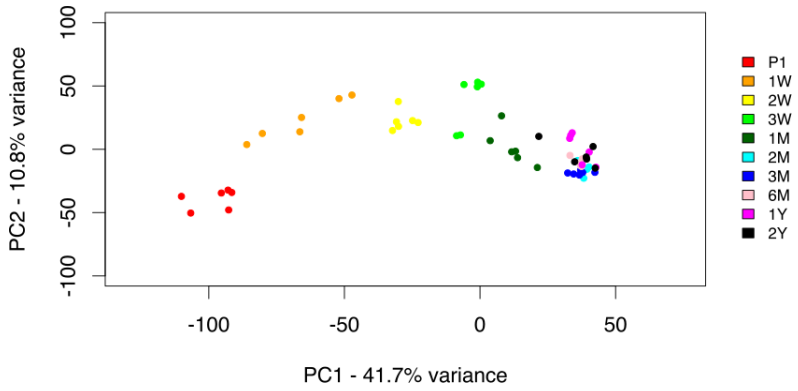


Figure 4.13: Principal component analysis of the female mouse liver gene expression set generated at EMBL-Monterotondo. Each dot represents one of the 60 female mouse liver samples in a multidimensional gene expression space projected on the principal plane formed by the first and second principal axes. The dots are coloured according to the age group the sample belongs to as enumerated in Table 4.2.

Third, a look to the individual gene expression profiles of the 931 orthologues of pharmacologically relevant genes (DT - 629, ME - 285, and TP - 86), revealed that only 8.3% of those genes changed their expression with age. Their expression patterns were classified into four main behaviours: a) no alterations in expression with age, 850 genes in this category (91.6%); b) the expression increases rapidly with age (i.e. at 2 or 3 weeks of age the expression of the gene reaches adult levels), 31 genes in this category (3.2%); c) the expression increases with age once the mice reaches puberty (transition from 1 month to 2 months of age), 15 genes

in this category (1.6%); and d) the expression decreases with age, 30 genes in this category (3.1%). A representative gene expression profile for each of these categories is depicted in figure 4.14.

In terms of the distribution of the drug targets, metabolic enzymes and transporters in these four categories, it was noteworthy the fact that the majority of genes in category b) and c) were metabolic enzymes and transporters (57%, and 22% respectively), such as sulfotransferase family cytosolic 1C member 2 (Sult1c2), and solute carrier family 22 member 26 (Slc22a26), while category d) was predominantly composed of drug targets (53%), like carbonic anhydrase 1 (Car1).

Also, it was interesting to observe that some metabolic enzymes associated with drugs with reported safety issues in paediatric populations (DDR - Safety), such as the members of the sulfotransferase family (SULT), in female mouse liver had an interesting expression behaviour, i.e. 33% of them increased their expression quickly with age. However, it is important to mention too, that no member of the glucuronosyltransferase family (UGT), also associated with drugs with reported safety issues in paediatric populations, was detected to alter their expression with age. This seems to suggest, that contrary to what is known in humans, i.e., sulfate conjugation in humans is the preferred metabolic route for the second phase of metabolism in infants, and glucoronidation is the primary pathway in adolescents and adults [14], in female mice, this switched is not present, at least judging by the lack of changes in the expression levels of the members of these two families as a function of age.

Additionally, there were four genes whose behaviour with age did not fall into any of the above categories. A noteworthy case being the cytochrome P450 2c55 (Cyp2c55), which increased its expression quickly with age, to later decrease it, and the case of solute carrier organic anion transporter family member 1a1 (Slco1a1), which increased its expression in adulthood, and decreased it in old age (fig. 4.15).

Integrating the analysis of this female mouse liver expression dataset generated at EMBL-Monterotondo with the analysis of the liver experiments from the public domain, one can then verify the reproducibility of the overall signals detected in mouse liver. Using the same visualisation approach as used for the human data, and for the previous animal models data, one can see from figure 4.16, that the majority of metabolic enzymes increased their expression with age consistently

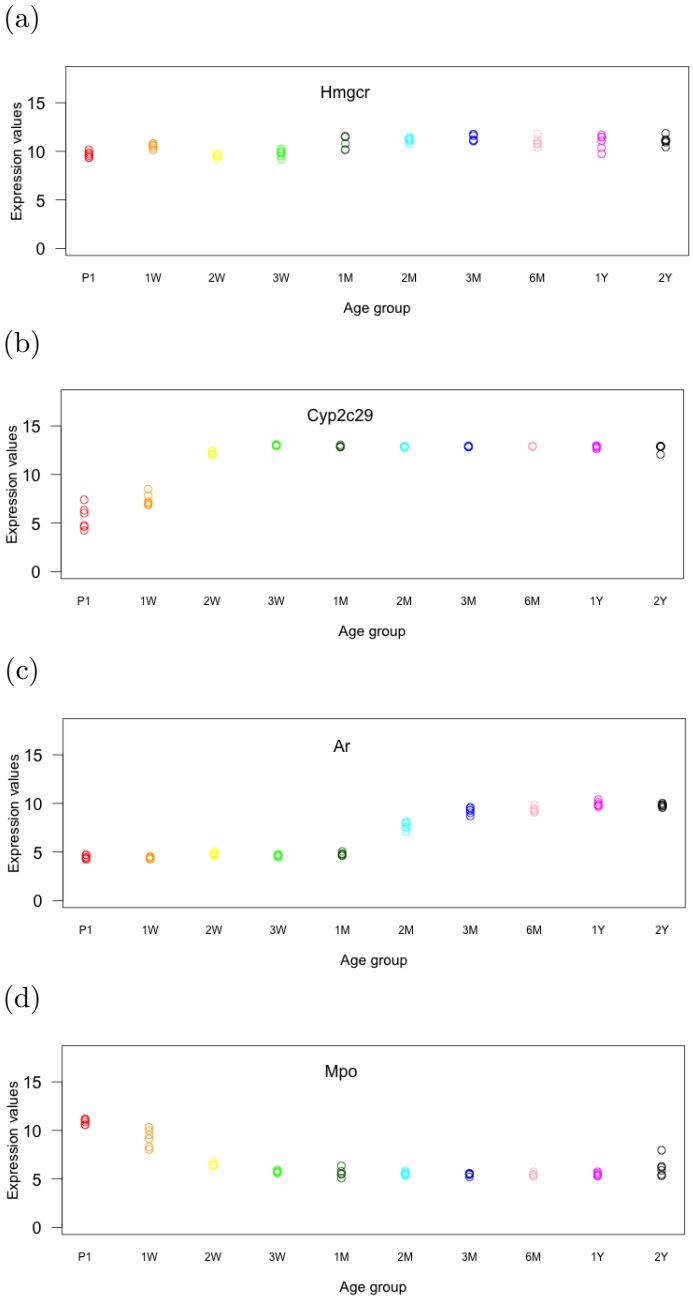


Figure 4.14: Overview of the main expression profiles as a function of age, among the pharmacologically relevant genes in female mice liver from the gene expression set generated at EMBL-Monterotondo. a) Expression profile of HMG-CoA reductase (Hmgcr); b) Expression profile of Cytochrome P450 2c29 (Cyp2c29); c) Expression profile of Androgen receptor (Ar); d) Expression profile of Myeloperoxidase (Mpo).

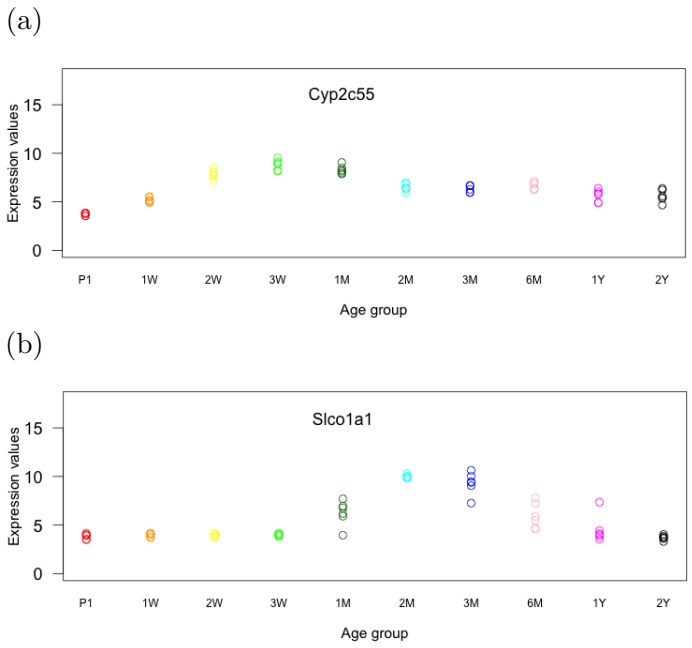


Figure 4.15: Overview of the unusual expression patterns detected in female mice liver from the gene expression set generated at EMBL-Monterotondo. a) Expression profile of cytochrome P450 2c55 (Cyp2c55); b) Expression profile of solute carrier organic anion transporter family member 1a1(Slco1a1).

across different experiments, with the exception of experiments only containing adult and elderly data (E-BAIR-10, E-MEXP-839, and to some extend E-MEXP-153, and E-MEXP-1504), where the tendency was to decrease their expression with age.

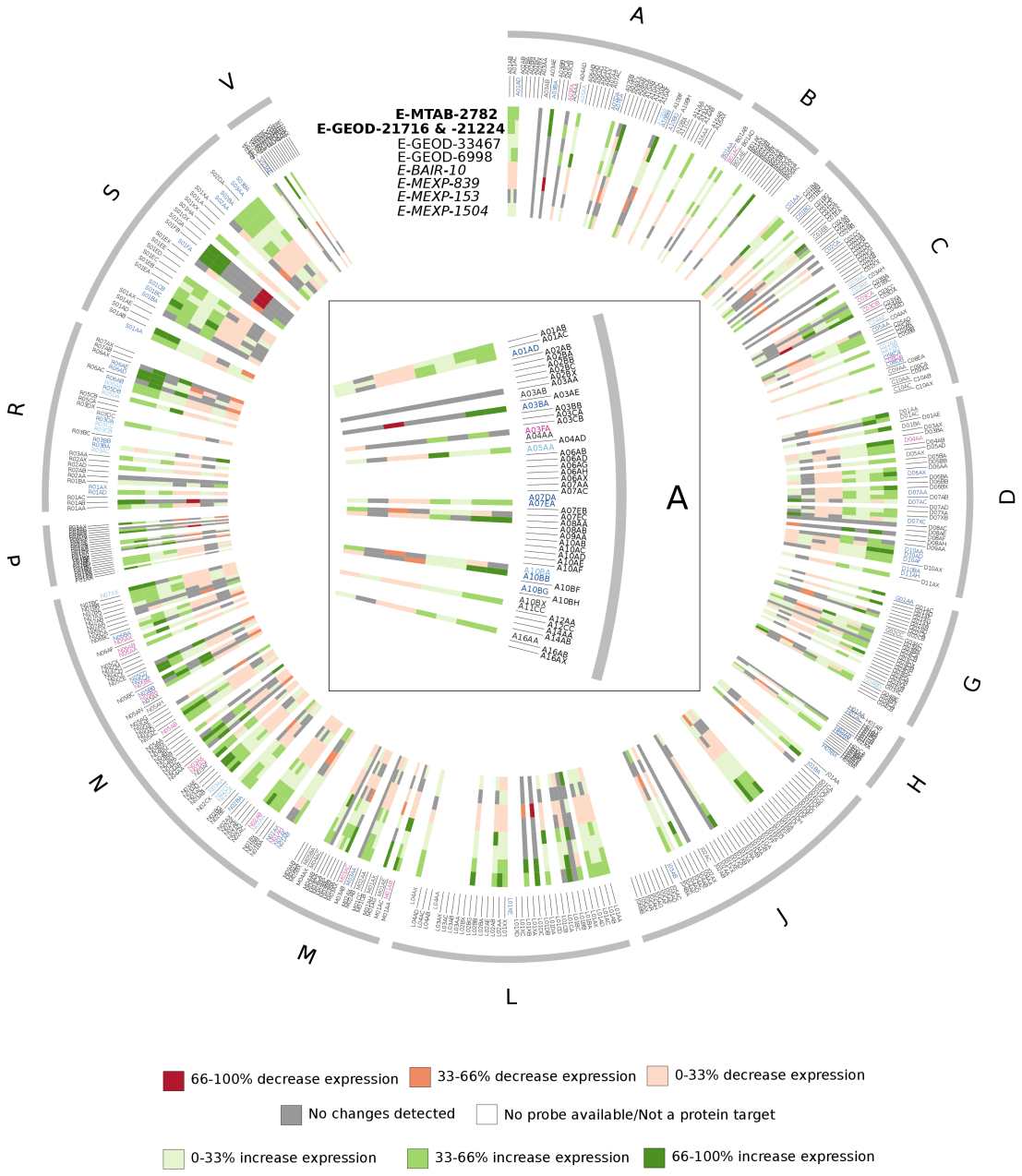


Figure 4.16: Overview of gene expression changes as a function of age of metabolic enzymes in mouse liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 4.9 for further details.

In addition, when comparing the female mouse liver expression dataset generated at EMBL-Monterotondo with the male mouse liver from the combination of experiments E-GEOD-21716 and E-GEOD-21224, it was possible to study the

differences in the expression patterns detected between male and female. From figure 4.16 again, it was clear that not many differences between gender were detected. However, there were two genes which had a significant change in behaviour: flavin containing monooxygenase 3 (Fmo3), which increased its expression with age in female mouse liver, while in male mouse liver there were no changes detected (fig. 4.17), and sulfotransferase family cytosolic 2A member 1 (Sult2a1), which increased its expression with age in female mouse liver, while in male mouse liver it decreased its expression. The Fmo3 behaviour is particular interesting, because in humans, it is the major flavin monooxygenase isoform in human liver microssomes [103; 131]. Therefore, if mouse is chosen as an animal model to study drug metabolism, and later to extrapolate this knowledge to humans, it is important to be aware of the fact that in males this particular isoform is not a major metabolic player, and that, if the presence of this isoform is actually desirable for the study in question, then female should be the gender selected.

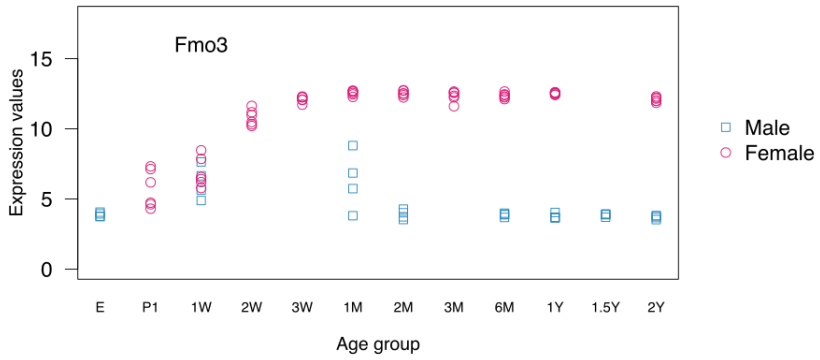


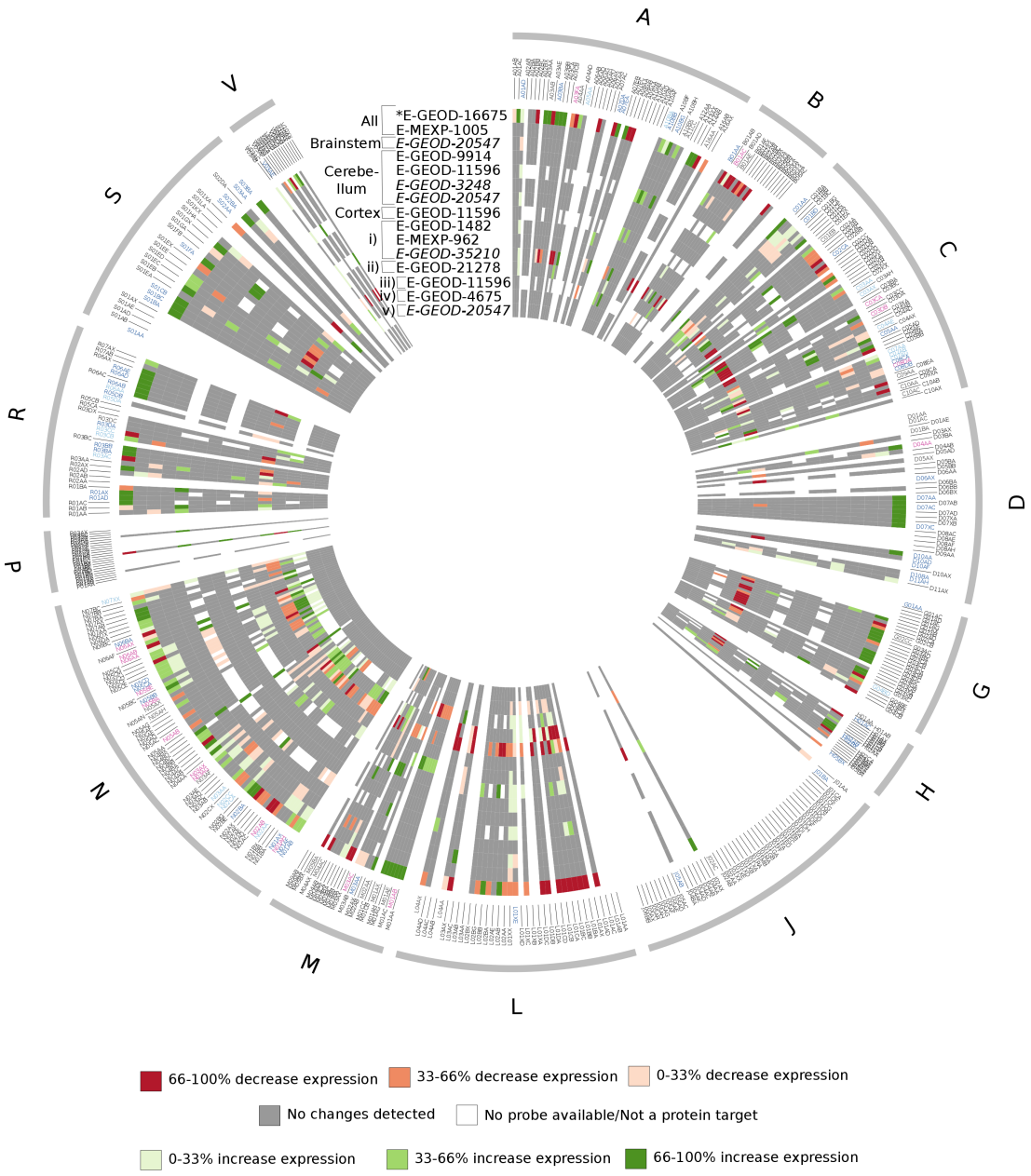
Figure 4.17: Expression profile of flavin containing monooxygenase 3 (Fmo3) in male and female mouse liver. The female data originates the gene expression set generated at EMBL-Monterotondo, and the male data from experiments E-GEOD-21716 and E-GEOD-21224.

Drug targets in mouse liver were also found to alter their expression with age, like carbonic anhydrase 3 (Car3), which increased its expression with age in 50% of the experiments (fig. 14 in Appendix C). Some transporters were also detected to alter their expression as a function of age in mouse liver (fig. 15 in Appendix C), such as ATP-binding cassette sub-family C member 3 (Abcc3), which increased

its expression in 37.5% of the experiments.

In mouse brain, there seemed to be a trend among drug targets to increase rather than decrease their expression with age, opposite to what I observed in human brain (fig. 4.18). Furthermore, there were more drug targets in mouse hippocampus whose expression changed with age, than in human hippocampus, but a comparable behaviour in the two other regions in common between human and mouse data, cerebellum and cortex. In terms of the expression of metabolic enzymes in mouse brain, despite the fact that brain is not a tissue specialised in metabolism, some alterations in expression of metabolic enzymes as a function of age were also detected (fig. 16 in Appendix C). Curiously, and in contrast to the results observed in human brain, not all metabolic enzymes seemed to increase their expression with age. Finally, among transporters, there were more alterations in gene expression detected as a function of age in mouse brain than in human brain (fig. 17 in Appendix C).

(a)


A
B
C
D
G
H
J
L
M
N
P
R
S
V
All
Brainstem
Cerebellum
Cortex
i)
ii)
iii)
iv)
v)
*E-GEOD-16675
E-MEXP-1005
E-GEOD-20547
E-GEOD-9914
E-GEOD-11596
E-GEOD-3248
E-GEOD-20547
E-GEOD-11596
E-GEOD-1482
E-MEXP-962
E-GEOD-35210
E-GEOD-21278
E-GEOD-11596
E-GEOD-4675
E-GEOD-20547
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

(b)

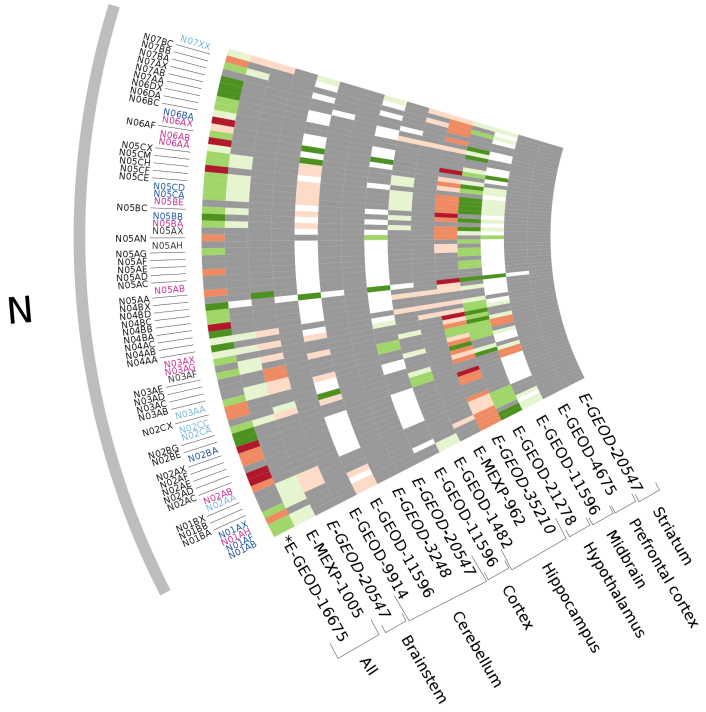


Figure 4.18: Overview of gene expression changes as a function of age of drug targets in (a) mouse brain per ATC level 4, and (b) in the nervous system category (ATC code 'N'). See caption from figure 4.9 for further details. i) Hippocampus, ii) Hypothalamus, iii) Midbrain, iv) Prefrontal cortex, v) Striatum.

In mouse kidney, pharmacologically relevant genes tended to increase their expression with age (fig. 18 in Appendix C), in a bigger proportion than what was detected in human, and rat kidney, despite the majority of those actually not being associated with drugs with reported DDRs.

In mouse heart, opposite to what was observed in human and rat data, one could observe alterations in gene expression as a function of age among drug targets (fig. 19 in Appendix C). The majority of changes increases with age. For example, peroxisome proliferator-activated receptor alpha (Ppara) increased its expression with age in both mouse heart expression sets. However, this drug target is associated with drugs in category "C10AB" of the ATC code, for which no information on the presence or absence of DDRs in humans was found in the literature. Based on my observations if one is to use mouse as an animal model of the human to

study the heart, in an experiment where this particular receptor is important, it may be wise to carefully select the age of the animals.

In mouse lungs, the trend was once more, among the pharmacologically relevant genes, to increase their expression with age (fig. 20 in Appendix C).

Finally, when comparing the expression of the top drug targets associated with drugs with reported DDRs when given to paediatric populations in mouse tissues, with their expression in human tissues (see Table 4.3 above), I observed that the GABA A receptor clearly increased its expression with age in mouse brain, opposed to what I observed in human brain (Table 15 in Appendix C). Among the glutamate [NMDA] receptor subunits, specially GRIN3A subunit, there was a shift from decreasing to increasing their expression with age in mouse brain, in comparison with the human data. The same behaviour could be detected among the voltage-gated calcium channel subunits (CACN*), and the sodium channel subunits (SCN*). Among the serotonin receptors, once more I noticed the opposite trend. In this case, in human brain the serotonin receptors tended to increase their expression with age, and in mouse brain, I observed them decreasing their expression with age, specially HTR3A and HTR1B. The serotonin transporter (Slc6a4) and the norepinephrine transporter (Slc6a2) were also detected to decrease their expression with age in a few experiments, oppose to the absence of alterations seen in human brain and rat brain data. Among transcription factors, the only one whose expression pattern differed between human and mouse in terms of expression as a function of age, was the glucocorticoid receptor (NR3C1), which seemed to increase its expression with age in mouse brain.

Among the top drug targets, associated with drugs with reported DDRs when given to the elderly (see Table 4.4 above), the expression patterns between human and mice, differ in the GABA A receptor as already mentioned, as well as, in the expression of the amiloride-sensitive sodium channel (SCNN1A), which increased its expression with age in mouse kidney; in the expression of the histamine H1 receptor (HRH1), which also increased its expression with age in mouse brain; and, in the expression of β2-adrenergic receptor (ADRB2), which was detected to increase its expression with age in mouse lungs, rather than decrease its expression like observed in one of the human lungs experiments (Table 16 in Appendix C). Curiously, there was no alteration detected in the target of sulfonylurea drugs

(ABCC8-KCNJ), contrary to what I had seen in rat liver, but in agreement with the expression patterns from the human liver data.

4.4 Summary

By querying gene expression public archives, I have identified 91 publicly available expression datasets, which allowed me to study the ontogeny of pharmacologically relevant genes in the context of age-related differential drug responses, in five primary tissues (brain, heart, lungs, liver and kidney), and in four different species (human, macaque, rat and mouse). The majority of the expression sets contained samples from brain and liver tissues, and mainly from human and mouse specimens. However, among the animal models data, not a single experiment covered the full life cycle of the animal, therefore, we decided to generated a new female liver mouse expression dataset to complement the data in the public domain. This expression set was analysed together with the data from the public domain, and allowed me to compare the expression profile of pharmacologically relevant genes between male and female mouse in liver, besides providing information on the expression profiles of pharmacologically relevant genes at all stages of the mouse life cycle.

Overall, among the human experiments, only in the brain expression datasets, I observed statistically and reproducible significant differential expression as a function of age of drug targets. In animal models, this trend was repeated and extended to other tissues as well. In term of metabolic enzymes, the majority seemed to increase their expression with age in all tissues and species.

Among the drug targets associated with drugs with reported DDRs when given to either paediatric or geriatric populations, it was rather curious to observe an overall decrease in expression with age among ion channels, such as GABA A receptor, and voltage-gated calcium channels in human brain, but detect the opposite trend in rodents. Additionally, it was interesting to observe that some serotonin receptors, like serotonin receptor 1D (HTR1D), increased their expression in the human brain, but in both rat and mouse, followed the opposite trend, and decreased their expression with age.

In the next chapter (chapter 5), these results, and its implications for drug response, and transferability of pharmacological knowledge between animal models

and humans, will be further discussed.

4.5 Material and Methods

4.5.1 Datasets from the public domain

ArrayExpress [116] was programmatically queried for gene expression experiments with *Homo sapiens*, *Macaca mulata*, *Rat norvegicus*, and/or *Mus musculus* samples. An XML file was retrieved for each species. A python script was subsequently used to parse the resulting XML files, and identify experiments with the following samples attributes: 'age'; and 'sex' or 'gender'. Experiments with these flags were manually inspected for information on tissue type, number of samples, age range covered, and disease status. If the experiment met the data requirements described in section 4.1.1, the experiment was selected for gene expression analysis.

4.5.2 Experimental methods

4.5.3 Ethics statement

All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of The European Molecular Biology Laboratory (EMBL, Heidelberg, Germany) and were performed in accordance with the Italian Ministry of Health and European regulations. All efforts were made to minimise suffering.

4.5.3.1 Mice

Wild-type C57BL/6JCrl mice from EMBL Monterotondo breeding colonies were obtained at 6 to 8 weeks of age and were mated, with pregnancy confirmed by the presence of a vaginal plug. Animals were kept in an animal house in groups of two to five per cage at 22-25°C on a 12-h light-dark cycle with water and food *ad libitum* (2019 Teklad Global 19% Protein Extruded Rodent Diet). Day of birth was recorded as P0 and pups were weaned at 21 days of age. The resulting 19 litters were randomly assigned to six groups, composed of six females and six males each, with a maximum of two animals from the same litter per gender and per group. Non-allocated pups returned to the breeding colonies. Animals were sacrificed

once they reached the age of the group to which they had been allocated to: 1 day old (P1), 1 week old (1W), 2 weeks old (2W), 3 weeks old (3W), 1 month old (1M), and 2 months old (2M). For the older age groups, wild-type C57BL/6JRcc females aged 2 months, 5 months, 11 months and 23 months were purchased from Harlan Laboratories (Harlan, Netherlands). The 24 animals, six per age group, were housed in the same conditions as the younger animals for one month, after which they were sacrificed.

4.5.3.2 Tissue extraction

All animals were sacrificed by exposure to dry ice in an euthanasia chamber followed by decapitation with surgical scissors. Dissection was performed on ice under aseptic conditions. Liver, brain, kidney, heart and lungs were extracted and washed in DEPC treated phosphate-buffered saline (pH 7.2) solution (PBS-DEPC), and quickly submerged in RNA*later* RNA stabilization reagent (Qiagen, Milano, Italy). The samples were kept at 4°C for one day in RNA*later*, before being transferred to a -80°C freezer as recommended by the manufacture.

4.5.3.3 RNA isolation

Total RNA was isolated from the 60 female mice liver samples using a standard RNeasy mini kit (Qiagen, Milano, Italy) protocol. The tissue was disrupted and homogenised with a rotor-stator homogeniser (Polytron, PT1200 E., Kinematica) and further blended using a syringe and needle. RNA concentration was determined with a NanoDrop 8-sample spectrophotometer (ND-8000, Thermo scientific). Quality of RNA was tested using Agilent Bioanalyzer 2100 (Agilent Technologies Inc., Santa Clara, USA). RNA integrity number (RIN) values were always above 6, with most samples having a RIN between 8 and 9.

4.5.3.4 Sample hybridisation

Samples with total RNA were processed for hybridisation onto Affymetrix GeneChip Mouse Gene 1.0 ST Arrays (Affymetrix, Santa Clara, CA), by the gene core facility at the European Molecular Biology Laboratory (EMBL, Heidelberg, Germany), according to manufacture instructions. Briefly, in the first cycle, cDNA

was synthesised using a priming method that specifically primes non-ribosomal RNA (including both polyA and non-polyA containing mRNAs). A T7 promoter sequence was also incorporated and used in the subsequent synthesis of antisense RNA. In the second cycle, a modified dUTP was incorporated in cDNA instead of dTTP. The modified dUTP incorporated in the DNA was subsequently recognised and cleaved by Urasil-DNA glycosylase (UDG) and human apurinic/apyrimidinic endonuclease 1 (APE 1), resulting in cDNA fragmentation. Each DNA fragment was end-labeled with biotin using terminal deoxynucleotidyl transferase (TdT), before being hybridised to the arrays. Chips were scanned on Affymetrix Complete GeneChip Instrument System generating digitised image files, which were later converted to CEL files with Affymetrix software. All array data was deposited into ArrayExpress [116], and is accessible under the accession number E-MTAB-2782.

4.5.4 Microarray data analysis

All data was processed and analysed in R [80], version 2.15.1, using multiple Bioconductor packages [132]. Quality of array data was examined with the *arrayQualityMetrics* package [133]. Samples which were marked as outliers when comparing the distance between arrays, the distribution of array intensities, and the spatial arrangement of arrays after a principal component analysis, were removed from the dataset and not used in the downstream analysis. Additionally, samples which did not met the criteria from section 4.1.1, such as, samples from cancerous tissues, or samples from knock-out animals, were also discarded. Samples which passed the quality control and data requirements criteria were analysed per experiment and per tissue. Raw affymetrix data was background corrected, normalised and summarised with the robust multi-array average (RMA) algorithm [134] provided with the *affy* package [135]. Raw illumina data was summarised, and quantile normalised with the algorithms provided in the *beadarray* package [136]. Raw agilent data was background corrected using local background estimates as described by Ritchie *et al* [137] and Silver *et al* [138], and quantile normalised between arrays [139] with the algorithms provided in the *limma* package [140]. In the absence of raw data, processed data was used for downstream analysis, if, for a given tissue and species, there was no alternative experiment with available raw data.

Age groups were defined as classes ranging from 0 years of age to 100 years of age, in increments of one year, for human data. For animal models, the age groups provided with the experimental information, were the groups used for downstream analysis. All expression datasets were subjected to principal component analysis (PCA), and the variance in the data was compared with changes in age. Pairwise comparisons between age groups were performed using the linear models for microarray data algorithm from the *limma* package [140]. The method of Benjamini-Hochberg [141] was used to control the false discovery rate at the level of 0.05 to correct for multiple hypothesis testing. A significantly differentially expressed gene in a given age group comparison was defined as having an adjusted p-value ≤ 0.05, and an absolute fold change ≥ 1.25. Overall, a gene was defined as differentially expressed as a function of age, if the average of the p-values from all pair-wise comparisons was ≤ 0.05, and the average of slopes from the pair-wise comparisons was in absolute terms ≥ 1. Genes with similar expression profiles were clustered with the fuzzy c-means algorithm [125] provided by the *Mfuzz* package [142], followed by manual inspection of the clusters identified, in order to detect the general trends in the data.

4.5.5 NGS data analysis

All data was processed and analysed in R [80], version 2.15.1, using multiple Bioconductor packages [132]. Read count data, already summarised per transcript, was inspected for quality control, and samples were discarded if the number of reads per sample was bellow 1×10^{-7}, and/or the sample was identified as an outlier when examining the distance between samples. Moreover, samples which did not met the criteria from section 4.1.1, such as, samples from cancerous tissues, or samples from knock-out animals, were also discarded. Additionally, low expressed genes were discarded from further analysis to minimise distortion on the detection of age-related effects. A gene was considered low expressed if the average expression value in all samples divided by the variance was bellow 0.5. Samples which passed the quality control and data requirements criteria were normalised by the effective library size (depth of sequencing) per experiment and per tissue, using the algorithms provided by the *DESeq* package [143].

Age groups were defined as classes ranging from 0 years of age to 100 years of

age, in increments of one year, for human data. For animal models, the age groups provided with the experimental information, were the groups used for downstream analysis. All expression datasets were subjected to principal component analysis (PCA), and the variance in the data was compared with changes in age. Pairwise comparisons between age groups were performed for each gene by fitting generalised linear models (GLMs) for each age group, which were then compared in order to infer whether the additional specification of the difference in age improved the fit, hence, whether it had a significant effect in the gene expression as described in the *DESeq* package [143]. The method of Benjamini-Hochberg [141] was used to control the false discovery rate at the level of 0.05 to correct for multiple hypothesis testing. Overall, a gene was defined as differentially expressed as a function of age, if the average of the p-values from all pair-wise comparisons was ≤ 0.05, and the average of slopes from the pair-wise comparisons was in absolute terms ≥ 1. Genes with similar expression profiles were clustered with the fuzzy c-means algorithm [125] provided by the *Mfuzz* package [142], followed by manual inspection of the clusters identified, in order to detect general trends in the data.

4.5.6 Linking drugs to gene expression data

Human orthologues of pharmacologically relevant genes of *Macaca mulata*, *Rat norvegicus*, and *Mus musculus* were extracted from ENSEMBL Compara version 72 [144; 145]. The drug maps compiled in chapter 3 (Appendix B - Table 3, 4, and 5) were used to link drugs to the human, macaque, rat and mouse gene expression patterns of pharmacologically relevant genes as a function of age. The WHO Anatomical Therapeutic Chemical Classification System code (ATC code) [62], used in chapter 2 as a standard indication of the drug's action, was used to link gene expression changes with age to the anatomical and biological system of the drug's action. All the figures were produced using the following visualisation programs: Circos [146] and Inkscape [147].

Chapter 5

General discussion

In this last chapter I will explore the implications of the main findings from the previous chapters when attempting to rationalise age-related differential drug responses. Specifically, I will discuss if the differences in the ontogeny of pharmacologically relevant genes detected between differently aged human populations in chapter 4, can explain why certain drugs have reported age-related differential drug responses (DDRs) when transferred from one aged population to another. In addition, I'll speculate how the differences in the ontogeny of pharmacologically relevant genes between human and animal models, could influence pre-clinical studies. Furthermore, I will scrutinise the pitfalls of the systematic analysis employed to tackle the research question of this thesis, and elaborate on the influence of other factors, such as race and gender, on differential drug responses. Finally, I will discuss how other emerging fields, such as proteomics and metabolomics could bring alternative perspectives into this project in the near future.

5.1 Overview of gene expression changes in the context of age-related differential drug responses

5.1.1 In paediatric populations

As mentioned in chapter 2, 40 drugs were identified as having shown severe or unseen side effects in paediatric populations when transferred from adult populations, and 31 as having shown a lack of efficacy. The large majority were drugs modulating the nervous system (category N of the ATC code), such as antidepressants and antiepileptics. Their therapeutic targets were mainly GABA A receptors, calcium and potassium channels, and serotonin related targets, such as the serotonin transporter and serotonin receptors (see chapter 3). In the control group, composed of 31 drugs, there was no clear prevalent therapeutic indication, which ranged from treatment of hyperparathyroidism associated with chronic kidney disease (CKD), to attention deficit hyperactivity disorder (ADHD). Among their drug targets, glutamate [NMDA] receptor, sodium channels, dopamine type 2 receptor, and serotonin receptor 2A stood out.

The expression of those drug targets with age were examined in chapter 4. For the majority of the drugs, no alterations in the expression of their drug targets were detected in human tissues, as in case of the selective serotonin re-uptake inhibitors (SSRIs), whose drug target, serotonin transporter, did not alter its expression with age in the human brain. This observation seems to suggest that the reason why the majority of SSRIs, with the exception of fluoxetine, are not effective in children is not due to pharmacodynamic changes in their drug target. However, it was interesting to observe that for the minority of drugs whose drug target altered its expression with age in the 'DDR-safety' and/or in the 'DDR-efficacy' groups (33%, $n = 19$), there were more drugs, whose target expression decreased with age when compared with the control group (fig. 5.1).

Table 5.1 lists all the drugs in the three DDR categories where alterations in the expression of the drug targets has been detected in the human tissues. Among these drugs there is a clear over representation of drugs which modulate the nervous system. Partially because it is the biggest class of drugs, among the drugs with

reported safety or efficacy issues when transferred from the adult populations to paediatric populations. But also because only in the human brain I detected significant and reoccurring changes in expression of pharmacologically relevant genes, when analysing human data, as already briefly discussed in chapter 4.

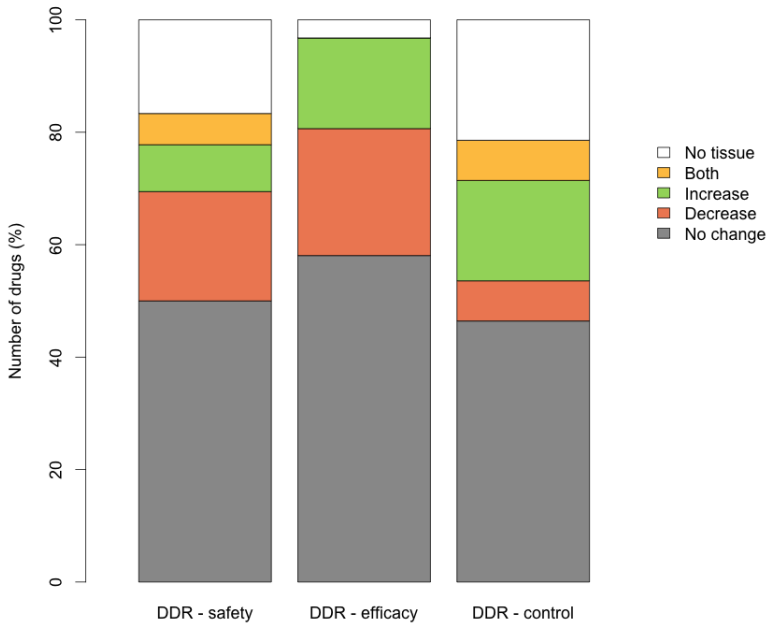


Figure 5.1: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, per expression patterns of the drug targets with age in relevant human tissues. No change - No changes detected in the expression levels of drug targets as a function of age; Decrease - At least one of the drug targets (or one of the drug target subunits) decrease expression levels with age; Increase - At least one of the drug targets (or one of the drug target subunits) increase expression levels with age; Both - An equal proportion of drug targets (or an equal proportion of drug target subunits) decrease and increase expression levels with age; No tissue - No relevant tissue was analysed, which allowed me to study the ontogeny of a particular drug target. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.55); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.058); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.24).

Table 5.1: List of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatrics when compared with adult populations, and the expression patterns of their drug targets in relevant human, rat and mouse tissues as a function of age. The pattern depicted is the most prevalent among the drug targets (or subunits), which change with age. H. Exp. - Expression pattern in human tissues ; R. Exp. - Expression pattern in rat tissues; M. Exp. - Expression pattern in mouse tissues.

Drug	**ATC code**	**DDR category**	**Target gene symbol**	**H. Exp.**	**R. Exp.**	**M. Exp.**
Buspirone	N05BE01	Safety, Efficacy	HTR1A	↓	↑	↓
Desflurane	N01AB07	Safety	GABR*, KCNK*	↓	↑	↑
Gabapentin	N03AX12	Safety, Efficacy	CACN*	↓	↓	↑
Midazolam	N05CD08	Safety	GABR*	↓	↑	↑
Mirtazapine	N06AX11	Safety, Efficacy	ADRA2*, HTR2A, HTR2C	↓	↑	↑
Propofol	N01AX10	Safety	GABR*	↓	↑	↑
Sevoflurane	N01AB08	Safety	GABR*, KCNK*	↓	↑	↑
Metoclopramide	A03FA01	Safety, Efficacy	DRD2, HTR3A, HTR4	↑	↑	↓
Nefazodone	N06AX06	Safety, Efficacy	HTR2A, HTR2C, SLC6A2, SLC6A4	↑	↑	↓

Continued on next page

Table 5.1 – *Continued from previous page*

Drug	ATC code	DDR category	Target gene symbol	H. Exp.	R. Exp.	M. Exp.
Theophylline	R03DA02, R03DA04, R03DA54, R03DA74	Safety	ADORA*, PDE3*, PDE4*	↑	?	-
Fluticasone Propionate	D07AC17, R01AD08, R01AD58, R03BA05	Safety	NR3C1, PGR	↑ & ↓	?	-
Lamotrigine	N03AX09	Safety	SCN*	↑ & ↓	↓	↑
Dextromethorphan	N07XX59, R05DA09	Efficacy	GRIN2A, SIGMAR1	↓	↑ & ↓	↑ & ↓
Diphenhydramine	D04AA32, R06AA02, R06AA52	Efficacy	HRH1	↓	?	-
Lorazepam	N05BA06	Efficacy	GABR*	↓	↑	↑
Phenobarbital	N03AA02	Efficacy	GABR*	↓	↑	↑
Codeine	N02AA59, N02AA79, R05DA04	Efficacy	OPRD1, OPRK1, OPRM1	↑	↑	↓
Rizatriptan	N02CC04	Efficacy	HTR1B, HTR1D	↑	↑	↓
Sumatriptan	N02CC01	Efficacy	HTR1B, HTR1D	↑	↑	↓
Felbamate	N03AX10	Control	GRIN*	↓	↓	↑ & ↓

Continued on next page

Table 5.1 – *Continued from previous page*

Drug	ATC code	DDR category	Target gene symbol	H. Exp.	R. Exp.	M. Exp.
Ketamine	N01AX03	Control	GRIN*	↓	↓	↑ & ↓
Clozapine	N05AH02	Control	DRD2, DRD3, DRD4, HTR2A	↑	↑	↓
Ibuprofen	C01EB16, G02CC01, M01AE01, M01AE51, M02AA13	Control	PTGS1, PTGS2	↑	↑ & ↓	↑
Montelukast	R03DC03	Control	CYSLTR1	↑	?	-
Olanzapine	N05AH03	Control	DRD2, DRD3, DRD4, HTR2A, HTR2C	↑	↑	↑
Risperidone	N05AX08	Control	DRD2, DRD3, DRD4, HTR2A, HTR2C	↑	↑	↑
Glycopyrrolate	A03AB02, R03BB06	Control	CHRM1, CHRM3	↑ & ↓	↓	-
Rufinamide	N03AF03	Control	SCN*	↑ & ↓	↓	↑

GABR* - GABA A receptor; KCNK* - Potassium channel subfamily K members; CACN* - Voltage-gated calcium channel; ADRA2* - α-adrenergic receptor 2A, 2B and 2C; ADORA* - Adenosine receptor; PDE3* - Phosphodiesterase 3; PDE4* - Phosphodiesterase 4; SCN* - Sodium channel; GLR* - Glycine receptor; ↑ - Increase expression with age; ↓ - Decrease expression with

age; '-' - No change in expression with age detected; ? - No relevant tissue to analyse gene expression.

However, those changes in expression are not equal throughout the human brain. In fact, a lot of the changes I detected in human brain data were mainly in expression datasets from the prefrontal cortex, cerebral cortex and frontal cortex. Interestingly, the prefrontal cortex activity is affected during depression, i.e. its activity increases after successful antidepressant therapies [148], but it is also an important region in attention-hyperactive disorders (ADHD). Curiously, α-adrenergic receptor 2A (ADRA2A) is one of the drug targets, whose expression decreases in all three expression datasets from human prefrontal cortex, and it is a target for gaunfancine. Gaunfancine was initially approved for the management of hypertension, however, it was not considered the most effective in the elderly among its drug class [149], but in the last years it has been prescribed off-label to children with ADHD with renown therapeutic success [150]. This observation seems to correlate with the expression pattern of this target.

When comparing the human data with the animal models data, at least in terms of the expression with age of drug targets associated with drugs with reported DDRs in paediatrics, overall rat seems closer to human than mouse (Table 5.1). Although, for certain ion channels, both rat and mouse have an opposite expression pattern to what was detected in human. In particular, in humans the GABA A receptors, and the potassium channels decrease their expression with age, while in rodents they increase. Also, in humans, sodium channels were detected to both increase and decrease its expression with age, while in rat there is a tendency to decrease with age, and in mouse to increase. It is not clear why these differences arouse, possibly different regulation mechanisms are in place, and/or the regulatory machinery might have diverged between species [151]. Consequently, to study drug response in the context of children *vs.* adult, one may actually need to use an adult mouse to model children's data, and an infant mouse to model adults's responses. However, a word of caution is necessary, as the brain regions covered by the human expression datasets do not overlap 100% with the ones from rat and mouse. In short, in rat the cerebral cortex is the region that changes the most with age, and there is no prefrontal cortex expression dataset to compare it with. Likewise in

mouse, hypothalamus, hippocampus and 'all brain' are the regions where most changes are detected. Additionally, despite having one mouse expression dataset with samples from prefrontal cortex, the signals identified do not correlate with what was detected in the human equivalent dataset. Similarly, in rodents several changes in the metabolic machinery were detected in liver with age, but nothing could be verified in the human data. Possibly because of confounder factors such as diet, but also, as already mentioned in chapter 4, the majority of switches seem to happen very early in life, consequently given that we only have access to a few samples from that period, it is likely that there simply isn't enough statistical power to detect those changes. In any case, one needs to keep in mind these differences in age when considering transferability of studies between species. Specially when doing toxicology studies in rat, one should be aware of the fact that not all metabolic enzymes are expressed before the rat reaches adulthood (i.e. 2.5 months).

5.1.2 In geriatric populations

A total of 60 drugs were identified in chapter 2 as having reported either severe and unseen side effects ($n = 47$), and/or as having a lack of efficacy ($n = 22$) in geriatric populations when compared with adult populations. The majority of the drugs modulate the cardiovascular system (ATC category C), such as antiarrhythmics or antihypertensives, followed by the ones modulating the nervous system (ATC category N), such as antidepressants and anxiolytics. In terms of their drug targets, GABA A receptor was once more a frequent drug target, followed by L-type calcium channels, sulfonylurea receptor, histamine receptors, opioid receptors and β-adrenergic receptors.

The expression of those drug targets were also examined in chapter 4. Similar to that observed in section 5.1.1, for the majority of drugs, their drug targets do not change their expression with age. However, for 32% ($n = 19$) of the drugs in either DDR-safety and/or DDR-efficacy groups, and similar to that observed in the section above, the drug targets tend to decrease their expression with age, in contrast to the behaviour of the drug targets of the drugs in the control group (fig. 5.2). This pattern, specially in the DDR-safety category, is driven by the presence of many benzodiazepines in this category, since the majority of GABA A

subunits that change with age, tend to decrease its expression with age in human brain. While in the DDR-efficacy category, the pattern is mainly driven by the beta adrenergic inhibitors, whose target also decreases with age in human lungs.

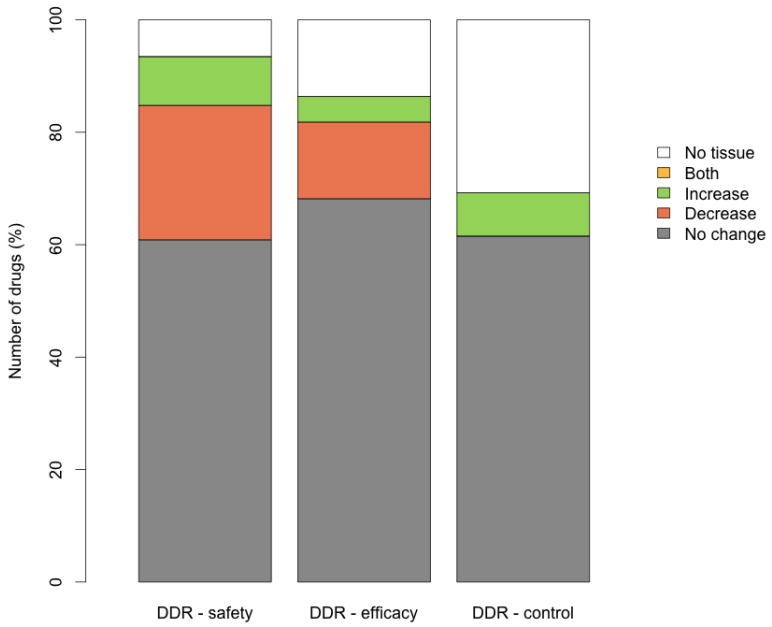


Figure 5.2: Distribution of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, per expression patterns of the drug targets with age n relevant human tissues. No change - No changes detected in the expression levels of drug targets as a function of age; Decrease - At least one of the drug targets (or one of the drug target subunits) decrease its expression with age; Increase - At least one of the drug targets (or one of the drug target subunits) increase its expression with age; Both - An equal proportion of drug targets (or an equal proportion of drug target subunits) decrease and increase its expression with age; No tissue - No relevant tissue was analysed, which allowed me to study the ontogeny of a particular drug target. DDR - safety *vs.* DDR - control (Fisher's exact test, p-value = 0.031); DDR - efficacy *vs.* DDR - control (Fisher's exact test, p-value = 0.41); DDR - safety *vs.* DDR - efficacy (Fisher's exact test, p-value = 0.59).

Table 5.2 lists all the drugs in the three DDR categories where alterations in the expression of the drug targets has been detected in the human tissues. Despite the fact that a big portion of drugs with reported safety or efficacy issues when given to geriatric populations, modulate the cardiovascular system, no alterations in drug targets in the human heart could be identified, consequently the table is mainly dominated by drugs which modulate the nervous system. This trend seemed to suggest that either pharmacodynamic changes are not the reason why these drugs have reported DDRs when transferred from adults to the elderly, or intrinsic issues with the quality of data. For example, as mentioned in chapter 2 and in chapter 4, β-2 adrenergic receptor antagonists are not considered effective in the elderly for the treatment of hypertension, and according with the literature, they loose their efficacy because, β-2 adrenergic receptor supposedly decreases its expression with age [13]. This pattern was, however, not detected in any of the human heart experiments available in the public domain, and the reason for this apparent discrepancy is not clear, apart from the possibility of influence on gene expression by confounder factors, such as non-reported medical comorbidities, or simply differences in study design (i.e. number of subjects per age group).

Table 5.2: List of drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatrics when compared with adult populations, and the expression patterns of their drug targets in relevant human, rat and mouse tissues as a function of age. The pattern depicted is the most prevalent among the drug targets (or subunits), which change with age. H. Exp. - Expression pattern in human tissues ; R. Exp. - Expression pattern in rat tissues; M. Exp. - Expression pattern in mouse tissues.

Drug	ATC code	DDR category	Target gene symbol	H. Exp.	R. Exp.	M. Exp.
Chlordiazepoxide	N05BA02	Safety	GABR*	↓	↑	↑
Chlorpheniramine	R06AB04, R06AB54	Safety	HRH1	↓	?	-

Continued on next page

Table 5.2 – *Continued from previous page*

Drug	ATC code	DDR category	Target gene symbol	H. Exp.	R. Exp.	M. Exp.
Cyclizine	R06AE03, R06AE53	Safety	HRH1	↓	?	-
Diazepam	N05BA01	Safety	GABR*	↓	↑	↑
Etomidate	N01AX07	Safety	GABR*	↓	↑	↑
Flurazepam	N05CD01	Safety	GABR*	↓	↑	↑
Midazolam	N05CD08	Safety	GABR*	↓	↑	↑
Promethazine	D04AA10, R06AD02, R06AD52	Safety	HRH1	↓	?	-
Propofol	N01AX10	Safety	GABR*	↓	↑	↑
Thiopental	N01AF03, N05CA19	Safety	GABR*	↓	↑	↑
Triazolam	N05CD05	Safety	GABR*	↓	↑	↑
Aspirin	A01AD05, B01AC06, N02BA01, N02BA51, N02BA71	Safety	PTGS1, PTGS2	↑	↑ & ↓	↑
Ipratropium	R01AX03, R03BB01	Safety	CHRM3	↑	?	-
Ketorolac	M01AB15, S01BC05	Safety	PTGS1, PTGS2	↑	↑ & ↓	↑
Theophylline	R03DA02, R03DA04, R03DA54, R03DA74	Safety	ADORA*, PDE3*, PDE4*	↑	?	-

Continued on next page

Table 5.2 – *Continued from previous page*

Drug	ATC code	DDR category	Target gene symbol	H. Exp.	R. Exp.	M. Exp.
Metoclopramide	A03FA01	Safety	DRD2, HTR3A, HTR4	↑	↑	↓
Albuterol	R03AC02, R03CC02	Efficacy	ADRB2	↓	?	↑
Ergoloid	C04AE01, C04AE51	Efficacy	ADRA*, DRD*, HTR1A, HTR1B	↓	↑	↓
Isoproterenol	R03CB51	Efficacy	ADRB*	↓	?	↑
Trazodone	N06AX05	Control	HTR2A, HTR2C, SLC6A4	↑	↑	↓

GABR* - GABA A receptor; ADORA* - Adenosine receptor; PDE3* - Phosphodiesterase 3; PDE4* - Phosphodiesterase 4; ADRA* - α-adrenergic receptors; DRD* - Dopamine receptors; ADRB* - β-adrenergic receptors; ↑ - Increase expression with age; ↓ - Decrease expression with age; '-' - No change in expression with age detected; ? - No relevant tissue to analyse gene expression.

Similarly to observations in the previous section, when looking to drugs with reported DDRs in geriatrics, the expression patterns detected with age of their drug targets in humans, is more comparable with the ones in rat, than with the ones in mouse. Likewise, the same observations apply in terms of the expression of the GABA A receptor in humans and in rodents, as well as, the same cautionary tail regarding expression patterns in different brain regions. In terms of drug response, for drugs interacting with the GABA A receptor, for instance, once more the effects in rodents, in this case in older rat or mice, would be slightly different than in humans, i.e. as the expression does not seem to decrease with age, the increase in sensitivity seen with these type of medication in this population would potentially not be mimicked in rodents.

5.1.3 Clinical implications

From the two sections above, it was clear that among the drugs with reported safety and/or efficacy issues when transferred from one aged population to another, whose drug target changed their expression with age, the prevalent expression pattern among those drug targets was to decrease their expression with age. This observation seems to suggest that one of the reasons why a drug may lack efficacy in children, for instance, may be because the dose given was not high enough to occupy all the expressed receptors, and thus fully activate the necessary downstream signals for an effective therapy. Likewise, this trend also seems to suggest, that for geriatric populations, a smaller dose may be required to prevent over inhibition or over activation of these drug targets. Figure 5.3 illustrates this working hypothesis.

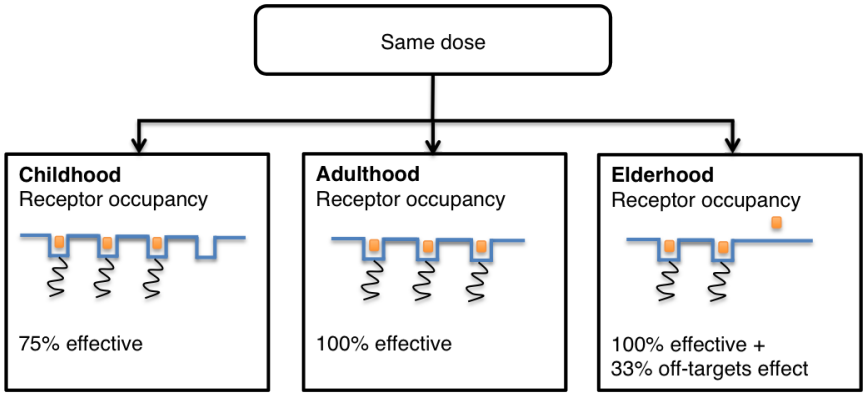


Figure 5.3: Diagram illustrating receptor occupancy in children, adults, and in the elderly, for drug targets, whose expression levels decrease with age, such as the GABA A receptor in human brain.

At first glance, this model is rather provocative, given that it only seems to explain a small fraction of the differential drug responses described in the literature (17%, $n = 18$), and, at least in case of children, it is more intuitive to decrease their dose, given their smaller size. However, recent studies demonstrated that increasing the dose of phenobarbital and gabapentin, actually helped, i.e. these medicines were considered effective in children after the dose being adjusted [64; 152]. It is

important to mention, however, that in case of phenobarbital, there is another factor contributing to the fact that a higher dose is required in children. In short, the pH in the stomach is almost neutral until the child reaches two years of age, contributing to decrease the absorption of medicines with a weakly acidic composition, such as phenobarbital. Thus the paediatric populations require a higher dose to achieve a therapeutic plasma concentration [41]. However, this physiological difference only supports the increased dose of phenobarbital in children up to two years of age. Consequently, the proposed hypothesis based on the pharmacodynamics detected in the expression changes of the GABA A receptor with age in human brain, could explain the need to still have higher doses in children older than 2 years of age, in order to achieve the same therapeutic response as in adults. In addition, as the elderly appear to be more sensitive to benzodiazepines, i.e. they need a 2-3 fold lower plasma concentration to achieve the same safe and effective response as an average adult, and given that for some drugs the need for this lower dose could not be explained by pharmacokinetic changes [32], the proposed model seems to provide a plausible explanation for it.

For the remaining drugs with reported DDRs in the database ($n = 88$), and in particular to the ones whose drug target increases, or equally increases and decreases their expression with age ($n = 13$), such as the case of Lamotrigene and Rufinamide[1], this model is not adequate. Additionally, there are three drugs in the DDR-efficacy category when given to the elderly (Albuterol, Isoproterenol and Ergoloid), for which the drug target seems to decrease its expression with age, which suggests that maybe the concentration of target is not able to transmit enough signals for a full therapeutic effect. In this case, though, a dose adjustment does not seem the most appropriated course of action, being suggested to switch to a different drug class to achieve the same therapeutic effect. Lastly, for the 54 drugs[2], for which no changes in the expression of their drug targets with age were detected in human tissues, the data seems to suggest two hypothesis: 1) either the DDR reported is a consequence of physiological differences between differently

[1]Both drugs have sodium channels as drug targets, and SCN1A goes up with age, while SCN3A goes down in almost the same number of experiments (5 *vs.* 4).

[2]There are extra 12 drugs for which the tissue where the drug exert its therapeutic action was not available to study the ontogeny of drug targets, and 6 other drugs whose target is not a human protein.

aged populations, which substantially affects the pharmacokinetic profile of the drug in question, rather than due to pharmacodynamic differences, and thus if one would develop appropriated dosing algorithms for that drug for different age populations this effect could be corrected; 2) or, the data used in this thesis to study the ontogeny of pharmacologically relevant genes is too noisy to detect any mild transcriptional switch with age, i.e., it is difficult to disentangle the age signal from other confounding factors, such as gender, race, medical history, diet, *etc.*. Based on the literature, the first case seems to apply to diuretics, which seem to be lacking efficacy in the elderly not because of gene expression changes, but because the size of the kidney changes [13], while for antihypertensive agents belonging to the beta-adrenergic antagonists category, the last scenario seems plausible to explain the fact that I was unable to detect this receptor decreasing its expression with age in human heart, contrary to what has been described in the literature [13]. In addition, specially in case of drugs with reported unseen or more severe safety issues in paediatric or geriatric patients, one also has to consider the possibility that the reported DDR is a product of off-target effects, which could differ between differently aged populations in the same fashion as the ontogeny of the main targets may differ.

Nevertheless, for the drugs whose DDR could potentially be explained by this model, complementing what is known about the drugs pharmacokinetic profiles and its alterations with age, with information on pharmacodynamic changes from this work, could be used to guide practitioners on how to better adjust the doses of the medicines here listed, and potentially improve existing dosing algorithms in the future.

5.1.4 Pitfalls

As in every project studying a complex biological question, there are limitations in the methods employed. This systematic analysis on the ontogeny of pharmacologically relevant genes is, therefore, not an exception to the rule. The first obvious limitation is the data type, and data quality. Ideally, one would have access to both messenger RNA (mRNA) and protein levels, given that those are only weakly positively correlated [114; 115], and to extensive phenotypic annotations of all the samples of interest (i.e. medical history, diet, race, *etc*). In reality, as mentioned

in chapter 4, there are not many proteomic studies available in the public domain yet, which could be re-used to address the research question of this project, nor do I have access to even a third of the desired phenotypic annotations. This particular point, as alluded before, may help to explain why for certain human tissues, no significant gene expression changes with age were detected, which have been previously reported in the literature, such as the lower expression levels of the beta adrenergic receptor in human heart in the elderly [13]. Meaning, without being possible to investigate the individual effect of confounder factors, it is not possible to decouple noise from true positive age signals. The presence of confounder factors, may thus, contribute to lower the number of drugs with reported DDRs, for which the DDR could potentially be rationalised based on the expression of the respective drug target. Although, it is more likely that for the majority of drugs with reported DDRs, differences in physiology play a bigger role, in the final pharmacokinetics of the drug, and ultimately therapeutic effect, than alterations in the pharmacodynamics of their drug targets.

The second pitfall, is a rather pertinent one. Basically, I have established with this analysis a baseline expression of pharmacologically relevant genes with age, however, I do not know, if these patterns are actually maintained in the course of a disease, and thus if the knowledge generated with this analysis would be transferable in the context of illness, or even, if it would only be pertinent in certain diseases.

The third point is around the fact that for the drugs where we seemed to be able to explain their lack of efficacy in children, and/or their unsafe profile in the elderly, their drug targets are predominantly protein complexes. This imposes one extra level of complexity when trying to rationalise differential drug responses, since, as mentioned in chapter 4, not all subunits change their expression with age in the same direction. Despite that fact, I decided to consider that the overall receptor would increase or decrease its expression based on the strongest trend among the expression of the individual subunits, and use that final trend to rationalise drug response in the context of age. However, different drugs may have different affinities to different GABA A receptor subunits, for instance, as exemplified in a recent study where the affinities of clobazam and N-desmethylclobazam were measured against different GABA A receptor subtypes [153]. This difference in

affinity can thus affect the final drug response, and it is difficult to predict the true effect and contribution of these differences in subunit expression with age in the final therapeutic effect.

Additionally, this analysis investigated how the pharmacologically relevant genes change their expression with age, but not how their signalling pathways may be affected by it. For instance, in case of G protein-coupled receptors, like the adrenergic receptor family, it would also be important to study how the G proteins change their expression with age. Although, access to such information, may hinder the interpretation of results, specially if the G-protein expression with age does not follow the same pattern as the receptor it is coupled to. For example, if the receptor is higher expressed in a given aged population, but not the G-protein, should one assume that the downstream signal is the same independent of the levels of the receptor? Or on the contrary, should one expect a decrease in effect, since some of the receptors that the drug may bind to, may not have a G-protein coupled to it to transmit the signal?

Finally, the last point is related with the provenance of the list of drugs with reported differential drug responses. As described in chapter 2, the data was collected from the literature, with the inherited hindrances. However, as mentioned before, this data is dynamic, and lists compiled at different times may reveal different patterns. For instance, SSRIs were considered safe and effective in the paediatric populations until 2004, when the FDA reviewed clinical trial data and decided to impose a boxed warning against the members of this class, due to the increased risk of suicides detected among the paediatric population when treated with this type of medicine [63]. Consequently, the results of this thesis can only be interpreted in the context of the drugs analysed. Moreover, the literature only reflects the most known cases, thus it would be interesting to do a similar analysis on hospital records, and confirm if the main trends detected in chapter 2 in terms of therapeutic indications, and consequently in chapter 3 in terms of the prevalence of drug targets, remain the same.

5.2 Looking beyond age

5.2.1 Gender-related differential drug responses

Besides age-related differential drug responses, one may also be confronted with gender-related differential drug responses. Despite of its awareness, historically, women's health research has focus on diseases affecting fertility and reproduction. Consequently, there are not as many studies on gender differences on drug response in other disease areas. However, there are some known differences on the incidence of diseases between genders, and also in drug activity. For example, women are more prone to depression, i.e., the burden of depression is 50% higher for females than males, as well as being more susceptible to osteoporosis, breast cancer, anxiety disorders, migraines, Alzheimer's disease and other dementias [60; 154; 155; 156]. In terms of differences in drug activity, tricyclic antidepressants seem more effective in men, while SSRIs are more effective in women, and class I and III antiarrhythmics require a lower dose in women to guarantee a safe therapeutic effect [157]. In addition, women are 50 to 75% more likely to experience adverse reactions, specially with medications known to prolong the QT interval [157]. Some of these differences in drug activity, are explained by physiological differences. For example, the fact that women have a slower renal clearance, results in renally-excreted medications, such as digoxin, to remain longer in the body, and thus requiring an initial lower loading-dose regime to guarantee a safe response. Likewise, the fact that women have on average more adipose tissue than men, contributes to prolonged action of lipophilic medications, such as benzodiazepines, and neuromuscular blockers in women [157]. In terms of molecular differences, though, very little is known, despite being hypothesised that the fact that women have an enhanced effectiveness of beta blockers and opioids, may be connected to pharmacodynamic differences between genders [157].

In the beginning of my PhD, gender was thus a factor which I've considered to investigate when exploring differences in expression patterns of pharmacologically relevant genes in differently aged populations. However, initial preliminary studies, with a human brain expression dataset, indicated that gender, opposed to age, explained very little of the variance detected in the data. This observation was also corroborated by the group of one of our collaborators, Dr. Nenad Sestan from

the Medical School of Yale University in New Haven, USA, which arrived to the same conclusion, when studying an independent human brain expression dataset [158]. Therefore, I decided not to pursue that avenue, given that the initial data seemed to suggest very few differences in gene expression between male and female subjects, rendering it unlikely for this type of analysis to help rationalise gender-related differential drug responses. This observation seemed to be extended to animal models. At least in mouse liver, as discussed in chapter 4, when comparing male and female differences in the expression of pharmacologically relevant genes as a function of age, only 2 genes out of 931 had a significant different profile (Fmo3 and Sult2a1)[1]. This overall trend, seems to suggest that pharmacokinetic studies in both genders, may provide a bigger understanding on how to minimise gender-related differential drug responses in the future, than a closer look into pharmacodynamic differences.

5.2.2 Race-related differential drug responses

Race is another factor which may also influence drug response. Due to its sensitive nature, although, it is less discussed in the community, and rarely provided as a phenotypic annotation in functional genomic experiments. Thus, rendering it impossible at the moment, to study the expression of pharmacologically relevant genes as a function of race, in the same fashion as employed in this thesis to study the effects of age. Consequently, it is not straightforward to estimate how much different ethnical backgrounds actually influence drug target and metabolic machinery, and ultimately drug response.

There have been a couple of recent studies, though, addressing this topic, and raising awareness to the importance of gaining a better understanding of pharmacokinetic and pharmacodynamic differences in populations with different ethnic backgrounds. For example, when patients with coronary artery disease were treated with calcium channel blockers (verapamil and trandolapril), or with β-adrenergic receptor inhibitors, such as atenolol, the practitioners noticed that

[1]The comparison of experiment E-MTAB-2782 with E-GEOD-21716 in combination with E-GEOD-21224 does not allow one to compare mRNA levels between genders per age groups, just, the overall profile or mRNA ratios, given that the samples were not hybridise in the same array type, nor at the same time.

hispanic patients achieved a better control of blood pressure than non-hispanic ones [159]. This is not the only study which suggests that ethnicity influences the pharmacological action of antihypertensive agents [160]. In addition, these observations have been extended to and reproduced in other therapeutic areas as well. For instance, a recent trial conducted in Japan and in the USA, which used the same doses and schedules for the administration of carboplatin plus paclitaxel for the treatment of small cell lung cancer, concluded that asian patients were more susceptible to experience neutropenia[1] than caucasian patients [161]. Furthermore, clinical trials of cetuximab, indicated for the treatment of non-small cell lung cancer, and of bevacizumab, for the treatment of stomach cancer, have shown that these agents are only active in caucasians [161]. Differences in sensitivity at the drug target level have been suggested to explain the difference in responses, however, the concrete mechanism has yet to be clarified.

In the future, as the community grows more aware of the need to understand this race-related drug responses, and as the access to the ethnical background as a phenotypic annotation in functional genomic or proteomic experiences becomes available, one may attempt to study pharmacodynamic differences between populations with different ethnic backgrounds, which may illuminate some of the differences previously discussed.

5.3 Looking beyond genomics

5.3.1 Proteomics

As mentioned in the introduction of chapter 4, to study the ontogeny of pharmacologically relevant genes, one could also use proteomic profiling studies reporting on the levels of the proteins of interest at different developmental stages. At the present time, however, the number of available studies in public archives is rather modest, for example, only 93 human experiments have been deposited so far in ArrayExpress[2], and similarly with the data from functional genomic studies, only a minority of those have annotations about age and gender. However, as tech-

[1] A granulocyte disorder characterised by an abnormally low number of neutrophils.
[2] Accessed on the 31st of July of 2014.

nology progresses and more protein array studies, and mass spectrometry based proteomic studies become available in public repositories, in five to ten years time, it should be possible to repeat the same type of analysis, but at the protein level, and compare how much of the switches in the proteome with age correlate with what was observed at the transcriptome level, and conversely, how much it is actually buffered by the intrinsic complexity of biological networks.

5.3.2 Metabolomics

Metabolomic studies are another type of experiments, which could potentially be used to study how pharmacologically relevant pathways may differ between differently aged populations, by measuring the levels of the metabolites in a given tissue at a given developmental stage. This information could then be used to compare the flux of pathways in children versus adults, for instance, and hypothesise on how the different levels of pharmacologically relevant mRNAs or proteins actually influence the pathway output, or if the network is able to buffer those pharmacodynamic differences. However, this type of experiments are only now becoming available in the public domain. For example, MetaboLights, a repository for metabolomic studies hosted at the EBI [162], has only 13 human metabolic experiments deposited so far[1]. Additionally, at the moment, there is very little meta-data associated with these experiments, i.e., phenotypic annotation of the subjects (age, gender, disease status, *etc*) from where the sample was taken to measure the level of metabolites, is not captured for the majority of experiments, requiring the user to extract that information from the publication (if available). With time, though, and once the minimal deposition standards for a metabolic experiment [163] become routinely employed, and more experiments become available with phenotypic data, one may be able to, in 10 to 15 years time, re-analyse the deposit experiments in the context of age-related differential drug responses, similarly to what has been presented in this thesis with functional genomic experiments.

[1] Accessed on the 31st of July of 2014.

Chapter 6

Conclusion and future plans

The transfer of medicines between adult and paediatric populations, or between adult and geriatric populations is not a straightforward process. The lack of understanding of how some of the factors which influence drug response change with age, such as the physical size, weight, lipid fraction, drug target and metabolic machinery expression, *etc.*, has been a hurdle in the safe translation of medicines from one population to another, giving rise to a few documented cases of differential drug responses: increased number of side effects, or lack of efficacy.

In this thesis, I have systematically analysed the ontogeny of drug targets, metabolic enzymes and transporters in five relevant human tissues, in order to study if their changes in expression as a function of age, would potentially explain some of the differential drug responses reported in the literature. In parallel, I have also investigated on how those changes translated to animal models (macaque, rat and mouse) commonly used in drug discovery, with the aim of identifying the best age-driven animal model.

On that account, as described in chapter 2, I have probed the medical literature, in order to first compile a list of drugs with reported differential drug responses when transferred from the adult population to the paediatric and/or geriatric population. The effort resulted in the identification of 106 drugs with reported age-related differential drug responses (DDRs) and 43 drugs without. A closer look to this list, revealed that the majority of drugs have shown more safety issues when transferred from one aged population to another, than a lack of efficacy. In terms of therapeutic indications, drugs modulating the nervous system were more prevalent

among the drugs with reported DDRs when given to paediatric populations, while among the drugs with reported DDRs when given to the geriatric populations, the predominance was of drugs which interact with the cardiovascular system. Additionally, drugs in the control group seemed to be newer than drugs with reported DDRs, which suggests an effort from the community to try to address the issue and minimise the incidence of such differential drug responses.

In chapter 3, after having compiled an up-to-date and comprehensive list of pharmacologically relevant human genes responsible for the drugs' efficacy and safety, and an up-to-date map of drug targets, metabolic enzymes and transporters against FDA approved drugs, I have explored the distribution of those drug targets, metabolic enzymes and transporters families, in the context of age-related differential drug responses. In particular, I have identified the ones predominantly associated with reported safety issues and/or a lack of efficacy when comparing different aged populations. Overall, among drugs which have shown a lack of efficacy and/or an increased incidence of side effects in children, it was interesting to observe that the majority of their drug targets were targets from the serotonin pathway (receptors and transporter), and to a certain degree, the GABA A receptor. In terms of the metabolic enzymes, only members of the UGT family stood out. On the other hand, for drugs which have shown a lack of efficacy and/or an increased incidence of side effects in the elderly, the prevalent drug targets were the GABA A receptor, sulfonylureas receptor, voltage-gated L-type calcium channel, histamine receptor, opioid receptors and β-adrenergic receptors, with no noteworthy pattern among metabolic enzymes and transporters. Consequently, in chapter 4, those targets and metabolic enzymes were the ones whose expression was closely followed in relevant human and animal tissues.

However, to study the expression of those targets, I had to first query gene expression public archives for relevant expression datasets in human, macaque, rat and mouse, which met the necessary requirements, i.e., samples were from healthy subjects, with age and age units clearly indicated, and from one of the five primary tissues (brain, heart, lungs, liver and kidney). A total of 91 publicly available expression datasets were identified, which allowed me to study the ontogeny of pharmacologically relevant genes in the context of age-related differential drug responses. The majority of the expression sets contained samples from

brain and liver tissues, and mainly from human and mouse specimens. However, among the animal models data, not a single experiment covered the full life cycle of the animal, therefore, we decided to generated a new female liver mouse expression dataset to complement the data in the public domain. This expression set was analysed together with the data from the public domain, and allowed me to compare the expression profile of pharmacologically relevant genes between male and female mouse in liver, besides providing information on the expression profiles of pharmacologically relevant genes at all stages of the mouse life cycle. Overall, among the human experiments, only in the brain expression datasets, I observed statistically and reproducible significant differential expression as a function of age of drug targets. In animal models, this trend was repeated and extended to other tissues as well. In terms of metabolic enzymes, the majority seemed to increase their expression with age in all tissues and species. Finally, among the drug targets associated with drugs with reported DDRs when given to either paediatric or geriatric populations, it was rather curious to observe an overall decrease in expression with age among ion channels, such as GABA A receptor, and voltage-gated calcium channels in human brain, but detect the opposite trend in rodents. Additionally, it was interesting to observe that some serotonin receptors, like serotonin receptor 1D (HTR1D), increased their expression in the human brain, but in both rat and mouse, followed the opposite trend, and decreased their expression with age.

From the general discussion, however, it became clear that for only 17% (n = 18) of drugs with reported differential drug responses, the identified expression patterns of pharmacologically relevant genes with age, contributed to rationalise why there was a difference in response. In short, the results seem to suggest that one of the reasons why a drug may lack efficacy in children, for instance, may be because the dose given was not high enough to occupy all the expressed receptors, and thus fully activate the necessary downstream signals for an effective therapy. Likewise, this trend also seemed to suggest, that for geriatric populations, a smaller dose may be required to prevent over inhibition or over activation of these drug targets. For the remaining 83% of drugs with reported DDRs ($n = 88$), the results of this systematic analysis, do not seem to contribute to increase our understanding of DDRs. Specially for the cases whose drug target increased their expression with

age, given that it was a pattern also present in the control group.

I have also alluded to the main pitfalls of this approach in the general discussion, and hypothesised on the reasons why this analysis is unable to explain more differential drug responses. Briefly, the data collected has many confounders, such as gender, race, medical history, diet, *etc.*, contributing to increase the noise in the data and mask potential age-related signals. For example, in human heart, contrary to what has been described in the literature [13], I was unable to detect β-adrenergic receptors decreasing their expression with age, possibly due to the presence of those confounder factors. In addition, at least in human liver, the expected changes in expression with age happen in a very short time-frame, making it impossible to detect the signals described in the literature (i.e. switch of CYP3A7 to CYP3A4 [164]), given that I didn't have access to enough samples covering those age periods. Alternatively, for the majority of drugs, the physiological differences may affect more the pharmacokinetics of the drug, and consequently its therapeutic effect, than potential pharmacodynamic changes of the drug target. Furthermore, this systematic analysis on the ontogeny of pharmacologically relevant genes has a couple of limitations. For instance, what if the detected changes in expression with age, are not valid in the context of a disease? Also, some of the signals detected were just identified in a few protein complex subunits, mainly from ion channels, and as explained in chapter 4, I decided to consider that the expression of the complete receptor would increase or decrease with age based on the behaviour of only those subunits, and on the most prevalent trend among them. However, what if a drug has more affinity to a subunit whose expression does not change with age, than to the one that does? How would that affect the final drug response? Another hindrance, which I have not explored, is the effect of age in the downstream signalling pathways, i.e., do the G_i/G_o coupled proteins to the serotonin 1a receptor (HTR1A), for instance, also decrease their expression with age? What happens if not? Does the fact that the levels of HTR1A are higher in children still influence the final drug response, if the levels of the G_i/G_o coupled proteins do not alter? Lastly, this project was build on the information available in the medical literature, which may not be an accurate, and up to date source of information in terms of DDRs, thus, would these results still be reproduced if the source of information would be hospital records instead?

This project could, however, be further explored and developed, which could result in the clarification of some of the issues discussed above. For instance, to address the point made about how this baseline expression of pharmacologically relevant genes with age could change in the context of a disease, one could study the expression of those genes in non-healthy subjects as well, where controls are available, and compare the differences. For example, analysing experiments E-GEOD-35974 or E-GEOD-35977 with adult and elder data from patients with and without schizophrenia or bipolar disorder could be a good starting point to generate new perspectives on this issue.

Likewise, to tackle the fact that subunits in protein complexes may alter their expression differently with age, which makes it difficult to predict how the overall behaviour of the complex is affected by it, and thus drug response, one could test the affinity of relevant drugs for this project against different configurations of the receptors. In particular, one could do radioligand assays to test the affinity of drugs against GABA A receptors containing the alpha 5 subunit (GBRA5) and the gamma 1 (GBRG1) or gamma 2 (GBRG2) subunits. With this kind of assay, one could then see if the drugs which have shown a lack of efficacy in children, or increased side effects in the elderly have a higher affinity for this receptor configuration, which would support the hypothesis that maybe in this case, one needs to increase the dose in children and decrease the dose in the elderly to compensate for this pharmacodynamic changes. Similarly, it would be interesting to perform identical assays with different configurations of the glutamate [NMDA] receptor, specially the ones containing the GRIN2A, and/or the GRIN3A subunits, and verify if the drugs in our database, which belong to the control group, have a low affinity for these subunits which change with age. However, these proposed experiments can be quite time consuming and challenging to set up. For example, in case of the GABA A receptor, given that there are at least two commercial human cell lines with the configuration of interest, i.e. $\alpha_5/\beta_2/\gamma_2$ (Merck: CYL3073 — PrecisION^{TM} hGABAA $\alpha_5/\beta_2/\gamma_2$-HEK Recombinant Stable Cell Line; ChanTest: Human GABAA ($\alpha_5/\beta_2/\gamma_2$) Receptor Cell Line), it would be possible to set up a radioligand assay, and collect affinity data within approximately three months, and thus evaluate how different drugs have different affinities for different subtypes of the GABA A receptor, comparable to what was done by Jensen *et al*

[153] or explored by Fitzgerald *et al* [165]. However, if there is no cell line with the intended target and/or no radioligand to probe the target, like in the case of glutamate [NMDA] receptor, then it could take up to one year and half to set up the assay, without any guarantees that affinity data could actually be collected in the end.

Similarly, to address the last pitfall of this project design, i.e. the information on differential drug response was collected from the literature, and may therefore be bias towards the most studied cases, and not provide a complete picture of the reality, one could repeat the same analysis in ten years time using electronic hospital records from projects such as the eHospital initiative [166], and compare how the list of drugs may differ and change over time, and in light of that, investigate if the results of this project are still relevant then.

In addition, the results of this project also generated some interesting fundamental biology research questions, which could also be further explored. For example, within the same family of membrane receptors, such as the case of the adrenergic receptors, why does α-adrenergic receptor 2A (ADRA2A) has a different expression pattern with age? How does the regulatory machinery (promoters, microRNA, *etc*) differ within the same family of receptors? Same applies to the expression of different carbonic anhydrase isoforms in mouse liver, what are the regulatory elements which play a role in the decrease of expression with age of Car1, Car2, and Car3, and in the increase of Car5a? A better understanding on the regulatory machinery which controls the expression of pharmacologically relevant genes may gives us news perspectives on age modulation.

Likewise, when comparing the human expression patterns with the ones detected in the mouse data, the differences identified also prompt a series of research questions about the evolution of regulatory elements. This subject has been the focus of recent studies [167; 168], and so far substantial divergence between human and mouse tissue-specific transcription factors has been identified, for instance. However, it has been rather challenging to link some of the variations in regulation with phenotypic differences. For example, a recent study on the regulation of the expression of the GABA A subunits, comparing human and mouse patterns, points exactly to that fact, where for certain genes there are differences in the regulatory elements but they don't necessarily translate to differences in the

phenotype [151]. The authors of this study suggest that these "inconsistencies" illustrate the impressive ability of higher order network properties to buffer the variation of their constitutes (i.e. a 10-fold difference in transcript may be muted at the protein level).

In conclusion, this systematic analysis on the ontogeny of pharmacologically relevant genes, despite its limitations, brings a new perspective on age-related differential drug responses, and challenges some of the approaches used when extrapolating knowledge from adult studies to other aged populations.

Appendix A

Age-related words:

adolescence
adolescent
adolescents
adult
adults
age
babies
baby
child
childhood
children
elderly
geriatric
infant
infants
juvenille
kid
kids
menarche
neonate
neonates
newborn
newborns
peadiatric
pediatric
puberty
teenager
teenagers
young

Drug-related words:

differential
dose
dosing
effect
effective
response
safe

Appendix A

Stop-words:

1
2
3
4
5
6
7
8
9
a
about
above
across
after
again
against
all
almost
alone
along
already
also
although
always
am
among
an
and
another
any
anybody
anyone
first
for
four
from
full
fully
further
furthered
furthering
furthers
g
gave
general
generally
get
gets
give
given
gives
go
going
good
goods
got
great
greater
greatest
group
grouped
grouping
groups
h
had
newer
anything
anywhere
are
area
areas
aren't
around
as
ask
asked
asking
asks
at
away
b
back
backed
backing
backs
be
became
because
become
becomes
been
before
began
behind
being
beings
below
best
hadn't
has
hasn't
have
haven't
having
he
he'd
he'll
he's
her
here
here's
hers
herself
high
higher
highest
him
himself
his
how
how's
however
i
i'd
i'll
i'm
i've
if
important
in
interest
ought
better
between
big
both
but
by
c
came
can
can't
cannot
case
cases
certain
certainly
clear
clearly
come
could
couldn't
d
did
didn't
differ
different
differently
do
does
doesn't
doing
don't
done
interested
interesting
interests
into
is
isn't
it
it's
its
itself
j
just
k
keep
keeps
kind
knew
know
known
knows
l
large
largely
last
later
latest
least
less
let
let's
lets
like
likely
really
down
downed
downing
downs
during
e
each
early
either
end
ended
ending
ends
enough
even
evenly
ever
every
everybody
everyone
everything
everywhere
f
face
faces
fact
facts
far
felt
few
find
finds
long
longer
longest
m
made
make
making
man
many
may
me
member
members
men
might
more
most
mostly
mr
mrs
much
must
mustn't
my
myself
n
necessary
need
needed
needing
needs
never
new
since

newest
next
no
nobody
non
noone
nor
not
nothing
now
nowhere
number
numbers
o
of
off
often
on
once
one
only
open
opened
opening
opens
or
order
ordered
ordering
orders
other
others
they'll
they're
they've
thing
things
think
thinks
this
those
though
thought
thoughts
three
through
thus
to
today
together
too
took
toward
turn
turned
turning
turns
two
u
under
until
up
upon
us
use
our
ours
ourselves
out
over
own
p
part
parted
parting
parts
per
perhaps
place
places
point
pointed
pointing
points
possible
present
presented
presenting
presents
problem
problems
put
puts
q
quite
r
rather
used
uses
v
very
w
want
wanted
wanting
wants
was
wasn't
way
ways
we
we'd
we'll
we're
we've
well
wells
went
were
weren't
what
what's
when
when's
where
where's
whether
which
while
who
right
room
rooms
s
said
same
saw
say
says
second
seconds
see
seem
seemed
seeming
seems
sees
several
shall
shan't
she
she'd
she'll
she's
should
shouldn't
show
showed
showing
shows
side
sides
who's
whole
whom
whose
why
why's
will
with
within
without
won't
work
worked
working
works
would
wouldn't
x
y
year
years
yet
you
you'd
you'll
you're
you've
your
yours
yourself
yourselves
z
small
smaller
smallest
so
some
somebody
someone
something
somewhere
state
states
still
such
sure
t
take
taken
than
that
that's
the
their
theirs
them
themselves
then
there
there's
therefore
these
they
they'd

Table 2: List of drugs with reported age-related differential drug responses (DDR). The presence of an asterisk (*) indicates that the drug has been discontinued, and is no longer available in the U.S.A. market.

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Acebutolol	CHEMBL642	For the treatment of hypertension	C07AB04	1984	Not effective in the elderly	PMID:17996666	DT	ADRB1	CHEMBL2094118
Acetaminophen	CHEMBL112	Antipyretic and analgesic	N02BE01, N02BE51, N02BE71	1968	Lower risk of hepatoxicity in infants	PMID:12761605	DT	FAAH, PTGS1, PTGS2, TRPV1	CHEMBL2243, CHEMBL2094253, CHEMBL2094253, CHEMBL4794
							ME	CYP1A1, CYP2A6, CYP2C8, CYP2C9, CYP2D6, CYP3A4, CYP3A5, GST, UGT1A6, UGT1A9, SULT1A1, CYP2E1, CYP1A2	CHEMBL2231, CHEMBL5282, CHEMBL3721, CHEMBL3397, CHEMBL289, CHEMBL340, CHEMBL3019, -, CHEMBL1743316, CHEMBL1743319, CHEMBL1743291, CHEMBL5281, CHEMBL3356
							TP	ABCB1, ABCC1, ABCC2, ABCC3, ABCC4	CHEMBL4302, CHEMBL3004, CHEMBL5748, CHEMBL5918, CHEMBL1743128
Albuterol	CHEMBL714	For the treatment of asthma	R03AC02, R03CC02	1981	Not effective in the elderly	PMID:17996666	DT	ADRB2	CHEMBL2094118
Alfentanil	CHEMBL634	Analgesic	N01AH02	1986	Equally effective in old people and adults when the dose is adjusted	PMID:17996666	DT	OPRM1	CHEMBL233
Amiodarone	CHEMBL633	Antiarrhythmic	C01BD01	1985	Prone for side effects in the elderly	PMID:22376048	DT	KCNH2	CHEMBL240
Amitriptyline	CHEMBL629	Antidepressant	N06AA09	1961	Prone for side effects in children and old people	PMC:3565674	DT	SLC6A2, SLC6A4	CHEMBL222, CHEMBL228

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
							ME	CYP2B6, CYP3A5, CYP3A4, CYP2E1, CYP1A2, UGT, CYP2C8, CYP2C9, CYP2C19, CYP2D6	CHEMBL4729, CHEMBL3019, CHEMBL340, CHEMBL5281, CHEMBL3356, -, CHEMBL3721, CHEMBL3397, CHEMBL3622, CHEMBL289
Amlodipine	CHEMBL1491	For the treatment of hypertension	C08CA01	1992	Prone for side effects in the elderly	PMID:17996666	DT	CACNA1C, CACNA1S, CACNA1F, CACNA1D	CHEMBL2095229, CHEMBL2095229, CHEMBL2095229, CHEMBL2095229
Argatroban	CHEMBL1166	Anticoagulant	B01AE03	2000	Equally effective in adults and old people	PMID:17996666	DT	F2	CHEMBL204
Aspirin	CHEMBL25	Antipyretic and analgesic and anti-inflamatory	A01AD05, B01AC06, N02BA01, N02BA51, N02BA71	1950	Prone for side effects in the elderly	PMID:18218287	DT	PTGS2, PTGS1	CHEMBL2094253, CHEMBL2094253
							ME	SUL, NAT, GST, CYP2C8, CYP2C9, CYP2C19, UGT	-, -, -, CHEMBL3721, CHEMBL3397, CHEMBL3622, -
Atomoxetine	CHEMBL641	ADHD	N06BA09	2002	Equally effective in children and adults	PMC:3513681	DT	SLC6A2	CHEMBL222
Atorvastatin	CHEMBL1487	Cholesterol lowering agent	C10AA05	1996	Equally effective in children and adults and old people	PMID:12915827	ME	CYP2D6	CHEMBL289
							DT	HMGCR	CHEMBL402
Atracurium	CHEMBL1360	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC04	1983	Equally effective in adults and old people	PMID:17996666	ME	CYP3A4	CHEMBL340
							TP	SLCO1B1, ABCG2	CHEMBL1697668, CHEMBL5393
							DT	CHRNG, CHRNE, CHRND, CHRNA1, CHRNB1	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Atropine	CHEMBL517712	Bradicardia and Bronchial secretions lowering agent	A03BA01, S01FA01	1960	Increased risk of side effects in children and old people	PMID:24105354	DT	CHRM1, CHRM2, CHRM3	CHEMBL216, CHEMBL211, CHEMBL245
							ME	CYP1A2	CHEMBL3356
Benazepril	CHEMBL838	For the treatment of hypertension	C09AA07	1991	Equally effective in adults and old people	PMID:17996666	DT	ACE	CHEMBL1808
Benzonatate	CHEMBL1200369	Antitussive	R05DB01	1958	Increased risk of death in children under 10 years old	PMC:3526931	DT		CHEMBL2364682
Benzyl Alcohol	CHEMBL720	For the treatment of head lice	Unclassified	2009	Prone for side effects in children	ISBN:978-92-4-156343-7			
Betamethasone	CHEMBL632	For the treatment of eczema	A07EA04, C05AA05, D07AC01, D07XC01, H02AB01, R01AD06, R03BA04, S01BA06, S01CB04, S02BA07, S03BA03	1961	Supresses adrenal activity in children under 12 years old;	PMID:21086850	DT	NR3C1	CHEMBL2034
							ME	CYP19A1, CYP3A4	CHEMBL1978, CHEMBL340
Bortezomib	CHEMBL325041	Anticancer	L01XX32	2003	Lower risk of neurotoxicity for children	PMID:22490765	DT	, PSMB5	CHEMBL2364701, CHEMBL2364701
Bumetanide	CHEMBL1072	Diuretic	C03CA02	1983	Prone for side effects in the elderly, not very effective to lower heart pressure	PMID:12208585	DT	SLC12A1	CHEMBL1874
Buspirone	CHEMBL49	For the treatment of anxiety	N05BE01	1986	Not effective or safe in children between 6 and 17 years old	PMID:21086850	DT	HTR1A	CHEMBL214
Calcitriol	CHEMBL846	For the treatment of hyperparathyroidism associated with Chronic Kidney Disease (CKD)	A11CC04, D05AX03	1978	Equally effective in children and adults	PMID:23381010	DT	VDR	CHEMBL1977

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Carvedilol	CHEMBL723	For the treatment of chronic heart failure (CHF)	C07AG02	1995	Not effective in children and adolescents, but effective in the elderly	PMID:22025597	DT	ADRB3, ADRB2, ADRB1, ADRA1D, ADRA1B, ADRA1A	CHEMBL2094118, CHEMBL2094118, CHEMBL2094118, CHEMBL2094251, CHEMBL2094251, CHEMBL2094251
Chloram-phenicol	CHEMBL130	Broad spectrum antibiotic	D06AX02, D10AF03, G01AA05, J01BA01, S01AA01, S02AA01, S03AA08	1950	Causes the grey baby syndrome in premature and newborn infants	PMID:20008037	DT		CHEMBL2363135
Chlor-diazepoxide	CHEMBL451	For the treatment of anxiety	N05BA02	1960	Prone for side effects in the elderly	PMID:18218287	ME	CYP2C8, CYP2C9, CYP3A4, CYP2C19, UGT2B7	CHEMBL3721, CHEMBL3397, CHEMBL340, CHEMBL3622, CHEMBL4370
							DT	GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRQ, GABRA1, GABRA2, GABRA3, GABRA4, GABRA5, GABRA6, GABRB1	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Chlor-pheniramine	CHEMBL505	Antihistaminic	R06AB04, R06AB54	1950	Prone for side effects in the elderly	PMID:18218287	ME	CYP2D6, CYP3A4	CHEMBL289, CHEMBL340
							DT	HRH1	CHEMBL231
							ME	CYP2D6, CYP3A5, CYP3A7, CYP3A4	CHEMBL289, CHEMBL3019, -, CHEMBL340
							TP	ABCC1	CHEMBL3004

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Chlorpropamide	CHEMBL498	For the treatment of type II diabetes	A10BB02	1958	Prone for side effects in the elderly	PMID:22376048	DT	KCNJ11, ABCC8	CHEMBL2096972, CHEMBL2096972
							ME	CYP2C19, CYP2C9	CHEMBL3622, CHEMBL3397
Cinacalcet	CHEMBL1201284	For the treatment of hyperparathyroidism associated with Chronic Kidney Disease (CKD)	H05BX01	2004	Prone for side effects in children	PMID:23381010	DT	CASR	CHEMBL1878
Cisapride*	CHEMBL1729	Gastroprokinetic agent	A03FA02	1993	Not effective in children when treating gastro-oesophageal reflux	PMID:20393933	DT	HTR4	CHEMBL1875
Cisatracurium	CHEMBL1201248	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC11	1995	Slower response in old people	PMID:17996666	ME	CYP3A4	CHEMBL340
							DT	CHRNG, CHRND, CHRNB1, CHRNA1, CHRNE	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Citalopram	CHEMBL549	Antidepressant	N06AB04	1998	Not effective in children, and prone for side effects	PMID:22064376	DT	SLC6A4	CHEMBL228
Clomipramine	CHEMBL415	For the treatment of Obsessive-Compulsive Disorder (OCD)	N06AA04	1989	Prone for side effects in children	PMC:3513681	DT	SLC6A4	CHEMBL228
Clozapine	CHEMBL42	For the treatment of schizophrenia	N05AH02	1989	Equally effective in children and adults	PMC:3513681	ME	CYP2C19, GST, UGT, CYP2D6, CYP1A2, CYP3A4	CHEMBL3622, -, -, CHEMBL289, CHEMBL3356, CHEMBL340
							TP	SLCO6A1	-
							DT	HTR2A, DRD4, DRD3, DRD2	CHEMBL224, CHEMBL2331075, CHEMBL2331075, CHEMBL2331075
Codeine	CHEMBL485	Antitussive	N02AA59, N02AA79, R05DA04	1952	Not effective in children	PMID:22895922	DT	OPRM1, OPRK1, OPRD1	CHEMBL2095181, CHEMBL2095181, CHEMBL2095181

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
							ME	CYP2D6, CYP3A4, CYP3A5, CYP3A7, UGT	CHEMBL289, CHEMBL340, CHEMBL3019, -, -
Colchicine	CHEMBL107	For the treatment of familial mediterranean fever (FMV)	M04AC01	1961	Equally effective in children and adults	PMID:22738946	DT	, TUBB, TUBB1, TUBB2A, TUBB2B, TUBB3, TUBB4A, TUBB4B, TUBB6, TUBB8	CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182, CHEMBL2095182
							ME	CYP3A4, CYP1A1, CYP2B6, CYP2E1, CYP2C9, CYP2C8	CHEMBL340, CHEMBL2231, CHEMBL4729, CHEMBL5281, CHEMBL3397, CHEMBL3721
							TP	ABCB1, ABCB11	CHEMBL4302, CHEMBL6020
Colesevelam Hydrochloride	CHEMBL1201473	Cholesterol lowering agent	C10AC04	2000	Equally effective in adults and old people	PMID:22530704	DT		CHEMBL2364175
Cosyntropin	CHEMBL2103784	For the treatment and diagnosis of primary adrenal insufficiency	H01AA02	1970	Equally effective in children and adults	PMID:23740534	DT	MC2R	CHEMBL1965
Cyclizine*	CHEMBL648	Antihistaminic	R06AE03, R06AE53	1982	Prone for side effects in the elderly	PMID:18218287	DT	HRH1	CHEMBL231
							ME	CYP2C9	CHEMBL3397
Cysteamine	CHEMBL602	For the management of nephropathic cystinosis	A16AA04	1994	Equally effective in children and adults	PMID:23075296	DT		CHEMBL2364039
							ME	GST, NAT	-, -

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Desflurane	CHEMBL1200733	Anaesthetic	N01AB07	1992	Increased incidence of respiratory adverse events in children	PMID:22025597	DT	GABRP, GABRQ, GABRG3, GABRG2, GABRG1, GABRE, GABRD, GABRB3, GABRB2, GABRB1, GABRA6, GABRA5, GABRA4, GABRA3, GABRA2, GABRA1, KCNK9, KCNK3, KCNK2, KCNK18, KCNK10, GLRB, GLRA1	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2321614, CHEMBL2321613, CHEMBL2321615, CHEMBL2331042, CHEMBL2331041, CHEMBL2363052, CHEMBL2363052
Dextro-methorphan	CHEMBL52440	Antitussive	N07XX59, R05DA09	1957	Not effective in children	PMID:15495019	DT	SIGMAR1, GRIN2A	CHEMBL287, CHEMBL1972
							ME	CYP2C19, CYP2B6, CYP2C9, CYP2C8, CYP46A1, CYP3A7, CYP3A5, CYP3A4, CYP2E1, CYP2D6	CHEMBL3622, CHEMBL4729, CHEMBL3397, CHEMBL3721, -, -, CHEMBL3019, CHEMBL340, CHEMBL5281, CHEMBL289

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Diazepam	CHEMBL12	For the treatment of anxiety	N05BA01	1963	Prone for side effects in the elderly	PMID:18218287	DT	GABRA4, GABRA6, GABRB1, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRQ, GABRA1, GABRA2, GABRA3, GABRA5	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
							ME	NAT, CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C18, CYP2C19, CYP2E1, CYP3A4, CYP3A5, CYP3A7, UGT	-, CHEMBL3356, CHEMBL4729, CHEMBL3721, CHEMBL3397, CHEMBL2408, CHEMBL3622, CHEMBL5281, CHEMBL340, CHEMBL3019, -, -
Digoxin	CHEMBL1751	For the treatment of atrial fibrilation or heart failure	C01AA05	1975	Prone for side effects in the elderly	PMID:18218287	DT	, ATP1A1, ATP1A2, ATP1A3, ATP1A4	CHEMBL2095186, CHEMBL2095186, CHEMBL2095186, CHEMBL2095186, CHEMBL2095186
Dihydro-ergotamine	CHEMBL1732	For the treatment of migranes	N02CA01, N02CA51	1946	Not effective in children and adolescents orally	PMID:16061583	DT	HTR1D	CHEMBL1983
							ME	CYP3A4	CHEMBL340
Diltiazem	CHEMBL23	For the treatment of angina	C08DB01	1982	Prone for side effects in the elderly	PMID:18218287	DT	CACNA1S, CACNA1F, CACNA1C, CACNA1D	CHEMBL2095229, CHEMBL2095229, CHEMBL2095229, CHEMBL2095229
Diphen-hydramine	CHEMBL657	Antihistaminic	D04AA32, R06AA02, R06AA52	1946	Not effective in children	PMID:22962927	DT	HRH1	CHEMBL231

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
							ME	CYP2D6	CHEMBL289
Diphenoxylate	CHEMBL1201294	For the treatment of diarrhea	A07DA01	1960	Prone for side effects in the elderly	PMID:18218287	DT	OPRM1	CHEMBL233
Dipyridamole	CHEMBL932	Antithrombotic	B01AC07	1961	Not effective in the elderly	PMID:18218287	DT	SLC29A1, PDE5A	CHEMBL1997, CHEMBL1827
							TP	ABCB1, ABCC1, ABCG2	CHEMBL4302, CHEMBL3004, CHEMBL5393
Dolasetron	CHEMBL1237118	Antiemetic	A04AA04	1997	Equally effective in children and adults	PMID:20434730	DT	HTR3A	CHEMBL1899
							ME	FMO5, FMO3, FMO4, CYP2D6, CYP3A4, FMO2, FMO1	-, -, -, CHEMBL289, CHEMBL340, -, -
Doxacurium*	CHEMBL1237123	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC07	1991	Slower response in old people	PMID:17996666	DT	CHRNG, CHRNE, CHRND, CHRNB1, CHRNA1	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Doxepin	CHEMBL1628227	Antidepressant	N06AA12	1969	Prone for side effects in the elderly	PMID:22376048	DT	SLC6A2	CHEMBL222
Doxercalciferol	CHEMBL1200810	For the treatment of hyperparathyroidism associated with Chronic Kidney Disease (CKD)	H05BX03	1999	Equally effective in children and adults	PMID:23381010	DT	VDR	CHEMBL1977
Duloxetine	CHEMBL1175	Antidepressant	N06AX21	2004	Equally effective in children, adults and old people	PMID:18416656	DT	SLC6A4, SLC6A2	CHEMBL228, CHEMBL222
							ME	CYP1A2	CHEMBL3356
Edrophonium	CHEMBL1104	For the differential diagnosis of myasthenia gravis	Unclassified	1951	Equally effective in old people and adults when the dose is adjusted	PMID:17996666	DT	ACHE	CHEMBL220
Enalapril	CHEMBL578	For the treatment of hypertension	C09AA02	1985	Equally effective in adults and old people	PMID:17996666	DT	ACE	CHEMBL1808

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Ergoloid	CHEMBL2364968	For the treatment of dementia, and recovery after stroke	C04AE01, C04AE51	1953	Not effective in the elderly	PMID:22376048	DT	ADRA1B, ADRA1D, ADRA2A, ADRA2B, ADRA2C, DRD1, DRD5, DRD2, DRD3, DRD4, ADRA1A, HTR1A, HTR1B	CHEMBL2095203, CHEMBL2095203, CHEMBL2095203, CHEMBL2095203, CHEMBL2095203, CHEMBL2096905, CHEMBL2096905, CHEMBL2096905, CHEMBL2096905, CHEMBL2096905, CHEMBL2095203, CHEMBL214, CHEMBL1898
Erlotinib	CHEMBL553	Anticancer	L01XE03	2004	Different side effects in children	PMID:22490765	DT	EGFR	CHEMBL203
Escitalopram	CHEMBL1508	Antidepressant	N06AB10	2002	Equally effective in teenagers, adults and old people, but prone for side effects in the young	PMID:23271894	DT	SLC6A4	CHEMBL228
Etodolac	CHEMBL622	For the treatment of rheumatoid arthritis	M01AB08	1991	Equally effective in children and adults when the dose is adjusted	PMID:21086850	DT	PTGS2	CHEMBL230
Etomidate	CHEMBL681	Anaesthetic	N01AX07	1982	Safe in the elderly when the dose is reduced	PMID:17996666	DT	GABRA2, GABRA1, GABRA4, GABRA5, GABRA6, GABRB1, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRQ, GABRA3	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Exenatide	CHEMBL414357	For the treatment of type II diabetes	A10BX04	2005	Equally effective in children and adults	PMC:3684414	DT	GLP1R	CHEMBL1784

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Felbamate	CHEMBL1094	Antiepileptic	N03AX10	1993	Equally effective in children and adults	PMID:18637038	DT	GRIN3A, GRIN3B, GRIN1, GRIN2D, GRIN2C, GRIN2B, GRIN2A	CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124
Fentanyl	CHEMBL596	For the management of breakthrough pain in cancer patients	N01AH01, N01AH51, N02AB03	1968	Equally effective in old people and adults when the dose is adjusted; Prone for side effects in children	PMID:17996666	DT	OPRM1	CHEMBL233
							ME	CYP3A5, CYP3A7, CYP3A4	CHEMBL3019, -, CHEMBL340
Fesoterodine	CHEMBL1201764	Overactive bladder	G04BD11	2008	Equally effective in children and adults	PMID:23204858	DT	CHRM3, CHRM2	CHEMBL245, CHEMBL211
Fluconazole	CHEMBL106	Antifungal	D01AC15, J02AC01	1990	Equally effective in children and adults	PMID:23325436	DT	ERG11	CHEMBL1780
Fluoxetine	CHEMBL41	Antidepressant	N06AB03	1987	Equally effective in children and adults, but prone for side effects in the young	PMID:22427983	DT	SLC6A4	CHEMBL228
Flurazepam	CHEMBL968	For the treatment of anxiety	N05CD01	1970	Prone for side effects in the elderly	PMID:18218287	DT	GABRB2, GABRA3, GABRA2, GABRA1, GABRA6, GABRB1, GABRQ, GABRP, GABRG3, GABRG2, GABRG1, GABRE, GABRD, GABRB3, GABRA5, GABRA4	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Fluticasone Propionate	CHEMBL1473	For the treatment of asthma	D07AC17, R01AD08, R01AD58, R03BA05	1990	Equally effective in children and adults, but impairs growth in children	PMID:22664162	DT	NR3C1, PGR	CHEMBL2034, CHEMBL208

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Fluvoxamine	CHEMBL814	Antidepressant	N06AB08	1994	Equally effective in children and adults when the dose is adjusted	PMID:21086850	DT	SLC6A4	CHEMBL228
Furosemide	CHEMBL35	Diuretic	C03CA01	1966	Prone for side effects in the elderly, not very effective to lower heart pressure	PMID:12208585	DT	SLC12A1	CHEMBL1874
Gabapentin	CHEMBL940	Antiepileptic	N03AX12	1993	Equally effective in children and adults when the dose is adjusted; Prone to side effects in children	PMID:21086850	ME	CYP2E1	CHEMBL5281
							TP	SLC22A8, SLC22A6, ABCC4	CHEMBL1641348, CHEMBL1641347, CHEMBL1743128
							DT	CACNA1C, CACNA1D, CACNA2D4, CACNA2D3, CACNA2D2, CACNA2D1, CACNG8, CACNG7, CACNG6, CACNG5, CACNG4, CACNG3, CACNG2, CACNG1, CACNA1I, CACNA1H, CACNA1G, CACNA1E, CACNA1A, CACNA1B, CACNB4, CACNB3, CACNB2, CACNA1F, CACNA1S, CACNB1	CHEMBL2363032
Glimepiride	CHEMBL1481	For the treatment of type II diabetes	A10BB12	1995	Prone for side effects in the elderly	PMC:3448454	DT	ABCC8, KCNJ11	CHEMBL2096972, CHEMBL2096972
Glyburide	CHEMBL472	For the treatment of type II diabetes	A10BB01	1984	Prone for side effects in the elderly	PMID:22376048	DT	KCNJ11, ABCC8	CHEMBL2096972, CHEMBL2096972

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Glycopyrrolate	CHEMBL1201335	For the treatmenf of chronic severe drooling	A03AB02, R03BB06	1961	Equally effective in children and adults	PMID:22445330	DT	CHRM1, CHRM3	CHEMBL216, CHEMBL245
Griseofulvin	CHEMBL562	Antifungal	D01AA08, D01BA01	1962	Equally effective in children and adults	PMID:23278514	DT	TERG_07904, TERG_02189, TERG_07773	CHEMBL2364703, CHEMBL2364703, CHEMBL2364703
							ME	CYP2C8, CYP1A2, CYP2C9, CYP3A4	CHEMBL3721, CHEMBL3356, CHEMBL3397, CHEMBL340
Hydrochloro-thiazide	CHEMBL435	For the treatment of hypertension	C03AA03, C03AX01	1959	Not effective in children	PMID:24488616	DT	SLC12A3	CHEMBL1876
							TP	SLC12A1	CHEMBL1874
Hydroxyzine	CHEMBL896	Antihistaminic	N05BB01, N05BB51	1956	Prone for side effects in the elderly	PMID:22376048	DT	HRH1	CHEMBL231
							ME	CYP2D6	CHEMBL289
Ibuprofen	CHEMBL521	Antipyretic and analgesic and anti-inflamatory	C01EB16, G02CC01, M01AE01, M01AE51, M02AA13	1974	Equally effective in children and adults	PMID:10638387	DT	PTGS2, PTGS1	CHEMBL2094253, CHEMBL2094253
Imipramine	CHEMBL11	Antidepressant	N06AA02	1959	Not effective in children, and prone for side effects	PMID:22427983	DT	SLC6A4, SLC6A2	CHEMBL228, CHEMBL222
							ME	UGT, NAT, GST, CYP3A7, CYP1A2, CYP2E1, CYP2D6, CYP2C19, CYP2C18, CYP2B6, CYP3A4	-, -, -, -, CHEMBL3356, CHEMBL5281, CHEMBL289, CHEMBL3622, CHEMBL2408, CHEMBL4729, CHEMBL340
Ipratropium	CHEMBL1615433	COPD	R01AX03, R03BB01	1986	It can exacerbate glaucoma in the elderly	PMID:18218287	DT	CHRM3	CHEMBL245
Isoproterenol	CHEMBL434	For the treatment of hypertension	R03CB51	1956	Not effective in the elderly	PMID:17996666	DT	ADRB1, ADRB2, ADRB3	CHEMBL2094118, CHEMBL2094118, CHEMBL2094118
Isotretinoin	CHEMBL547	For the treatment of acne	D10AD04, D10AD54, D10BA01	1982	Higher incidence of depression in children/teenagers	PMID:23680205	DT	RARB	CHEMBL2008

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Ketamine	CHEMBL742	Anaesthetic	N01AX03	1970	Equally effective in children and adults	PMID:22303887	DT	GRIN3A, GRIN2D, GRIN2C, GRIN3B, GRIN1, GRIN2A, GRIN2B	CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124, CHEMBL2094124
							ME	CYP2C9, CYP2B6, CYP3A4	CHEMBL3397, CHEMBL4729, CHEMBL340
Ketorolac	CHEMBL469	Anti-inflamatory	M01AB15, S01BC05	1989	Prone for side effects in the elderly	PMID:22376048	DT	PTGS2, PTGS1	CHEMBL2094253, CHEMBL2094253
Labetalol	CHEMBL429	For the treatment of hypertension	C07AG01	1984	Not effective in the elderly	PMID:17521291	DT	ADRA1A, ADRA1B, ADRA1D, ADRA2A, ADRB3, ADRA2C, ADRB1, ADRB2, ADRA2B	CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074, CHEMBL2331074
Lamotrigine	CHEMBL741	Antiepileptic	N03AX09	1994	Increased incidence of rash in children	PMID:12060006	DT	SCN10A, SCN1A, SCN9A, SCN8A, SCN7A, SCN5A, SCN4A, SCN3A, SCN11A, SCN2A	CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043
							ME	UGT1A4	CHEMBL3619
Lanthanum Carbonate	CHEMBL2096647	For the control of serum phosphorus in patients with chronic kidney disease (CKD)	V03AE03	2004	Prone for side effects in children	PMID:23381010			
Lapatinib	CHEMBL554	Anticancer	L01XE07	2007	Equally effective in children and adults	PMID:22490765	DT	ERBB2, EGFR	CHEMBL1824, CHEMBL203
Lenalidomide	CHEMBL848	Anticancer	L04AX04	2005	Equally effective in children and adults	PMID:22490765			
Levetiracetam	CHEMBL1286	Antiepileptic	N03AX14	1999	Not effective in children, and prone for side effects	PMID:23159714	DT	SV2A, CACNA1B	CHEMBL1998, CHEMBL4478

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Lisdexamfetamine	CHEMBL1201222	ADHD	N06BA12	2007	Equally effective in children and adults	PMID:21566420	DT	SLC6A3, SLC18A2, SLC6A2	CHEMBL238, CHEMBL1893, CHEMBL222
Loperamide	CHEMBL841	For the treatment of diarrhea	A07DA03, A07DA53	1976	Prone for side effects in the elderly	PMID:18218287	DT	OPRM1	CHEMBL233
Lorazepam	CHEMBL580	Antiepileptic	N05BA06	1977	Not effective in children	PMID:22661762	DT	GABRA3, GABRQ, GABRG3, GABRG2, GABRG1, GABRE, GABRD, GABRA2, GABRA1, GABRB3, GABRB2, GABRB1, GABRA6, GABRA5, GABRA4, GABRP	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Meperidine	CHEMBL607	Analgesic	N02AB02, N02AB52, N02AB72	1942	Prone for side effects in the elderly and not effective with the current dosing scheme	PMID:22376048	DT	OPRM1	CHEMBL233
							ME	CYP3A5, CYP1A2, CYP2B6, CYP2C19, CYP3A4	CHEMBL3019, CHEMBL3356, CHEMBL4729, CHEMBL3622, CHEMBL340

Continued on next page

Table 2 – *Continued from previous page*

Drug	Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
									NDUFA9,	CHEMBL2363065,
									NDUFB1,	CHEMBL2363065,
									NDUFB10,	CHEMBL2363065,
									NDUFB11,	CHEMBL2363065,
									NDUFB2,	CHEMBL2363065,
									NDUFB3,	CHEMBL2363065,
									NDUFB4,	CHEMBL2363065,
									NDUFB5,	CHEMBL2363065,
									NDUFB6,	CHEMBL2363065,
									NDUFB7,	CHEMBL2363065,
									NDUFB8,	CHEMBL2363065,
									Q9Y6M9,	CHEMBL2363065,
									NDUFC1,	CHEMBL2363065,
Metformin		CHEMBL1431	For the treatment of type II diabetes and non-alcoholic fatty liver disease (NAFLD)	A10BA02	1995	Not effective in children with NAFLD	PMID:23075296	DT	NDUFC2,	CHEMBL2363065,
									NDUFV1,	CHEMBL2363065,
									NDUFV3,	CHEMBL2363065,
									NDUFS2,	CHEMBL2363065,
									NDUFS3,	CHEMBL2363065,
									NDUFS4,	CHEMBL2363065,
									NDUFS6,	CHEMBL2363065,
									NDUFS7,	CHEMBL2363065,
									NDUFS8,	CHEMBL2363065,
									MT-ND2,	CHEMBL2363065,
									MT-ND3,	CHEMBL2363065,
									MT-ND4,	CHEMBL2363065,
									MT-ND4L,	CHEMBL2363065,
									MT-ND6,	CHEMBL2363065,

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Metformin (continuation)	CHEMBL1431	For the treatment of type II diabetes and non-alcoholic fatty liver disease (NAFLD)	A10BA02	1995	Not effective in children with NAFLD	PMID:23075296	DT	NDUFAB1, NDUFAF1, NDUFAF2, NDUFAF3, NDUFAF4, NDUFA1, NDUFA10, NDUFA11, NDUFA12, NDUFA13, NDUFA2, NDUFA3, NDUFA4, NDUFA4L2, NDUFA5, NDUFA6, NDUFA7, NDUFA8, NDUFV2, NDUFS5, MT-ND1, MT-ND5	CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065, CHEMBL2363065
							ME	CYP3A4	CHEMBL340
							TP	SLC22A1, SLC22A2, ABCB1, ABCG2	CHEMBL5685, CHEMBL1743122, CHEMBL4302, CHEMBL5393
Methyl-phenidate	CHEMBL796	ADHD	N06BA04	1955	Prone for side effects in children	ISBN:9780470018033	DT	SLC6A2, SLC6A3	CHEMBL222, CHEMBL238
							ME	CYP2D6	CHEMBL289
Methyl-prednisolone	CHEMBL650	Anti-inflamatory	D07AA01, D10AA02, H02AB04, H02BX01	1957	Greater cortisol suppression in the elderly	PMID:17996666	DT	NR3C1	CHEMBL2034
							ME	CYP3A4	CHEMBL340
Metoclopramide	CHEMBL86	Antiemetic	A03FA01	1979	Not effective in children; Prone for side effects in the elderly with Parkinsonism	PMID:22815462	DT	DRD2, HTR3A, HTR4	CHEMBL217, CHEMBL1899, CHEMBL1875

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
							ME	CYP2D6, CYP11B1, CYP17A1, NAT, CYP11B2	CHEMBL289, CHEMBL1908, CHEMBL3522, -, CHEMBL2722
Metocurine*	CHEMBL1259	Adjunt to general anaesthesia (facilitate endotracheal intubation)	Unclassified	1982	Equally effective in adults and old people	PMID:17996666	DT	CHRNG, CHRNA1, CHRNB1, CHRND, CHRNE	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Midazolam	CHEMBL655	For the treatment of anxiety	N05CD08	1985	Increased risk of life-threatening adverse events in children with congenital heart disease or pulmonary hypertension; Prone for side effects in the elderly	PMID:21086850	DT	GABRG3, GABRA2, GABRE, GABRD, GABRB3, GABRA6, GABRA5, GABRA4, GABRA3, GABRB2, GABRP, GABRB1, GABRG2, GABRG1, GABRA1, GABRQ	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Mirabegron	CHEMBL2095212	Overactive bladder	G04BD12	2012	Equally effective in adults and old people	PMID:24113752	ME	CYP3A4	CHEMBL340
							DT	ADRB3	CHEMBL246
Mirtazapine	CHEMBL654	Antidepressant	N06AX11	1996	Not effective in children, and prone for side effects	PMID:22064376	DT	HTR2A, HTR2C, ADRA2C, ADRA2B, ADRA2A	CHEMBL224, CHEMBL225, CHEMBL2095158, CHEMBL2095158, CHEMBL2095158
Montelukast	CHEMBL787	For the treatment of asthma	R03DC03	1998	Equally effective in children and adults	PMID:22734451	DT	CYSLTR1	CHEMBL1798
Nefazodone	CHEMBL623	Antidepressant	N06AX06	1994	Not effective in children, and prone for side effects	PMID:22064376	DT	SLC6A2, HTR2C, HTR2A, SLC6A4	CHEMBL222, CHEMBL225, CHEMBL224, CHEMBL228
Nifedipine	CHEMBL193	Antiarrhythmic	C08CA05, C08CA55	1981	Prone for side effects in the elderly	PMID:22376048	DT	CACNA1S, CACNA1C, CACNA1F, CACNA1D	CHEMBL2095229, CHEMBL2095229, CHEMBL2095229, CHEMBL2095229

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Nimesulide	CHEMBL56367	Antipyretic and analgesic	M01AX17, M02AA26	1983	Equally effective in children and adults	PMID:10638387	DT	PTGS2	CHEMBL2094253
Olanzapine	CHEMBL715	For the treatment of schizophrenia/autism	N05AH03	1996	Equally effective in children and adults	PMID:16239162	DT	HTR2C, DRD4, DRD2, HTR2A, DRD3	CHEMBL225, CHEMBL2331075, CHEMBL2331075, CHEMBL224, CHEMBL2331075
Pamidronic Acid	CHEMBL834	For the treatment of hypercalcemia (loss of bone mass)	M05BA03	1991	Equally effective in children and adults	PMID:22330186	DT	FDPS	CHEMBL1782
Pancuronium	CHEMBL185073	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC01	1972	Equally effective in adults and old people	PMID:17996666	DT	CHRNA1, CHRNB1, CHRND, CHRNE, CHRNG	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Paricalcitol	CHEMBL1200622	For the treatment of hyperparathyroidism associated with Chronic Kidney Disease (CKD)	H05BX02	1998	Equally effective in children and adults	PMID:23381010	DT	VDR	CHEMBL1977
Paroxetine	CHEMBL490	Antidepressant	N06AB05	1992	Not effective in children, and prone for side effects	PMID:22427983	DT	SLC6A4	CHEMBL228
Perindopril	CHEMBL1581	For the treatment of hypertension	C09AA04	1993	Equally effective in adults and old people	PMID:17996666	DT	ACE	CHEMBL1808
Phenobarbital	CHEMBL40	Antiepileptic	N03AA02	1912	Equally effective in children and adults when the dose is adjusted	PMID:12968992	DT	GABRD, GABRA2, GABRP, GABRG3, GABRG2, GABRG1, GABRE, GABRB3, GABRB2, GABRB1, GABRA6, GABRA5, GABRA4, GABRA3, GABRQ, GABRA1	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
							ME	CYP2C19	CHEMBL3622

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Pimecrolimus	CHEMBL1200686	For the treatment of eczema	D11AH02	2001	Prone for side effects in children under 2 years old	PMID:21086850	DT	FKBP1A	CHEMBL1902
Pipecuronium*	CHEMBL1201206	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC06	1991	Equally effective in adults and old people	PMID:17996666	DT	CHRNG, CHRND, CHRNE, CHRNB1, CHRNA1	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Prazosin	CHEMBL2	For the treatment of hypertension	C02CA01	1976	Prone for side effects in the elderly	PMID:17996666	DT	ADRA1D, ADRA1B, ADRA1A	CHEMBL2094251, CHEMBL2094251, CHEMBL2094251
Prochlorperazine	CHEMBL728	For the treatment of migranes	N05AB04	1956	Not as effective in children as in adults; Prone for side effects in the elderly with Parkinsonism	PMID:18218287	DT	DRD2	CHEMBL217
Promethazine	CHEMBL643	Antihistaminic	D04AA10, R06AD02, R06AD52	1951	Prone for side effects in the elderly	PMID:22376048	ME DT	CYP3A4, CYP2D6 HRH1	CHEMBL340, CHEMBL289 CHEMBL231
Propofol	CHEMBL526	Anaesthetic	N01AX10	1989	Increased risk of death in children; safe in the elderly when the dose is reduced	PMID:21086850	ME DT	CYP2B6, CYP2D6, CYP2C9 GABRD, GABRA6, GABRA5, GABRA3, GABRA2, GABRA1, GABRQ, GABRP, GABRG2, GABRE, GABRB2, GABRG3, GABRG1, GABRB3, GABRA4, GABRB1	CHEMBL4729, CHEMBL289, CHEMBL3397 CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Propranolol	CHEMBL27	For the treatment of hypertension	C07AA05	1967	Not effective in the elderly	PMID:7955191	ME DT	UGT1A9 ADRB1, ADRB2	CHEMBL1743319 CHEMBL213, CHEMBL210

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
							ME	CYP1A1, CYP1A2, CYP2D6, CYP3A4, CYP3A5, CYP3A7, CYP2C19	CHEMBL2231, CHEMBL3356, CHEMBL289, CHEMBL340, CHEMBL3019, -, CHEMBL3622
							TP	SLC22A2	CHEMBL1743122
Remifentanil	CHEMBL1005	Analgesic	N01AH06	1996	Equally effective in old people and adults when the dose is adjusted	PMID:17996666	DT	OPRM1	CHEMBL233
Ribavirin	CHEMBL1643	For the treatment of hepatitis C	J05AB04	1985	Increased risk of suicidal ideation in children	PMID:21086850	DT	IMPDH1, , NS5B	CHEMBL1822, CHEMBL2363073, CHEMBL5375
Risperidone	CHEMBL85	For the treatment of schizophrenia/autism	N05AX08	1993	Equally effective in children and adults	PMID:24350813	DT	DRD3, DRD4, HTR2C, HTR2A, DRD2	CHEMBL2331075, CHEMBL2331075, CHEMBL225, CHEMBL224, CHEMBL2331075
Rizatriptan	CHEMBL905	For the treatment of migranes	N02CC04	1998	Not effective in children and adolescents orally	PMID:16061583	DT	HTR1B, HTR1D	CHEMBL1898, CHEMBL1983
Rocuronium	CHEMBL1201244	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC09	1994	Equally effective in adults and old people	PMID:17996666	DT	CHRNG, CHRNE, CHRND, CHRNB1, CHRNA1	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Rosiglitazone	CHEMBL121	For the treatment of type II diabetes	A10BG02	1999	Prone for side effects in the elderly	PMID:22530704	DT	PPARG	CHEMBL235
Rufinamide	CHEMBL1201754	Antiepileptic	N03AF03	2008	Equally effective in children and adults	PMID:23083943	DT	SCN8A, SCN9A, SCN10A, SCN1A, SCN7A, SCN5A, SCN4A, SCN3A, SCN2A, SCN11A	CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043, CHEMBL2331043
Sertraline	CHEMBL809	Antidepressant	N06AB06	1991	Not effective in children, and prone for side effects	PMID:22427983	DT	SLC6A4	CHEMBL228

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Sevelamer Hydrochloride	CHEMBL1201492	For the control of serum phosphorus in patients with chronic kidney disease (CKD)	V03AE02	1998	Equally effective in children and adults	PMID:23381010	DT		CHEMBL2364184
Sevoflurane	CHEMBL1200694	Anaesthetic	N01AB08	1995	Increased risk of seizures in children	PMID:21086850	DT	KCNK9, GABRA1, GABRQ, GABRA3, GABRA4, GABRA5, GABRA6, GABRB1, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRA2, GLRA1, GLRB, KCNK10, KCNK3, KCNK18, KCNK2	CHEMBL2321614, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2363052, CHEMBL2363052, CHEMBL2331041, CHEMBL2321613, CHEMBL2331042, CHEMBL2321615
Sirolimus	CHEMBL413	Prophylaxis of organ rejection	L04AA10	1999	Equally effective in children and adults	PMID:22490765	DT	FKBP1A	CHEMBL1902
Somatropin	CHEMBL1201621	For hormone replacement therapy	H01AC01	1976	Prone for side effects in the elderly	PMID:22376048	DT	GHR	CHEMBL1976
Sotalol	CHEMBL471	Antiarrhythmic	C07AA07, C07AA57	1992	Equally effective in children and adults	PMID:16198851	DT	KCNH2, ADRB2, ADRB1	CHEMBL240, CHEMBL210, CHEMBL213
Sumatriptan	CHEMBL128	For the treatment of migranes	N02CC01	1992	Not effective in children and adolescents orally, but effective when given in nasal-spray form	PMID:24163282	DT	HTR1B, HTR1D	CHEMBL1898, CHEMBL1983

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Sunitinib	CHEMBL535	Anticancer	L01XE04	2006	Different side effects in children	PMID:22490765	DT	CSF1R, FLT1, FLT3, RET, FLT4, PDGFRA, KIT, KDR, PDGFRB	CHEMBL1844, CHEMBL2095227, CHEMBL1974, CHEMBL2041, CHEMBL2095227, CHEMBL2095189, CHEMBL1936, CHEMBL2095227, CHEMBL2095189
Theophylline	CHEMBL190	For the treatment of apnea	R03DA02, R03DA04, R03DA54, R03DA74	1976	Increased risk of methylxanthines-related side effects in neonates; Prone for side effects in the elderly	PMID:448554	DT	PDE4D, ADORA2A, ADORA2B, ADORA3, PDE4A, PDE3A, PDE3B, PDE4B, ADORA1, PDE4C	CHEMBL2093863, CHEMBL2111329, CHEMBL2111329, CHEMBL2111329, CHEMBL2093863, CHEMBL2094125, CHEMBL2094125, CHEMBL2093863, CHEMBL2111329, CHEMBL2093863
							ME	CYP3A4, CYP2C8, CYP1B1, CYP1A2, CYP2D6, CYP1A1, CYP2C9, CYP2E1, NAT	CHEMBL340, CHEMBL3721, CHEMBL4878, CHEMBL3356, CHEMBL289, CHEMBL2231, CHEMBL3397, CHEMBL5281, -
							TP	SLC22A7	-

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Thiopental*	CHEMBL441	Anaesthetic	N01AF03, N05CA19	1982	Safe in the elderly when the dose is reduced	PMID:17996666	DT	GABRA4, GABRA3, GABRA6, GABRB1, GABRP, GABRQ, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRA1, GABRA2, GABRA5	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
							ME	CYP2E1, CYP3A4	CHEMBL5281, CHEMBL340
Tolterodine	CHEMBL1382	For the treatment of urinary incontinence	G04BD07	1998	Not effective in children	PMID:18310202	DT	CHRM2, CHRM3	CHEMBL211, CHEMBL245
							ME	CYP2D6	CHEMBL289
Trandolapril	CHEMBL1519	For the treatment of hypertension	C09AA10	1996	Equally effective in adults and old people	PMID:17996666	DT	ACE	CHEMBL1808
Trazodone	CHEMBL621	Antidepressant	N06AX05	1981	Equally effective in adults and old people	PMID:23192413	DT	HTR2C, HTR2A, SLC6A4	CHEMBL225, CHEMBL224, CHEMBL228
Triamterene	CHEMBL585	Diuretic	C03DB02	1964	Prone for side effects in the elderly, not very effective to lower heart pressure	PMID:12208585	DT	, SCNN1A	CHEMBL2107836, CHEMBL2107836
							ME	CYP1A2	CHEMBL3356

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Triazolam	CHEMBL646	For the short-term treatment of insomnia	N05CD05	1982	Prone for side effects in the elderly	PMID:17521291	DT	GABRB1, GABRA6, GABRA5, GABRA4, GABRA3, GABRA2, GABRA1, GABRB2, GABRB3, GABRD, GABRE, GABRG3, GABRG2, GABRQ, GABRP, GABRG1	CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872, CHEMBL2093872
Tubocurarine*	CHEMBL339427	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AA02	1945	Equally effective in adults and old people	PMID:17996666	DT	CHRNA1, CHRNB1, CHRNE, CHRNG, CHRND	CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Ursodiol	CHEMBL1551	For the treatment of primary biliary cirrhosis	A05AA02	1987	Not effective in children	PMID:10839869	TP DT ME TP	SLC22A1 NR1H4 CYP2E1, CYP7A1 SLCO1B1, SLCO6A1, ABCC2, ABCB1	CHEMBL5685 CHEMBL2047 CHEMBL5281, - CHEMBL1697668, -, CHEMBL5748, CHEMBL4302
Valproic Acid	CHEMBL109	Antiepileptic	N03AG01	1978	Increased incidence of hepatoxicity in children	PMID:13679531	DT ME	ALDH5A1 CYP3A4, UGT2B7, GST, CYP1A2, CYP2C19, CYP2C18, CYP2C9, CYP2C8, CYP2A6, CYP4A11	CHEMBL1911 CHEMBL340, CHEMBL4370, -, CHEMBL3356, CHEMBL3622, CHEMBL2408, CHEMBL3397, CHEMBL3721, CHEMBL5282, -

Continued on next page

Table 2 – *Continued from previous page*

Drug Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
Vecuronium	CHEMBL1201219	Adjunt to general anaesthesia (facilitate endotracheal intubation)	M03AC03	1984	Equally effective in adults and old people	PMID:17996666	TP DT	ABCB1 CHRNE, CHRNG, CHRND, CHRNB1, CHRNA1	CHEMBL4302 CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997, CHEMBL2362997
Venlafaxine	CHEMBL637	Antidepressant	N06AX16	1993	Not effective in children, and prone for side effects in children and in old people	PMID:22251022	DT	SLC6A2, SLC6A4	CHEMBL222, CHEMBL228
Verapamil	CHEMBL6966	For the treatment of angina and arrhythmias	C08DA01, C08DA51	1981	Prone for side effects in children and old people; Not effective in children	PMID:24113539	DT	CACNA1S, CACNA1F, CACNA1D, CACNA1C	CHEMBL2095229, CHEMBL2095229, CHEMBL2095229, CHEMBL2095229
							ME	CYP3A4, CYP1A2, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP3A5, CYP3A7, GST	CHEMBL340, CHEMBL3356, CHEMBL4729, CHEMBL3721, CHEMBL3397, CHEMBL3622, CHEMBL289, CHEMBL3019, -, -
							TP	ABCB4, ABCB11, SLC22A5, ABCB1, SLC2A4, ABCC1, SLCO6A1, ABCC2	CHEMBL1743129, CHEMBL6020, CHEMBL2073693, CHEMBL4302, CHEMBL5874, CHEMBL3004, -, CHEMBL5748
Vigabatrin	CHEMBL89598	Antiepileptic	N03AG04	2009	Equally effective in children and adults when the dose is adjusted	PMID:23740534	DT	ABAT	CHEMBL2044
Warfarin	CHEMBL1464	Anticoagulant	B01AA03	1954	Reduces bone density in children; Prone for side effects in the elderly	PMID:22315269	DT	VKORC1	CHEMBL1930

Continued on next page

Table 2 – *Continued from previous page*

Drug	Name	ChEMBL ID	Indication	ATC code	Approval year	DDR	Reference	Target type	Gene symbol	Target ChEMBL ID
								ME	CYP2C9, CYP1A2, CYP2C8, CYP2C18, CYP2C19, NAT, GST, UGT, SUL, CYP3A4	CHEMBL3397, CHEMBL3356, CHEMBL3721, CHEMBL2408, CHEMBL3622, -, -, -, -, CHEMBL340

Appendix B

Tables with the pharmacological relevant genes.

Table 3: List of human drug targets for all FDA approved drugs. Human orthologues of pharmacological relevant genes of *Macaca mulata*, *Rat norvegicus*, and *Mus musculus* were extracted from ENSEMBL Compara version 72 [144; 145]

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ABAT	P80404	4-aminobutyrate aminotransferase	ENSG00000183044	-	ENSRNOG00000002636	ENSMUSG00000057880
ABCA1	O95477	ATP-binding cassette, sub-family A (ABC1), member 1	ENSG00000165029	ENSMMUG00000020608	ENSRNOG00000018126	ENSMUSG00000015243
ABCC8	Q09428	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	ENSG00000006071	ENSMMUG00000023581	ENSRNOG00000021130	ENSMUSG00000040136
ABCC9	O60706	ATP-binding cassette, sub-family C (CFTR/MRP), member 9	ENSG00000069431	ENSMMUG00000020806	ENSRNOG00000036960	ENSMUSG00000030249
ABL1	P00519	c-abl oncogene 1, non-receptor tyrosine kinase	ENSG00000097007	ENSMMUG00000012499	ENSRNOG00000009371 ENSRNOG00000047356	ENSMUSG00000026842
ACE	P12821	angiotensin I converting enzyme	ENSG00000159640	ENSMMUG00000012608	ENSRNOG00000007467	ENSMUSG00000020681
ACHE	P22303	acetylcholinesterase	ENSG00000087085	ENSMMUG00000021257	ENSRNOG00000050841	ENSMUSG00000023328
ADA	P00813	adenosine deaminase	ENSG00000196839	ENSMMUG00000010603	ENSRNOG00000010265	ENSMUSG00000017697
ADH1A	P07327	alcohol dehydrogenase 1A (class I), alpha polypeptide	ENSG00000187758	-	ENSRNOG00000012464	ENSMUSG00000074207
ADH1B	P00325	alcohol dehydrogenase 1B (class I), beta polypeptide	ENSG00000196616	ENSMMUG00000018738	ENSRNOG00000012464	ENSMUSG00000074207
ADORA1	P30542	adenosine A1 receptor	ENSG00000163485	ENSMMUG00000008714	ENSRNOG00000003442	ENSMUSG00000042429
ADORA2A	P29274	adenosine A2a receptor	ENSG00000128271	ENSMMUG00000000261	ENSRNOG00000001302	ENSMUSG00000020178
ADORA2B	P29275	adenosine A2b receptor	ENSG00000170425	ENSMMUG00000020399	ENSRNOG00000002922	ENSMUSG00000018500
ADORA3	P33765	adenosine A3 receptor	ENSG00000121933	ENSMMUG00000003581	ENSRNOG00000048272 ENSRNOG00000043375	ENSMUSG00000074344
ADRA1A	P35348	adrenoceptor alpha 1A	ENSG00000120907	ENSMMUG00000023112	ENSRNOG00000009522	ENSMUSG00000045875
ADRA1B	P35368	adrenoceptor alpha 1B	ENSG00000170214	ENSMMUG00000005691	-	ENSMUSG00000050541
ADRA1D	P25100	adrenoceptor alpha 1D	ENSG00000171873	ENSMMUG00000020424	ENSRNOG00000021256	ENSMUSG00000027335
ADRA2A	P08913	adrenoceptor alpha 2A	ENSG00000150594	ENSMMUG00000005316	ENSRNOG00000047545	ENSMUSG00000033717
ADRA2B	P18089	adrenoceptor alpha 2B	ENSG00000222040	ENSMMUG00000011842	ENSRNOG00000013887	ENSMUSG00000058620
ADRA2C	P18825	adrenoceptor alpha 2C	ENSG00000184160	ENSMMUG00000014781	ENSRNOG00000009299	ENSMUSG00000045318
ADRB1	P08588	adrenoceptor beta 1	ENSG00000043591	-	ENSRNOG00000017002	ENSMUSG00000035283
ADRB2	P07550	adrenoceptor beta 2, surface	ENSG00000169252	ENSMMUG00000002214	ENSRNOG00000019217	ENSMUSG00000045730
ADRB3	P13945	adrenoceptor beta 3	ENSG00000188778	ENSMMUG00000005876	ENSRNOG00000012674	ENSMUSG00000031489
AGTR1	P30556	angiotensin II receptor, type 1	ENSG00000144891	ENSMMUG00000020680	ENSRNOG00000018346 ENSRNOG00000010640	ENSMUSG00000049115 ENSMUSG00000054988
ALDH2	P05091	aldehyde dehydrogenase 2 family (mitochondrial)	ENSG00000111275	ENSMMUG00000008336	ENSRNOG00000001344	ENSMUSG00000029455
ALDH5A1	P51649	aldehyde dehydrogenase 5 family, member A1	ENSG00000112294	ENSMMUG00000019368	ENSRNOG00000023538	ENSMUSG00000035936

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ALK	Q9UM73	anaplastic lymphoma receptor tyrosine kinase	ENSG00000171094	ENSMMUG00000010265	ENSRNOG00000008683	ENSMUSG00000055471
ALOX5	P09917	arachidonate 5-lipoxygenase	ENSG00000262552 ENSG00000012779	ENSMMUG00000013434	ENSRNOG00000012972	ENSMUSG00000025701
AMY2A	P04746	amylase, alpha 2A (pancreatic)	ENSG00000243480	ENSMMUG00000031908 ENSMMUG00000031904	ENSRNOG00000016168 ENSRNOG00000030590 ENSRNOG00000031336 ENSRNOG00000033529	ENSMUSG00000074268 ENSMUSG00000093931 ENSMUSG00000096770 ENSMUSG00000096569 ENSMUSG00000070360 ENSMUSG00000074264
AR	P10275	androgen receptor	ENSG00000169083	ENSMMUG00000011437	ENSRNOG00000005639	ENSMUSG00000046532
ATP1A1	P05023	ATPase, Na+/K+ transporting, alpha 1 polypeptide	ENSG00000163399	ENSMMUG00000008151	ENSRNOG00000030019	ENSMUSG00000033161
ATP1A2	P50993	ATPase, Na+/K+ transporting, alpha 2 polypeptide	ENSG00000018625	ENSMMUG00000017833	ENSRNOG00000007290	ENSMUSG00000007097
ATP1A3	P13637	ATPase, Na+/K+ transporting, alpha 3 polypeptide	ENSG00000105409	ENSMMUG00000017427	ENSRNOG00000020263	ENSMUSG00000040907
ATP1A4	Q13733	ATPase, Na+/K+ transporting, alpha 4 polypeptide	ENSG00000132681	ENSMMUG00000017836	ENSRNOG00000032378	ENSMUSG00000007107
ATP4A	P20648	ATPase, H+/K+ exchanging, alpha polypeptide	ENSG00000105675	ENSMMUG00000000845	ENSRNOG00000020985	ENSMUSG00000005553
AVPR1A	P37288	arginine vasopressin receptor 1A	ENSG00000166148	ENSMMUG00000000549	ENSRNOG00000004400	ENSMUSG00000020123
AVPR1B	P47901	arginine vasopressin receptor 1B	ENSG00000265027 ENSG00000198049	ENSMMUG00000002255	ENSRNOG00000048522 ENSRNOG00000049261	ENSMUSG00000026432
AVPR2	P30518	arginine vasopressin receptor 2	ENSG00000126895 ENSG00000269623	ENSMMUG00000005026	ENSRNOG00000037270	ENSMUSG00000031390
BCHE	P06276	butyrylcholinesterase	ENSG00000114200	ENSMMUG00000003742	ENSRNOG00000009826	ENSMUSG00000027792
BCR	P11274	breakpoint cluster region	ENSG00000186716	ENSMMUG00000023622 ENSMMUG00000020673	ENSRNOG00000001304	ENSMUSG00000009681
BDKRB2	P30411	bradykinin receptor B2	ENSG00000168398	ENSMMUG00000006611	ENSRNOG00000047300	ENSMUSG00000021070
BLK	P51451	B lymphoid tyrosine kinase	ENSG00000136573	ENSMMUG00000019206	ENSRNOG00000010798	ENSMUSG00000014453
BRAF	P15056	v-raf murine sarcoma viral oncogene homolog B	ENSG00000157764	-	ENSRNOG00000010957	ENSMUSG00000002413
C5	P01031	complement component 5	ENSG00000106804	ENSMMUG00000020124	ENSRNOG00000018899	ENSMUSG00000026874
CA1	P00915	carbonic anhydrase I	ENSG00000133742	ENSMMUG00000014383	ENSRNOG00000010698	ENSMUSG00000027556
CA12	O43570	carbonic anhydrase XII	ENSG00000074410	ENSMMUG00000021338	ENSRNOG00000017766	ENSMUSG00000032373
CA13	Q8N1Q1	carbonic anhydrase XIII	ENSG00000185015	ENSMMUG00000011143	ENSRNOG00000021323	ENSMUSG00000027555
CA14	Q9ULX7	carbonic anhydrase XIV	ENSG00000118298 ENSG00000264654	ENSMMUG00000021385	ENSRNOG00000023162	ENSMUSG00000038526
CA2	P00918	carbonic anhydrase II	ENSG00000104267	ENSMMUG00000014891	ENSRNOG00000009629	ENSMUSG00000027562
CA3	P07451	carbonic anhydrase III, muscle specific	ENSG00000164879	ENSMMUG00000014890	ENSRNOG00000010079	ENSMUSG00000027559
CA4	P22748	carbonic anhydrase IV	ENSG00000167434	ENSMMUG00000007917	ENSRNOG00000002916	ENSMUSG00000000805

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CA5A	P35218	carbonic anhydrase VA, mitochondrial	ENSG00000174990	ENSMMUG00000021404	ENSRNOG00000019098	ENSMUSG00000025317
CA5B	Q9Y2D0	carbonic anhydrase VB, mitochondrial	ENSG00000169239	ENSMMUG00000014052	ENSRNOG00000029330	ENSMUSG00000031373
CA6	P23280	carbonic anhydrase VI	ENSG00000131686	ENSMMUG00000019687	ENSRNOG00000021355	ENSMUSG00000028972
CA7	P43166	carbonic anhydrase VII	ENSG00000168748	ENSMMUG00000006749	ENSRNOG00000012371	ENSMUSG00000031883
CA9	Q16790	carbonic anhydrase IX	ENSG00000107159	ENSMMUG00000016320	ENSRNOG00000017073	ENSMUSG00000028463
CACNA1A	O00555	calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	ENSG00000141837	ENSMMUG00000003555	ENSRNOG00000046731	ENSMUSG00000034656
CACNA1B	Q00975	calcium channel, voltage-dependent, N type, alpha 1B subunit	ENSG00000148408	ENSMMUG00000009187	ENSRNOG00000004560	ENSMUSG00000004113
CACNA1C	Q13936	calcium channel, voltage-dependent, L type, alpha 1C subunit	ENSG00000151067	ENSMMUG00000012988	ENSRNOG00000007090	ENSMUSG00000051331
CACNA1D	Q01668	calcium channel, voltage-dependent, L type, alpha 1D subunit	ENSG00000157388	ENSMMUG00000019262	ENSRNOG00000013147	ENSMUSG00000015968
CACNA1E	Q15878	calcium channel, voltage-dependent, R type, alpha 1E subunit	ENSG00000198216	ENSMMUG00000014498	ENSRNOG00000002863	ENSMUSG00000004110
CACNA1F	O60840	calcium channel, voltage-dependent, L type, alpha 1F subunit	ENSG00000102001 ENSG00000269057	ENSMMUG00000008618	ENSRNOG00000010348	ENSMUSG00000031142
CACNA1G	O43497	calcium channel, voltage-dependent, T type, alpha 1G subunit	ENSG00000006283	ENSMMUG00000017256	ENSRNOG00000002981	ENSMUSG00000020866
CACNA1H	O95180	calcium channel, voltage-dependent, T type, alpha 1H subunit	ENSG00000196557	ENSMMUG00000014213	ENSRNOG00000033893	ENSMUSG00000024112
CACNA1I	Q9P0X4	calcium channel, voltage-dependent, T type, alpha 1I subunit	ENSG00000100346	ENSMMUG00000004878	ENSRNOG00000029914	ENSMUSG00000022416
CACNA1S	Q13698	calcium channel, voltage-dependent, L type, alpha 1S subunit	ENSG00000081248	-	ENSRNOG00000046231	ENSMUSG00000026407
CACNA2D1	P54289	calcium channel, voltage-dependent, alpha 2/delta subunit 1	ENSG00000153956	ENSMMUG00000018655	ENSRNOG00000033531	ENSMUSG00000040118
CACNA2D2	Q9NY47	calcium channel, voltage-dependent, alpha 2/delta subunit 2	ENSG00000007402	ENSMMUG00000007022	ENSRNOG00000015835	ENSMUSG00000010066

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CACNA2D3	Q8IZS8	calcium channel, voltage-dependent, alpha 2/delta subunit 3	ENSG00000157445	-	ENSRNOG00000031287	ENSMUSG00000021991
CACNA2D4	Q7Z3S7	calcium channel, voltage-dependent, alpha 2/delta subunit 4	ENSG00000151062	ENSMMUG00000017265	ENSRNOG00000008031	ENSMUSG00000041460
CACNB1	Q02641	calcium channel, voltage-dependent, beta 1 subunit	ENSG00000067191	ENSMMUG00000000289	ENSRNOG00000004518	ENSMUSG00000020882
CACNB2	Q08289	calcium channel, voltage-dependent, beta 2 subunit	ENSG00000165995	ENSMMUG00000007445	ENSRNOG00000018378	ENSMUSG00000057914
CACNB3	P54284	calcium channel, voltage-dependent, beta 3 subunit	ENSG00000167535	ENSMMUG00000008430	ENSRNOG00000012489	ENSMUSG00000003352
CACNB4	O00305	calcium channel, voltage-dependent, beta 4 subunit	ENSG00000182389	ENSMMUG00000010526	ENSRNOG00000007666	ENSMUSG00000017412
CACNG1	Q06432	calcium channel, voltage-dependent, gamma subunit 1	ENSG00000108878	ENSMMUG00000002958	ENSRNOG00000003245	ENSMUSG00000020722
CACNG2	Q9Y698	calcium channel, voltage-dependent, gamma subunit 2	ENSG00000166862	ENSMMUG00000031026	ENSRNOG00000006226	ENSMUSG00000019146
CACNG3	O60359	calcium channel, voltage-dependent, gamma subunit 3	ENSG00000006116	ENSMMUG00000013560	ENSRNOG00000012362	ENSMUSG00000066189
CACNG4	Q9UBN1	calcium channel, voltage-dependent, gamma subunit 4	ENSG00000075461	ENSMMUG00000002957	ENSRNOG00000003262	ENSMUSG00000020723
CACNG5	Q9UF02	calcium channel, voltage-dependent, gamma subunit 5	ENSG00000075429	ENSMMUG00000018072	ENSRNOG00000003288	ENSMUSG00000040373
CACNG6	Q9BXT2	calcium channel, voltage-dependent, gamma subunit 6	ENSG00000130433	ENSMMUG00000000292	ENSRNOG00000014326	ENSMUSG00000078815
CACNG7	P62955	calcium channel, voltage-dependent, gamma subunit 7	ENSG00000105605	ENSMMUG00000007975	ENSRNOG00000014416	ENSMUSG00000069806
CACNG8	Q8WXS5	calcium channel, voltage-dependent, gamma subunit 8	ENSG00000142408	-	ENSRNOG00000014348	ENSMUSG00000053395
CALCR	P30988	calcitonin receptor	ENSG00000004948	ENSMMUG00000014319	ENSRNOG00000010053	ENSMUSG00000023964
CASR	P41180	calcium-sensing receptor	ENSG00000036828	ENSMMUG00000007886	ENSRNOG00000002265	ENSMUSG00000051980

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CCKAR	P32238	cholecystokinin A receptor	ENSG00000163394	ENSMMUG00000013305	ENSRNOG00000043124	ENSMUSG00000029193
CCR5	P51681	chemokine (C-C motif) receptor 5 (gene/pseudogene)	ENSG00000160791	-	-	-
CD2	P06729	CD2 molecule	ENSG00000116824	ENSMMUG00000012272	ENSRNOG00000015821	ENSMUSG00000027863
CD33	P20138	CD33 molecule	ENSG00000105383	ENSMMUG00000009486	-	-
CD3E	P07766	CD3e molecule, epsilon (CD3-TCR complex)	ENSG00000198851	ENSMMUG00000017593	ENSRNOG00000016069	ENSMUSG00000032093
CD52	P31358	CD52 molecule	ENSG00000169442	ENSMMUG00000010240	-	-
CD80	P33681	CD80 molecule	ENSG00000121594	ENSMMUG00000016367	ENSRNOG00000001527	ENSMUSG00000075122
CD86	P42081	CD86 molecule	ENSG00000114013	ENSMMUG00000000912	ENSRNOG00000038835	ENSMUSG00000022901
CFTR	P13569	cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	ENSG00000001626	ENSMMUG00000011269	ENSRNOG00000008284	ENSMUSG00000041301
CHRM1	P11229	cholinergic receptor, muscarinic 1	ENSG00000168539	ENSMMUG00000017292	ENSRNOG00000018385	ENSMUSG00000032773
CHRM2	P08172	cholinergic receptor, muscarinic 2	ENSG00000181072	-	ENSRNOG00000046972 ENSRNOG00000011928	ENSMUSG00000045613
CHRM3	P20309	cholinergic receptor, muscarinic 3	ENSG00000133019	ENSMMUG00000021442	ENSRNOG00000049410	ENSMUSG00000046159
CHRNA1	P02708	cholinergic receptor, nicotinic, alpha 1 (muscle)	ENSG00000138435	ENSMMUG00000021796	ENSRNOG00000018286	ENSMUSG00000027107
CHRNA3	P32297	cholinergic receptor, nicotinic, alpha 3 (neuronal)	ENSG00000080644	ENSMMUG00000023736	ENSRNOG00000013829	ENSMUSG00000032303
CHRNA4	P43681	cholinergic receptor, nicotinic, alpha 4 (neuronal)	ENSG00000101204	ENSMMUG00000015916	ENSRNOG00000011202	ENSMUSG00000027577
CHRNB1	P11230	cholinergic receptor, nicotinic, beta 1 (muscle)	ENSG00000170175	ENSMMUG00000017690	ENSRNOG00000014698	ENSMUSG00000041189
CHRNB2	P17787	cholinergic receptor, nicotinic, beta 2 (neuronal)	ENSG00000160716	ENSMMUG00000016089	ENSRNOG00000020778	ENSMUSG00000027950
CHRNB4	P30926	cholinergic receptor, nicotinic, beta 4 (neuronal)	ENSG00000117971	ENSMMUG00000007541	ENSRNOG00000014427	ENSMUSG00000035200
CHRND	Q07001	cholinergic receptor, nicotinic, delta (muscle)	ENSG00000135902	ENSMMUG00000022147	ENSRNOG00000019527	ENSMUSG00000026251
CHRNE	Q04844	cholinergic receptor, nicotinic, epsilon (muscle)	ENSG00000108556	ENSMMUG00000002409	ENSRNOG00000003777	ENSMUSG00000014609
CHRNG	P07510	cholinergic receptor, nicotinic, gamma (muscle)	ENSG00000196811	ENSMMUG00000022149	ENSRNOG00000019602	ENSMUSG00000026253
CLCA1	A8K7I4	chloride channel accessory 1	ENSG00000016490	ENSMMUG00000006836	ENSRNOG00000013447	ENSMUSG00000028255
CLCA2	Q9UQC9	chloride channel accessory 2	ENSG00000137975	ENSMMUG00000001381	ENSRNOG00000013771	ENSMUSG00000036960
CLCA4	Q14CN2	chloride channel accessory 4	ENSG00000016602	ENSMMUG00000010306	ENSRNOG00000029889	ENSMUSG00000068547 ENSMUSG00000074195 ENSMUSG00000074196
CLCN2	P51788	chloride channel, voltage-sensitive 2	ENSG00000114859	ENSMMUG00000007507	ENSRNOG00000001742 ENSRNOG00000050854	ENSMUSG00000022843

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CNR1	P21554	cannabinoid receptor 1 (brain)	ENSG00000118432	ENSMMUG00000000763	ENSRNOG00000008223	ENSMUSG00000044288
COL11A1	P12107	collagen, type XI, alpha 1	ENSG00000060718	ENSMMUG00000004103	ENSRNOG00000023148	ENSMUSG00000027966
COL11A2	P13942	collagen, type XI, alpha 2	ENSG00000235708 ENSG00000204248 ENSG00000230930 ENSG00000206290 ENSG00000227801	ENSMMUG00000003850	ENSRNOG00000000463	ENSMUSG00000024330
COL15A1	P39059	collagen, type XV, alpha 1	ENSG00000204291	ENSMMUG00000011582	ENSRNOG00000039197	ENSMUSG00000028339
COL18A1	P39060	collagen, type XVIII, alpha 1	ENSG00000182871	ENSMMUG00000004472 ENSMMUG00000030502 ENSMMUG00000030501	ENSRNOG00000001229	ENSMUSG00000001435
COL1A1	P02452	collagen, type I, alpha 1	ENSG00000108821	ENSMMUG00000001467	ENSRNOG00000003897	ENSMUSG00000001506
COL1A2	P08123	collagen, type I, alpha 2	ENSG00000164692	-	ENSRNOG00000011292	ENSMUSG00000029661
COL24A1	Q17RW2	collagen, type XXIV, alpha 1	ENSG00000171502	ENSMMUG00000013205	ENSRNOG00000014143	ENSMUSG00000028197
COL27A1	Q8IZC6	collagen, type XXVII, alpha 1	ENSG00000196739	ENSMMUG00000031973	ENSRNOG00000007657	ENSMUSG00000045672
COL28A1	Q2UY09	collagen, type XXVIII, alpha 1	ENSG00000215018	-	ENSRNOG00000033618	ENSMUSG00000068794
COL2A1	P02458	collagen, type II, alpha 1	ENSG00000139219	ENSMMUG00000011444	ENSRNOG00000022282	ENSMUSG00000022483
COL3A1	P02461	collagen, type III, alpha 1	ENSG00000168542	ENSMMUG00000021286	ENSRNOG00000003357	ENSMUSG00000026043
COL4A1	P02462	collagen, type IV, alpha 1	ENSG00000187498	ENSMMUG00000022280	ENSRNOG00000016281	ENSMUSG00000031502
COL4A2	P08572	collagen, type IV, alpha 2	ENSG00000134871	-	ENSRNOG00000023972	ENSMUSG00000031503
COL4A3	Q01955	collagen, type IV, alpha 3 (Goodpasture antigen)	ENSG00000169031	ENSMMUG00000013235	ENSRNOG00000015365 ENSRNOG00000046521	ENSMUSG00000079465
COL4A4	P53420	collagen, type IV, alpha 4	ENSG00000081052	ENSMMUG00000013233 ENSMMUG00000031162	-	ENSMUSG00000067158
COL4A5	P29400	collagen, type IV, alpha 5	ENSG00000188153	ENSMMUG00000007219	ENSRNOG00000018951	ENSMUSG00000031274
COL4A6	Q14031	collagen, type IV, alpha 6	ENSG00000197565	ENSMMUG00000007205	ENSRNOG00000018856	ENSMUSG00000031273
COL5A1	P20908	collagen, type V, alpha 1	ENSG00000130635	ENSMMUG00000012147	ENSRNOG00000008749	ENSMUSG00000026837
COL5A2	P05997	collagen, type V, alpha 2	ENSG00000204262	ENSMMUG00000021290	ENSRNOG00000003736	ENSMUSG00000026042
COL5A3	P25940	collagen, type V, alpha 3	ENSG00000080573	-	ENSRNOG00000020525	ENSMUSG00000004098
COL6A1	P12109	collagen, type VI, alpha 1	ENSG00000142156	ENSMMUG00000000544	ENSRNOG00000001249	ENSMUSG00000001119
COL6A2	P12110	collagen, type VI, alpha 2	ENSG00000142173	ENSMMUG00000022317	ENSRNOG00000001254	ENSMUSG00000020241
COL6A3	P12111	collagen, type VI, alpha 3	ENSG00000163359	ENSMMUG00000014049	ENSRNOG00000019648 ENSRNOG00000050720	ENSMUSG00000048126
COL6A5	A8TX70	collagen, type VI, alpha 5	ENSG00000172752	-	ENSRNOG00000010663	ENSMUSG00000091345
COL6A6	A6NMZ7	collagen, type VI, alpha 6	ENSG00000206384	-	ENSRNOG00000023007	ENSMUSG00000043719
COMT	P21964	catechol-O-methyltransferase	ENSG00000093010	ENSMMUG00000017152	ENSRNOG00000001889	ENSMUSG00000000326
CPS1	P31327	carbamoyl-phosphate synthase 1, mitochondrial	ENSG00000021826	ENSMMUG00000010508	ENSRNOG00000013704	ENSMUSG00000025991
CSF1R	P07333	colony stimulating factor 1 receptor	ENSG00000182578	ENSMMUG00000011556	ENSRNOG00000018414	ENSMUSG00000024621
CSF2RA	P15509	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	ENSG00000198223	-	-	-

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CSF3R	Q99062	colony stimulating factor 3 receptor (granulocyte)	ENSG00000119535	ENSMMUG00000020434	ENSRNOG00000008759	ENSMUSG00000028859
CTLA4	P16410	cytotoxic T-lymphocyte-associated protein 4	ENSG00000163599	ENSMMUG00000015903	ENSRNOG00000010336	ENSMUSG00000026011
CXCR4	P61073	chemokine (C-X-C motif) receptor 4	ENSG00000121966	ENSMMUG00000011918	ENSRNOG00000003866	ENSMUSG00000045382
CYP11A1	P05108	cytochrome P450, family 11, subfamily A, polypeptide 1	ENSG00000140459	ENSMMUG00000003239	ENSRNOG00000008074	ENSMUSG00000032323
CYP11B1	P15538	cytochrome P450, family 11, subfamily B, polypeptide 1	ENSG00000160882	ENSMMUG00000013631	ENSRNOG00000046367	ENSMUSG00000022589
					ENSRNOG00000030111	ENSMUSG00000075604
					ENSRNOG00000049205	
CYP17A1	P05093	cytochrome P450, family 17, subfamily A, polypeptide 1	ENSG00000148795	ENSMMUG00000017601	ENSRNOG00000020035	ENSMUSG00000003555
CYP19A1	P11511	cytochrome P450, family 19, subfamily A, polypeptide 1	ENSG00000137869	ENSMMUG00000002553	ENSRNOG00000000196	ENSMUSG00000032274
CYP3A4	P08684	cytochrome P450, family 3, subfamily A, polypeptide 4	ENSG00000160868	-	ENSRNOG00000001379	ENSMUSG00000061292
					ENSRNOG00000032560	ENSMUSG00000054417
					ENSRNOG00000047745	ENSMUSG00000029727
					ENSRNOG00000000978	ENSMUSG00000029630
					ENSRNOG00000046643	ENSMUSG00000075551
					ENSRNOG00000049143	ENSMUSG00000075552
					ENSRNOG00000000969	ENSMUSG00000038656
						ENSMUSG00000070419
						ENSMUSG00000056035
CYP3A43	Q9HB55	cytochrome P450, family 3, subfamily A, polypeptide 43	ENSG00000021461	ENSMMUG00000022490	ENSRNOG00000001379	ENSMUSG00000061292
					ENSRNOG00000032560	ENSMUSG00000054417
					ENSRNOG00000047745	ENSMUSG00000029727
					ENSRNOG00000000978	ENSMUSG00000029630
					ENSRNOG00000046643	ENSMUSG00000075551
					ENSRNOG00000049143	ENSMUSG00000075552
					ENSRNOG00000000969	ENSMUSG00000038656
						ENSMUSG00000070419
						ENSMUSG00000056035
CYP3A5	P20815	cytochrome P450, family 3, subfamily A, polypeptide 5	ENSG00000106258	-	ENSRNOG00000001379	ENSMUSG00000061292
					ENSRNOG00000032560	ENSMUSG00000054417
					ENSRNOG00000047745	ENSMUSG00000029727
					ENSRNOG00000000978	ENSMUSG00000029630
					ENSRNOG00000046643	ENSMUSG00000075551
					ENSRNOG00000049143	ENSMUSG00000075552

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
					ENSRNOG00000000969	ENSMUSG00000038656
						ENSMUSG00000070419
						ENSMUSG00000056035
CYP3A7	P24462	cytochrome P450, family 3, subfamily A, polypeptide 7	ENSG00000160870	ENSMMUG00000022730	ENSRNOG00000001379	ENSMUSG00000061292
					ENSRNOG00000032560	ENSMUSG00000054417
					ENSRNOG00000047745	ENSMUSG00000029727
					ENSRNOG00000000978	ENSMUSG00000029630
					ENSRNOG00000046643	ENSMUSG00000075551
					ENSRNOG00000049143	ENSMUSG00000075552
					ENSRNOG00000000969	ENSMUSG00000038656
						ENSMUSG00000070419
						ENSMUSG00000056035
CYSLTR1	Q9Y271	cysteinyl leukotriene receptor 1	ENSG00000173198	ENSMMUG00000013387	ENSRNOG00000037845	ENSMUSG00000052821
DDC	P20711	dopa decarboxylase (aromatic L-amino acid decarboxylase)	ENSG00000132437	ENSMMUG00000019991	ENSRNOG00000004327	ENSMUSG00000020182
DDR2	Q16832	discoidin domain receptor tyrosine kinase 2	ENSG00000162733	ENSMMUG00000009330	ENSRNOG00000002881	ENSMUSG00000026674
DHFR	P00374	dihydrofolate reductase	ENSG00000228716	ENSMMUG00000004229 ENSMMUG00000018283	ENSRNOG00000013521	ENSMUSG00000021707
DHODH	Q02127	dihydroorotate dehydrogenase (quinone)	ENSG00000102967	ENSMMUG00000023312	ENSRNOG00000015063	ENSMUSG00000031730
DIO1	P49895	deiodinase, iodothyronine, type I	ENSG00000211452	ENSMMUG00000001415	-	ENSMUSG00000034785
DNMT1	P26358	DNA (cytosine-5-)-methyltransferase 1	ENSG00000130816	ENSMMUG00000014604	ENSRNOG00000039859	ENSMUSG00000004099
DNMT3A	Q9Y6K1	DNA (cytosine-5-)-methyltransferase 3 alpha	ENSG00000119772	ENSMMUG00000013565	ENSRNOG00000026649	ENSMUSG00000020661
DPEP1	P16444	dipeptidase 1 (renal)	ENSG00000015413	ENSMMUG00000014910	ENSRNOG00000015880	ENSMUSG00000019278
DPP4	P27487	dipeptidyl-peptidase 4	ENSG00000197635	ENSMMUG00000008387	ENSRNOG00000030763	ENSMUSG00000035000
DRD1	P21728	dopamine receptor D1	ENSG00000184845	ENSMMUG00000005147	ENSRNOG00000023688	ENSMUSG00000021478
DRD2	P14416	dopamine receptor D2	ENSG00000149295	ENSMMUG00000014334	ENSRNOG00000008428	ENSMUSG00000032259
DRD3	P35462	dopamine receptor D3	ENSG00000151577	ENSMMUG00000022936	ENSRNOG00000026772	ENSMUSG00000022705
DRD4	P21917	dopamine receptor D4	ENSG00000069696	ENSMMUG00000002658	ENSRNOG00000017927	ENSMUSG00000025496
DRD5	P21918	dopamine receptor D5	ENSG00000169676	ENSMMUG00000002783	ENSRNOG00000005338	ENSMUSG00000039358
EDNRA	P25101	endothelin receptor type A	ENSG00000151617	ENSMMUG00000003659	ENSRNOG00000012721	ENSMUSG00000031616
EDNRB	P24530	endothelin receptor type B	ENSG00000136160	ENSMMUG00000005131	ENSRNOG00000010997	ENSMUSG00000022122
EEF2	P13639	eukaryotic translation elongation factor 2	ENSG00000167658	ENSMMUG00000022767	ENSRNOG00000020266	ENSMUSG00000034994
EGFR	P00533	epidermal growth factor receptor	ENSG00000146648	ENSMMUG00000022394	ENSRNOG00000004332	ENSMUSG00000020122
EPHA1	P21709	EPH receptor A1	ENSG00000146904	ENSMMUG00000012638	ENSRNOG00000017525	ENSMUSG00000029859
EPHA10	Q5JZY3	EPH receptor A10	ENSG00000183317	ENSMMUG00000014342	ENSRNOG00000037340	-
EPHA2	P29317	EPH receptor A2	ENSG00000142627	ENSMMUG00000018245	ENSRNOG00000009222	ENSMUSG00000006445

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
EPHA3	P29320	EPH receptor A3	ENSG00000044524	ENSMMUG00000008242	ENSRNOG00000030285	ENSMUSG00000052504
EPHA4	P54764	EPH receptor A4	ENSG00000116106	ENSMMUG00000012364	ENSRNOG00000013213	ENSMUSG00000026235
EPHA5	P54756	EPH receptor A5	ENSG00000145242	ENSMMUG00000014468	ENSRNOG00000002024	ENSMUSG00000029245
EPHA6	Q9UF33	EPH receptor A6	ENSG00000080224	-	-	ENSMUSG00000055540
EPHA7	Q15375	EPH receptor A7	ENSG00000135333	ENSMMUG00000014867	ENSRNOG00000007030	ENSMUSG00000028289
EPHA8	P29322	EPH receptor A8	ENSG00000070886	ENSMMUG00000020754	ENSRNOG00000013036	ENSMUSG00000028661
EPHB1	P54762	EPH receptor B1	ENSG00000154928	ENSMMUG00000021334	ENSRNOG00000007865	ENSMUSG00000032537
EPHB2	P29323	EPH receptor B2	ENSG00000133216	ENSMMUG00000023382	ENSRNOG00000012531	ENSMUSG00000028664
EPHB3	P54753	EPH receptor B3	ENSG00000182580	ENSMMUG00000003304	ENSRNOG00000031801	ENSMUSG00000005958
EPHB4	P54760	EPH receptor B4	ENSG00000196411	ENSMMUG00000007375	ENSRNOG00000045952	ENSMUSG00000029710
EPHB6	O15197	EPH receptor B6	ENSG00000106123	ENSMMUG00000016218	ENSRNOG00000014367	ENSMUSG00000029869
EPOR	P19235	erythropoietin receptor	ENSG00000187266	-	ENSRNOG00000012619	ENSMUSG00000006235
ERBB2	P04626	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2	ENSG00000141736	ENSMMUG00000001348	ENSRNOG00000006450	ENSMUSG00000062312
					ENSRNOG00000049238	
ERBB3	P21860	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	ENSG00000065361	ENSMMUG00000020758	ENSRNOG00000004964	ENSMUSG00000018166
ERBB4	Q15303	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 4	ENSG00000178568	ENSMMUG00000001674	ENSRNOG00000014248	ENSMUSG00000062209
ESR1	P03372	estrogen receptor 1	ENSG00000091831	ENSMMUG00000019045	ENSRNOG00000019358	ENSMUSG00000019768
ESR2	Q92731	estrogen receptor 2 (ER beta)	ENSG00000140009	ENSMMUG00000022839	ENSRNOG00000005343	ENSMUSG00000021055
F10	P00742	coagulation factor X	ENSG00000126218	ENSMMUG00000001038	ENSRNOG00000033444	ENSMUSG00000031444
F2	P00734	coagulation factor II (thrombin)	ENSG00000180210	ENSMMUG00000019674	ENSRNOG00000016325	ENSMUSG00000027249
F5	P12259	coagulation factor V (proaccelerin, labile factor)	ENSG00000198734	-	ENSRNOG00000002815	ENSMUSG00000026579
F8	P00451	coagulation factor VIII, procoagulant component	ENSG00000185010	ENSMMUG00000010245	ENSRNOG00000031197	ENSMUSG00000031196
			ENSG00000269261			
FAAH	O00519	fatty acid amide hydrolase	ENSG00000117480	ENSMMUG00000002714	ENSRNOG00000011019	ENSMUSG00000034171
					ENSRNOG00000045949	
FASN	P49327	fatty acid synthase	ENSG00000169710	ENSMMUG00000012811	ENSRNOG00000045636	ENSMUSG00000025153
FDPS	P14324	farnesyl diphosphate synthase	ENSG00000160752	ENSMMUG00000023144	-	ENSMUSG00000059743
FGA	P02671	fibrinogen alpha chain	ENSG00000171560	ENSMMUG00000021303	ENSRNOG00000024848	ENSMUSG00000028001
FGB	P02675	fibrinogen beta chain	ENSG00000171564	ENSMMUG00000021302	ENSRNOG00000007092	ENSMUSG00000033831
FGFR1	P11362	fibroblast growth factor receptor 1	ENSG00000077782	ENSMMUG00000022478	ENSRNOG00000016050	ENSMUSG00000031565
FGFR2	P21802	fibroblast growth factor receptor 2	ENSG00000066468	ENSMMUG00000009594	ENSRNOG00000016374	ENSMUSG00000030849
FGFR3	P22607	fibroblast growth factor receptor 3	ENSG00000068078	ENSMMUG00000019945	ENSRNOG00000016818	ENSMUSG00000054252
FGG	P02679	fibrinogen gamma chain	ENSG00000171557	ENSMMUG00000021304	ENSRNOG00000025074	ENSMUSG00000033860

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
FGR	P09769	feline Gardner-Rasheed sarcoma viral oncogene homolog	ENSG00000000938	ENSMMUG00000014434	ENSRNOG00000009912	ENSMUSG00000028874
FKBP1A	P62942	FK506 binding protein 1A, 12kDa	ENSG00000088832	ENSMMUG00000015996 ENSMMUG00000016459	-	-
FLT1	P17948	fms-related tyrosine kinase 1	ENSG00000102755	ENSMMUG00000021538	ENSRNOG00000000940	ENSMUSG00000029648
FLT3	P36888	fms-related tyrosine kinase 3	ENSG00000122025	ENSMMUG00000016139	ENSRNOG00000000949	ENSMUSG00000042817
FLT4	P35916	fms-related tyrosine kinase 4	ENSG00000037280	ENSMMUG00000018693	ENSRNOG00000002511	ENSMUSG00000020357
FN1	P02751	fibronectin 1	ENSG00000115414	ENSMMUG00000012321	ENSRNOG00000014288	ENSMUSG00000026193
FRK	P42685	fyn-related kinase	ENSG00000111816	ENSMMUG00000014635	ENSRNOG00000000543	ENSMUSG00000019779
FSHR	P23945	follicle stimulating hormone receptor	ENSG00000170820	ENSMMUG00000017196	ENSRNOG00000016783	ENSMUSG00000032937
FYN	P06241	FYN oncogene related to SRC, FGR, YES	ENSG00000010810	ENSMMUG00000017017	ENSRNOG00000000596	ENSMUSG00000019843
GAA	P10253	glucosidase, alpha; acid	ENSG00000171298	ENSMMUG00000009589	ENSRNOG00000047656	ENSMUSG00000025579
GABBR1	Q9UBS5	gamma-aminobutyric acid (GABA) B receptor, 1	ENSG00000232569 ENSG00000206466 ENSG00000237051 ENSG00000232632 ENSG00000237112 ENSG00000206511 ENSG00000204681	ENSMMUG00000021483	ENSRNOG00000000774	ENSMUSG00000024462
GABBR2	O75899	gamma-aminobutyric acid (GABA) B receptor, 2	ENSG00000136928	ENSMMUG00000009832	ENSRNOG00000008431	ENSMUSG00000039809
GABRA1	P14867	gamma-aminobutyric acid (GABA) A receptor, alpha 1	ENSG00000022355	ENSMMUG00000007296	ENSRNOG00000003512	ENSMUSG00000010803
GABRA2	P47869	gamma-aminobutyric acid (GABA) A receptor, alpha 2	ENSG00000151834	ENSMMUG00000003614	ENSRNOG00000002349	ENSMUSG00000000560
GABRA3	P34903	gamma-aminobutyric acid (GABA) A receptor, alpha 3	ENSG00000011677	ENSMMUG00000001540	ENSRNOG00000016845	ENSMUSG00000031343
GABRA4	P48169	gamma-aminobutyric acid (GABA) A receptor, alpha 4	ENSG00000109158	ENSMMUG00000009526	ENSRNOG00000002336	ENSMUSG00000029211
GABRA5	P31644	gamma-aminobutyric acid (GABA) A receptor, alpha 5	ENSG00000186297	ENSMMUG00000023554	ENSRNOG00000010803	ENSMUSG00000055078
GABRA6	Q16445	gamma-aminobutyric acid (GABA) A receptor, alpha 6	ENSG00000145863	ENSMMUG00000002832	ENSRNOG00000003569	ENSMUSG00000020428
GABRB1	P18505	gamma-aminobutyric acid (GABA) A receptor, beta 1	ENSG00000163288	ENSMMUG00000021776	ENSRNOG00000002327	ENSMUSG00000029212
GABRB2	P47870	gamma-aminobutyric acid (GABA) A receptor, beta 2	ENSG00000145864	ENSMMUG00000018044	ENSRNOG00000003680	ENSMUSG00000007653
GABRB3	P28472	gamma-aminobutyric acid (GABA) A receptor, beta 3	ENSG00000166206	ENSMMUG00000004314	ENSRNOG00000014938	ENSMUSG00000033676

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
GABRD	O14764	gamma-aminobutyric acid (GABA) A receptor, delta	ENSG00000187730	ENSMMUG00000019086	ENSRNOG00000016385	ENSMUSG00000029054
GABRE	P78334	gamma-aminobutyric acid (GABA) A receptor, epsilon	ENSG00000102287	ENSMMUG00000010936	ENSRNOG00000016533	ENSMUSG00000031340
GABRG1	Q8N1C3	gamma-aminobutyric acid (GABA) A receptor, gamma 1	ENSG00000163285	ENSMMUG00000012820	ENSRNOG00000002360	ENSMUSG00000001260
GABRG2	P18507	gamma-aminobutyric acid (GABA) A receptor, gamma 2	ENSG00000113327	ENSMMUG00000009707	ENSRNOG00000003241	ENSMUSG00000020436
GABRG3	Q99928	gamma-aminobutyric acid (GABA) A receptor, gamma 3	ENSG00000182256	ENSMMUG00000023555	ENSRNOG00000014862	ENSMUSG00000055026
GABRP	O00591	gamma-aminobutyric acid (GABA) A receptor, pi	ENSG00000094755	ENSMMUG00000000570	ENSRNOG00000032417	ENSMUSG00000020159
GABRQ	Q9UN88	gamma-aminobutyric acid (GABA) A receptor, theta	ENSG00000147402	ENSMMUG00000021679	ENSRNOG00000016884	ENSMUSG00000031344
GART	P22102	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	ENSG00000262473 ENSG00000159131	ENSMMUG00000022702	ENSRNOG00000028292	ENSMUSG00000022962
GCGR	P47871	glucagon receptor	ENSG00000215644	ENSMMUG00000000256	ENSRNOG00000036692	ENSMUSG00000025127
GGCX	P38435	gamma-glutamyl carboxylase	ENSG00000115486	ENSMMUG00000004446	ENSRNOG00000012975	ENSMUSG00000053460
GHR	P10912	growth hormone receptor	ENSG00000112964	ENSMMUG00000001336	ENSRNOG00000015654	ENSMUSG00000055737
GHRHR	Q02643	growth hormone releasing hormone receptor	ENSG00000106128	ENSMMUG00000011764	ENSRNOG00000011808	ENSMUSG00000004654
GLP1R	P43220	glucagon-like peptide 1 receptor	ENSG00000112164	ENSMMUG00000023387	ENSRNOG00000001152	ENSMUSG00000024027
GLP2R	O95838	glucagon-like peptide 2 receptor	ENSG00000065325	ENSMMUG00000018090	ENSRNOG00000003683	ENSMUSG00000049928
GLRA1	P23415	glycine receptor, alpha 1	ENSG00000145888	ENSMMUG00000014704	ENSRNOG00000013588	ENSMUSG00000000263
GLRB	P48167	glycine receptor, beta	ENSG00000109738	ENSMMUG00000022901	ENSRNOG00000010199	ENSMUSG00000028020
GNRHR	P30968	gonadotropin-releasing hormone receptor	ENSG00000109163	ENSMMUG00000003779	ENSRNOG00000002011	ENSMUSG00000029255
GRIA1	P42261	glutamate receptor, ionotropic, AMPA 1	ENSG00000155511	ENSMMUG00000009287	ENSRNOG00000045816	ENSMUSG00000020524
GRIA2	P42262	glutamate receptor, ionotropic, AMPA 2	ENSG00000120251	ENSMMUG00000018272	ENSRNOG00000028589	ENSMUSG00000033981
GRIA3	P42263	glutamate receptor, ionotropic, AMPA 3	ENSG00000125675	ENSMMUG00000009054	ENSRNOG00000007682	ENSMUSG00000001986
GRIA4	P48058	glutamate receptor, ionotropic, AMPA 4	ENSG00000152578	ENSMMUG00000010987	ENSRNOG00000006957	ENSMUSG00000025892
GRIK1	P39086	glutamate receptor, ionotropic, kainate 1	ENSG00000171189	ENSMMUG00000004886	ENSRNOG00000001575	ENSMUSG00000022935
GRIK2	Q13002	glutamate receptor, ionotropic, kainate 2	ENSG00000164418	ENSMMUG00000009856	ENSRNOG00000000368	ENSMUSG00000056073

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
				ENSMMUG00000029528		
				ENSMMUG00000029527		
GRIK3	Q13003	glutamate receptor, ionotropic, kainate 3	ENSG00000163873	ENSMMUG00000009679	ENSRNOG00000008992	ENSMUSG00000001985
GRIK4	Q16099	glutamate receptor, ionotropic, kainate 4	ENSG00000149403	ENSMMUG00000022347	ENSRNOG00000030910	ENSMUSG00000032017
GRIK5	Q16478	glutamate receptor, ionotropic, kainate 5	ENSG00000105737	ENSMMUG00000017431	ENSRNOG00000020310	ENSMUSG00000003378
GRIN1	Q05586	glutamate receptor, ionotropic, N-methyl D-aspartate 1	ENSG00000176884	ENSMMUG00000001661	ENSRNOG00000011726	ENSMUSG00000026959
GRIN2A	Q12879	glutamate receptor, ionotropic, N-methyl D-aspartate 2A	ENSG00000183454	ENSMMUG00000010142	ENSRNOG00000033942	ENSMUSG00000059003
GRIN2B	Q13224	glutamate receptor, ionotropic, N-methyl D-aspartate 2B	ENSG00000150086	ENSMMUG00000019112	ENSRNOG00000008766	ENSMUSG00000030209
GRIN2C	Q14957	glutamate receptor, ionotropic, N-methyl D-aspartate 2C	ENSG00000161509	ENSMMUG00000005744	ENSRNOG00000003280	ENSMUSG00000020734
GRIN2D	O15399	glutamate receptor, ionotropic, N-methyl D-aspartate 2D	ENSG00000105464	ENSMMUG00000005983	ENSRNOG00000021063	ENSMUSG00000002771
GRIN3A	Q8TCU5	glutamate receptor, ionotropic, N-methyl-D-aspartate 3A	ENSG00000198785	ENSMMUG00000016675	ENSRNOG00000005723	ENSMUSG00000039579
GRIN3B	O60391	glutamate receptor, ionotropic, N-methyl-D-aspartate 3B	ENSG00000116032	-	ENSRNOG00000012562	ENSMUSG00000035745
GSK3A	P49840	glycogen synthase kinase 3 alpha	ENSG00000105723	ENSMMUG00000014597	ENSRNOG00000020417	ENSMUSG00000057177
GSK3B	P49841	glycogen synthase kinase 3 beta	ENSG00000082701	ENSMMUG00000001980	ENSRNOG00000002833	ENSMUSG00000022812
GSR	P00390	glutathione reductase	ENSG00000104687	-	ENSRNOG00000014915	ENSMUSG00000031584
GUCY1B3	Q02153	guanylate cyclase 1, soluble, beta 3	ENSG00000061918	ENSMMUG00000005447	ENSRNOG00000012060	ENSMUSG00000028005
GUCY2C	P25092	guanylate cyclase 2C (heat stable enterotoxin receptor)	ENSG00000070019	ENSMMUG00000009870	ENSRNOG00000009031	ENSMUSG00000042638
HCK	P08631	hemopoietic cell kinase	ENSG00000101336	ENSMMUG00000000117	ENSRNOG00000009331	ENSMUSG00000003283
HDAC1	Q13547	histone deacetylase 1	ENSG00000116478	ENSMMUG00000009948	-	-
HDAC10	Q969S8	histone deacetylase 10	ENSG00000100429	ENSMMUG00000014101	ENSRNOG00000031915	ENSMUSG00000062906
HDAC11	Q96DB2	histone deacetylase 11	ENSG00000163517	ENSMMUG00000001369	ENSRNOG00000006824	ENSMUSG00000034245
HDAC2	Q92769	histone deacetylase 2	ENSG00000196591	ENSMMUG00000008937	ENSRNOG00000000604	ENSMUSG00000019777
HDAC3	O15379	histone deacetylase 3	ENSG00000171720	ENSMMUG00000012588	ENSRNOG00000019618	ENSMUSG00000024454
				ENSMMUG00000029538		
HDAC4	P56524	histone deacetylase 4	ENSG00000068024	ENSMMUG00000006077	ENSRNOG00000020372	ENSMUSG00000026313
HDAC5	Q9UQL6	histone deacetylase 5	ENSG00000108840	ENSMMUG00000006806	ENSRNOG00000020905	ENSMUSG00000008855
HDAC6	Q9UBN7	histone deacetylase 6	ENSG00000094631	ENSMMUG00000016465	ENSRNOG00000047281	ENSMUSG00000031161
			ENSG00000269101		ENSRNOG00000048738	
HDAC7	Q8WUI4	histone deacetylase 7	ENSG00000061273	ENSMMUG00000032343	ENSRNOG00000008308	ENSMUSG00000022475
HDAC8	Q9BY41	histone deacetylase 8	ENSG00000147099	ENSMMUG00000007592	ENSRNOG00000003122	ENSMUSG00000067567
			ENSG00000268498			
HDAC9	Q9UKV0	histone deacetylase 9	ENSG00000048052	ENSMMUG00000011170	ENSRNOG00000004158	ENSMUSG00000004698

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
HMGCR	P04035	3-hydroxy-3-methylglutaryl-CoA reductase	ENSG00000113161	ENSMMUG00000006139	ENSRNOG00000016122	ENSMUSG00000021670
HPD	P32754	4-hydroxyphenylpyruvate dioxygenase	ENSG00000158104	ENSMMUG00000016454	ENSRNOG00000001338	ENSMUSG00000029445
HRH1	P35367	histamine receptor H1	ENSG00000196639	ENSMMUG00000004817	ENSRNOG00000007420	ENSMUSG00000053004
HRH2	P25021	histamine receptor H2	ENSG00000113749	ENSMMUG00000014245	ENSRNOG00000018260	ENSMUSG00000034987
HSD3B2	P26439	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2	ENSG00000203859	ENSMMUG00000016568	ENSRNOG00000019417	ENSMUSG00000027869
					ENSRNOG00000019441	ENSMUSG00000063730
					ENSRNOG00000042884	ENSMUSG00000095388
						ENSMUSG00000027871
						ENSMUSG00000095143
						ENSMUSG00000062410
						ENSMUSG00000038092
						ENSMUSG00000090817
HTR1A	P08908	5-hydroxytryptamine (serotonin) receptor 1A, G protein-coupled	ENSG00000178394	ENSMMUG00000001789	ENSRNOG00000010254	ENSMUSG00000021721
HTR1B	P28222	5-hydroxytryptamine (serotonin) receptor 1B, G protein-coupled	ENSG00000135312	ENSMMUG00000015717	ENSRNOG00000013042	ENSMUSG00000049511
HTR1D	P28221	5-hydroxytryptamine (serotonin) receptor 1D, G protein-coupled	ENSG00000179546	ENSMMUG00000032222	ENSRNOG00000012038	ENSMUSG00000070687
HTR1F	P30939	5-hydroxytryptamine (serotonin) receptor 1F, G protein-coupled	ENSG00000179097	ENSMMUG00000005238	ENSRNOG00000000716	ENSMUSG00000050783
HTR2A	P28223	5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled	ENSG00000102468	ENSMMUG00000004210	ENSRNOG00000010063	ENSMUSG00000034997
HTR2B	P41595	5-hydroxytryptamine (serotonin) receptor 2B, G protein-coupled	ENSG00000135914	ENSMMUG00000011773	ENSRNOG00000017625	ENSMUSG00000026228
HTR2C	P28335	5-hydroxytryptamine (serotonin) receptor 2C, G protein-coupled	ENSG00000147246	ENSMMUG00000009607	ENSRNOG00000030877	ENSMUSG00000041380
HTR3A	P46098	5-hydroxytryptamine (serotonin) receptor 3A, ionotropic	ENSG00000166736	ENSMMUG00000007080	ENSRNOG00000006595	ENSMUSG00000032269
HTR4	Q13639	5-hydroxytryptamine (serotonin) receptor 4, G protein-coupled	ENSG00000164270	ENSMMUG00000016913	ENSRNOG00000019134	ENSMUSG00000026322
IFNAR1	P17181	interferon (alpha, beta and omega) receptor 1	ENSG00000142166	ENSMMUG00000003961	ENSRNOG00000028594	ENSMUSG00000022967
IFNAR2	P48551	interferon (alpha, beta and omega) receptor 2	ENSG00000159110	ENSMMUG00000003956	ENSRNOG00000045668	ENSMUSG00000022971
IFNGR1	P15260	interferon gamma receptor 1	ENSG00000027697	ENSMMUG00000012098	ENSRNOG00000012074	ENSMUSG00000020009
IFNGR2	P38484	interferon gamma receptor 2 (interferon gamma transducer 1)	ENSG00000262795	ENSMMUG00000005508	ENSRNOG00000002032	ENSMUSG00000022965
			ENSG00000159128			
IGF1R	P08069	insulin-like growth factor 1 receptor	ENSG00000140443	ENSMMUG00000012305	ENSRNOG00000014187	ENSMUSG00000005533

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
IGHE	P01854		ENSG00000211891	ENSMMUG00000030305	ENSRNOG00000005328	ENSMUSG00000087642
IL11RA	Q14626	interleukin 11 receptor, alpha	ENSG00000137070	ENSMMUG00000020790	ENSRNOG00000015068	ENSMUSG00000073876
						ENSMUSG00000095623
						ENSMUSG00000073889
						ENSMUSG00000095456
						ENSMUSG00000078735
IL12B	P29460	interleukin 12B (natural killer cell stimulatory factor 2, cytotoxic lymphocyte maturation factor 2, p40)	ENSG00000113302	ENSMMUG00000021701	ENSRNOG00000004380	ENSMUSG00000004296
IL1B	P01584	interleukin 1, beta	ENSG00000125538	ENSMMUG00000019778	ENSRNOG00000004649	ENSMUSG00000027398
IL1R1	P14778	interleukin 1 receptor, type I	ENSG00000115594	ENSMMUG00000020737	ENSRNOG00000014504	ENSMUSG00000026072
IL2RA	P01589	interleukin 2 receptor, alpha	ENSG00000134460	ENSMMUG00000015151	ENSRNOG00000047647	ENSMUSG00000026770
IL2RB	P14784	interleukin 2 receptor, beta	ENSG00000100385	ENSMMUG00000005417	ENSRNOG00000048636	ENSMUSG00000068227
IL6R	P08887	interleukin 6 receptor	ENSG00000160712	ENSMMUG00000022978	ENSRNOG00000020811	ENSMUSG00000027947
IMPA1	P29218	inositol(myo)-1(or 4)-monophosphatase 1	ENSG00000133731	ENSMMUG00000002628	ENSRNOG00000010482	ENSMUSG00000027531
IMPDH1	P20839	IMP (inosine 5'-monophosphate) dehydrogenase 1	ENSG00000106348	ENSMMUG00000014509	ENSRNOG00000020032	ENSMUSG00000003500
IMPDH2	P12268	IMP (inosine 5'-monophosphate) dehydrogenase 2	ENSG00000178035	ENSMMUG00000005656	ENSRNOG00000031965	ENSMUSG00000062867
INSR	P06213	insulin receptor	ENSG00000171105	ENSMMUG00000003716	ENSRNOG00000029986	ENSMUSG00000005534
				ENSMMUG00000028907		
ITGA2B	P08514	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	ENSG00000005961	ENSMMUG00000012659	ENSRNOG00000022071	ENSMUSG00000034664
ITGB1	P05556	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	ENSG00000150093	ENSMMUG00000030624	ENSRNOG00000010966	ENSMUSG00000025809
				ENSMMUG00000017767		
ITGB2	P05107	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	ENSG00000160255	ENSMMUG00000000951	ENSRNOG00000001224	ENSMUSG00000000290
					ENSRNOG00000028203	ENSMUSG00000000157
ITGB3	P05106	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	ENSG00000259207	ENSMMUG00000003289	ENSRNOG00000048449	ENSMUSG00000020689
ITGB7	P26010	integrin, beta 7	ENSG00000139626	ENSMMUG00000005245	ENSRNOG00000012208	ENSMUSG00000001281
ITK	Q08881	IL2-inducible T-cell kinase	ENSG00000113263	ENSMMUG00000010995	ENSRNOG00000006860	ENSMUSG00000020395
JAK1	P23458	Janus kinase 1	ENSG00000162434	ENSMMUG00000018566	ENSRNOG00000011157	ENSMUSG00000028530
JAK2	O60674	Janus kinase 2	ENSG00000096968	ENSMMUG00000014558	ENSRNOG00000015547	ENSMUSG00000024789
JAK3	P52333	Janus kinase 3	ENSG00000105639	ENSMMUG00000010996	ENSRNOG00000018669	ENSMUSG00000031805

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
KCNA1	Q09470	potassium voltage-gated channel, shaker-related subfamily, member 1 (episodic ataxia with myokymia)	ENSG00000111262	ENSMMUG00000012274	ENSRNOG00000019750	ENSMUSG00000047976
KCNA10	Q16322	potassium voltage-gated channel, shaker-related subfamily, member 10	ENSG00000143105	ENSMMUG00000016056	ENSRNOG00000050416 ENSRNOG00000018283	ENSMUSG00000042861
KCNA2	P16389	potassium voltage-gated channel, shaker-related subfamily, member 2	ENSG00000177301	ENSMMUG00000018680	ENSRNOG00000018285	ENSMUSG00000040724
KCNA3	P22001	potassium voltage-gated channel, shaker-related subfamily, member 3	ENSG00000177272	ENSMMUG00000018684	ENSRNOG00000018116	ENSMUSG00000047959
KCNA4	P22459	potassium voltage-gated channel, shaker-related subfamily, member 4	ENSG00000182255	ENSMMUG00000031030	ENSRNOG00000004918	ENSMUSG00000042604
KCNA5	P22460	potassium voltage-gated channel, shaker-related subfamily, member 5	ENSG00000130037	ENSMMUG00000022887	ENSRNOG00000019719	ENSMUSG00000045534
KCNA6	P17658	potassium voltage-gated channel, shaker-related subfamily, member 6	ENSG00000151079	ENSMMUG00000020816	ENSRNOG00000047736	ENSMUSG00000038077
KCNA7	Q96RP8	potassium voltage-gated channel, shaker-related subfamily, member 7	ENSG00000104848	-	ENSRNOG00000020779	ENSMUSG00000038201
KCNB1	Q14721	potassium voltage-gated channel, Shab-related subfamily, member 1	ENSG00000158445	ENSMMUG00000006442	ENSRNOG00000046949	ENSMUSG00000050556
KCNB2	Q92953	potassium voltage-gated channel, Shab-related subfamily, member 2	ENSG00000182674	ENSMMUG00000013770	ENSRNOG00000028991	ENSMUSG00000092083
KCNC1	P48547	potassium voltage-gated channel, Shaw-related subfamily, member 1	ENSG00000129159	ENSMMUG00000006403	ENSRNOG00000011380	ENSMUSG00000058975
KCNC2	Q96PR1	potassium voltage-gated channel, Shaw-related subfamily, member 2	ENSG00000166006	ENSMMUG00000012362	ENSRNOG00000004077	ENSMUSG00000035681
KCNC3	Q14003	potassium voltage-gated channel, Shaw-related subfamily, member 3	ENSG00000131398	ENSMMUG00000001426	ENSRNOG00000019959	ENSMUSG00000062785
KCNC4	Q03721	potassium voltage-gated channel, Shaw-related subfamily, member 4	ENSG00000116396	ENSMMUG00000004614	ENSRNOG00000050924	ENSMUSG00000027895

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
KCND1	Q9NSA2	potassium voltage-gated channel, Shal-related subfamily, member 1	ENSG00000102057 ENSG00000267829	ENSMMUG00000004559	ENSRNOG00000039544	ENSMUSG00000009731
KCND2	Q9NZV8	potassium voltage-gated channel, Shal-related subfamily, member 2	ENSG00000184408	ENSMMUG00000023308	ENSRNOG00000029610	ENSMUSG00000060882
KCND3	Q9UK17	potassium voltage-gated channel, Shal-related subfamily, member 3	ENSG00000171385	ENSMMUG00000014490	ENSRNOG00000014686	ENSMUSG00000040896
KCNF1	Q9H3M0	potassium voltage-gated channel, subfamily F, member 1	ENSG00000162975	ENSMMUG00000002825	ENSRNOG00000024310	ENSMUSG00000051726
KCNG1	Q9UIX4	potassium voltage-gated channel, subfamily G, member 1	ENSG00000026559	ENSMMUG00000000683	ENSRNOG00000012144	ENSMUSG00000074575
KCNG2	Q9UJ96	potassium voltage-gated channel, subfamily G, member 2	ENSG00000178342	ENSMMUG00000032486	ENSRNOG00000017189	ENSMUSG00000059852
KCNG3	Q8TAE7	potassium voltage-gated channel, subfamily G, member 3	ENSG00000171126	ENSMMUG00000018583	ENSRNOG00000004535	ENSMUSG00000045053
KCNG4	Q8TDN1	potassium voltage-gated channel, subfamily G, member 4	ENSG00000168418	ENSMMUG00000000540	ENSRNOG00000015746	ENSMUSG00000045246
KCNH1	O95259	potassium voltage-gated channel, subfamily H (eag-related), member 1	ENSG00000143473	ENSMMUG00000016102	ENSRNOG00000003841	ENSMUSG00000058248
KCNH2	Q12809	potassium voltage-gated channel, subfamily H (eag-related), member 2	ENSG00000055118	ENSMMUG00000020436	ENSRNOG00000009872	ENSMUSG00000038319
KCNH3	Q9ULD8	potassium voltage-gated channel, subfamily H (eag-related), member 3	ENSG00000135519	ENSMMUG00000011097	ENSRNOG00000032075	ENSMUSG00000037579
KCNH4	Q9UQ05	potassium voltage-gated channel, subfamily H (eag-related), member 4	ENSG00000089558	ENSMMUG00000006401	ENSRNOG00000018790	ENSMUSG00000035355
KCNH5	Q8NCM2	potassium voltage-gated channel, subfamily H (eag-related), member 5	ENSG00000140015	ENSMMUG00000023518	ENSRNOG00000009542	ENSMUSG00000034402
KCNH6	Q9H252	potassium voltage-gated channel, subfamily H (eag-related), member 6	ENSG00000173826	ENSMMUG00000005183	ENSRNOG00000008078	ENSMUSG00000001901
KCNH7	Q9NS40	potassium voltage-gated channel, subfamily H (eag-related), member 7	ENSG00000184611	ENSMMUG00000019591	ENSRNOG00000007528	ENSMUSG00000059742
KCNH8	Q96L42	potassium voltage-gated channel, subfamily H (eag-related), member 8	ENSG00000183960	ENSMMUG00000005370	ENSRNOG00000012711	ENSMUSG00000035580
KCNJ11	Q14654	potassium inwardly-rectifying channel, subfamily J, member 11	ENSG00000187486	ENSMMUG00000023580	ENSRNOG00000021128	ENSMUSG00000096146

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
KCNJ8	Q15842	potassium inwardly-rectifying channel, subfamily J, member 8	ENSG00000121361	ENSMMUG00000020805	ENSRNOG00000013463	ENSMUSG00000030247
KCNK10	P57789	potassium channel, subfamily K, member 10	ENSG00000100433	ENSMMUG00000021010	ENSRNOG00000003813	ENSMUSG00000033854
KCNK18	Q7Z418	potassium channel, subfamily K, member 18	ENSG00000186795	ENSMMUG00000014815	ENSRNOG00000032706	ENSMUSG00000040901
KCNK2	O95069	potassium channel, subfamily K, member 2	ENSG00000082482	-	ENSRNOG00000002653	ENSMUSG00000037624
KCNK3	O14649	potassium channel, subfamily K, member 3	ENSG00000171303	ENSMMUG00000015933	ENSRNOG00000009790	ENSMUSG00000049265
KCNK9	Q9NPC2	potassium channel, subfamily K, member 9	ENSG00000169427	ENSMMUG00000005826	ENSRNOG00000009265	ENSMUSG00000036760
KCNN4	O15554	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 4	ENSG00000104783	ENSMMUG00000021267	ENSRNOG00000019440	ENSMUSG00000054342
KCNQ1	P51787	potassium voltage-gated channel, KQT-like subfamily, member 1	ENSG00000053918	ENSMMUG00000007742	ENSRNOG00000020532	ENSMUSG00000009545
KCNQ2	O43526	potassium voltage-gated channel, KQT-like subfamily, member 2	ENSG00000075043	ENSMMUG00000015918	ENSRNOG00000011624	ENSMUSG00000016346
KCNQ3	O43525	potassium voltage-gated channel, KQT-like subfamily, member 3	ENSG00000184156	ENSMMUG00000000892	ENSRNOG00000005206	ENSMUSG00000056258
KCNQ4	P56696	potassium voltage-gated channel, KQT-like subfamily, member 4	ENSG00000117013	ENSMMUG00000019536	ENSRNOG00000010576	ENSMUSG00000028631
KCNQ5	Q9NR82	potassium voltage-gated channel, KQT-like subfamily, member 5	ENSG00000185760	ENSMMUG00000015320	ENSRNOG00000013781	ENSMUSG00000028033
KCNS1	Q96KK3	potassium voltage-gated channel, delayed-rectifier, subfamily S, member 1	ENSG00000124134	ENSMMUG00000013214	ENSRNOG00000013681	ENSMUSG00000040164
KCNS2	Q9ULS6	potassium voltage-gated channel, delayed-rectifier, subfamily S, member 2	ENSG00000156486	ENSMMUG00000010267	ENSRNOG00000011369	ENSMUSG00000050963
KCNS3	Q9BQ31	potassium voltage-gated channel, delayed-rectifier, subfamily S, member 3	ENSG00000170745	ENSMMUG00000017198	ENSRNOG00000004899	ENSMUSG00000043673
KCNV1	Q6PIU1	potassium channel, subfamily V, member 1	ENSG00000164794	ENSMMUG00000018706	ENSRNOG00000004117	ENSMUSG00000022342
KCNV2	Q8TDN2	potassium channel, subfamily V, member 2	ENSG00000168263	ENSMMUG00000015105	ENSRNOG00000012566	ENSMUSG00000047298
KDR	P35968	kinase insert domain receptor (a type III receptor tyrosine kinase)	ENSG00000128052	ENSMMUG00000012669	ENSRNOG00000046829	ENSMUSG00000062960
KEAP1	Q14145	kelch-like ECH-associated protein 1	ENSG00000079999	ENSMMUG00000005218	ENSRNOG00000020878	ENSMUSG00000003308
KIT	P10721	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	ENSG00000157404	ENSMMUG00000003057	ENSRNOG00000002227	ENSMUSG00000005672

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
KLKB1	P03952	kallikrein B, plasma (Fletcher factor) 1	ENSG00000164344	ENSMMUG00000011633	ENSRNOG00000014118	ENSMUSG00000031640
LAMA1	P25391	laminin, alpha 1	ENSG00000101680	ENSMMUG00000011408 ENSMMUG00000011404	ENSRNOG00000017237	ENSMUSG00000032796
LAMA2	P24043	laminin, alpha 2	ENSG00000196569	ENSMMUG00000022136	ENSRNOG00000011134	ENSMUSG00000019899
LAMA3	Q16787	laminin, alpha 3	ENSG00000053747	ENSMMUG00000004321	ENSRNOG00000011300	ENSMUSG00000024421
LAMA4	Q16363	laminin, alpha 4	ENSG00000263699 ENSG00000112769	ENSMMUG00000013782	ENSRNOG00000000599	ENSMUSG00000019846
LAMA5	O15230	laminin, alpha 5	ENSG00000130702	ENSMMUG00000031416 ENSMMUG00000012627	ENSRNOG00000006479	ENSMUSG00000015647
LAMB1	P07942	laminin, beta 1	ENSG00000091136	ENSMMUG00000006567	ENSRNOG00000005678	ENSMUSG00000002900
LAMB2	P55268	laminin, beta 2 (laminin S)	ENSG00000172037	ENSMMUG00000002952	ENSRNOG00000047768	ENSMUSG00000052911
LAMB3	Q13751	laminin, beta 3	ENSG00000196878	ENSMMUG00000016925	ENSRNOG00000006025	ENSMUSG00000026639
LAMB4	A4D0S4	laminin, beta 4	ENSG00000091128	ENSMMUG00000021289	-	-
LAMC1	P11047	laminin, gamma 1 (formerly LAMB2)	ENSG00000135862	ENSMMUG00000015084	ENSRNOG00000002680	ENSMUSG00000026478
LAMC2	Q13753	laminin, gamma 2	ENSG00000058085	ENSMMUG00000001479	ENSRNOG00000002667	ENSMUSG00000026479
LAMC3	Q9Y6N6	laminin, gamma 3	ENSG00000050555	ENSMMUG00000002267	ENSRNOG00000009769	ENSMUSG00000026840
LCK	P06239	lymphocyte-specific protein tyrosine kinase	ENSG00000182866	-	ENSRNOG00000009705	ENSMUSG00000000409
LHCGR	P22888	luteinizing hormone/choriogonadotropin receptor	ENSG00000138039	ENSMMUG00000010368	ENSRNOG00000016712	ENSMUSG00000024107
LIPF	P07098	lipase, gastric	ENSG00000182333	ENSMMUG00000023148	ENSRNOG00000019448	ENSMUSG00000024768
LYN	P07948	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog	ENSG00000254087	ENSMMUG00000014149	ENSRNOG00000008180	ENSMUSG00000042228
MAOA	P21397	monoamine oxidase A	ENSG00000189221	ENSMMUG00000031960 ENSMMUG00000031961 ENSMMUG00000019489	ENSRNOG00000002848	ENSMUSG00000025037
MAOB	P27338	monoamine oxidase B	ENSG00000069535	ENSMMUG00000019492	ENSRNOG00000029778	ENSMUSG00000040147
MAP2K1	Q02750	mitogen-activated protein kinase kinase 1	ENSG00000169032	ENSMMUG00000029016	ENSRNOG00000010176	ENSMUSG00000004936
MAP2K2	P36507	mitogen-activated protein kinase kinase 2	ENSG00000126934	ENSMMUG00000019564	ENSRNOG00000020005	ENSMUSG00000035027
MAPK1	P28482	mitogen-activated protein kinase 1	ENSG00000100030	ENSMMUG00000012524	ENSRNOG00000001849	ENSMUSG00000063358
MAPK11	Q15759	mitogen-activated protein kinase 11	ENSG00000185386	ENSMMUG00000014107	ENSRNOG00000006984	ENSMUSG00000053137
MAPK3	P27361	mitogen-activated protein kinase 3	ENSG00000102882	ENSMMUG00000017758	-	ENSMUSG00000063065
MC2R	Q01718	melanocortin 2 receptor (adrenocorticotropic hormone)	ENSG00000185231	ENSMMUG00000016681	ENSRNOG00000016681	ENSMUSG00000045569
MET	P08581	met proto-oncogene (hepatocyte growth factor receptor)	ENSG00000105976	ENSMMUG00000006779	ENSRNOG00000007079	ENSMUSG00000009376

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
MGAM	O43451	maltase-glucoamylase (alpha-glucosidase)	ENSG00000257335	ENSMMUG00000016273	ENSRNOG00000026177	ENSMUSG00000068587
			ENSG00000259858	ENSMMUG00000029357		
				ENSMMUG00000029356		
MMP1	P03956	matrix metallopeptidase 1 (interstitial collagenase)	ENSG00000196611	ENSMMUG00000002037	ENSRNOG00000008881	ENSMUSG00000043089
					ENSRNOG00000032353	ENSMUSG00000041620
MMP13	P45452	matrix metallopeptidase 13 (collagenase 3)	ENSG00000137745	ENSMMUG00000004253	ENSRNOG00000008478	ENSMUSG00000050578
			ENSG00000262325			
MMP7	P09237	matrix metallopeptidase 7 (matrilysin, uterine)	ENSG00000137673	ENSMMUG00000003405	ENSRNOG00000010507	ENSMUSG00000018623
MMP8	P22894	matrix metallopeptidase 8 (neutrophil collagenase)	ENSG00000118113	ENSMMUG00000004489	ENSRNOG00000009907	ENSMUSG00000005800
MPL	P40238	myeloproliferative leukemia virus oncogene	ENSG00000117400	ENSMMUG00000003148	ENSRNOG00000028377	ENSMUSG00000006389
MS4A1	P11836	membrane-spanning 4-domains, subfamily A, member 1	ENSG00000156738	ENSMMUG00000001260	ENSRNOG00000020945	ENSMUSG00000024673
MT-ND1	P03886	NADH dehydrogenase, subunit 1 (complex I)	ENSG00000198888	ENSMMUG00000028699	ENSRNOG00000030644	ENSMUSG00000064341
MT-ND2	P03891	MTND2	ENSG00000198763	ENSMMUG00000028695	ENSRNOG00000031033	ENSMUSG00000064345
MT-ND3	P03897	NADH dehydrogenase, subunit 3 (complex I)	ENSG00000198840	ENSMMUG00000028680	ENSRNOG00000033615	ENSMUSG00000064360
MT-ND4	P03905	NADH dehydrogenase, subunit 4 (complex I)	ENSG00000198886	ENSMMUG00000028677	ENSRNOG00000029707	ENSMUSG00000064363
MT-ND4L	P03901	NADH dehydrogenase, subunit 4L (complex I)	ENSG00000212907	ENSMMUG00000028678	ENSRNOG00000031053	ENSMUSG00000065947
				ENSMMUG00000029085		ENSMUSG00000067736
				ENSMMUG00000031612		
MT-ND5	P03915	NADH dehydrogenase, subunit 5 (complex I)	ENSG00000198786	ENSMMUG00000028673	ENSRNOG00000029971	ENSMUSG00000064367
						ENSMUSG00000096449
MT-ND6	P03923	NADH dehydrogenase, subunit 6 (complex I)	ENSG00000198695	ENSMMUG00000028672	ENSRNOG00000029042	ENSMUSG00000064368
MTNR1A	P48039	melatonin receptor 1A	ENSG00000168412	ENSMMUG00000018050	ENSRNOG00000028744	ENSMUSG00000054764
MTNR1B	P49286	melatonin receptor 1B	ENSG00000134640	ENSMMUG00000017739	ENSRNOG00000008972	ENSMUSG00000050901
MTTP	P55157	microsomal triglyceride transfer protein	ENSG00000138823	ENSMMUG00000009091	ENSRNOG00000010655	ENSMUSG00000028158
NDUFA1	O15239	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 1, 7.5kDa	ENSG00000125356	ENSMMUG00000032503	ENSRNOG00000040005	ENSMUSG00000016427
				ENSMMUG00000032194		
				ENSMMUG00000032359		

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NDUFA10	O95299	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42kDa	ENSG00000130414	ENSMMUG00000010382	ENSRNOG00000016470	ENSMUSG00000026260
NDUFA11	Q86Y39	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 11, 14.7kDa	ENSG00000174886	ENSMMUG00000016665	ENSRNOG00000048320	ENSMUSG00000002379
NDUFA12	Q9UI09	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 12	ENSG00000184752	ENSMMUG00000018897	ENSRNOG00000007407	ENSMUSG00000020022
NDUFA13	Q9P0J0	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13	ENSG00000186010	ENSMMUG00000011350	ENSRNOG00000020602	ENSMUSG00000036199
					ENSRNOG00000021688	
NDUFA2	O43678	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa	ENSG00000131495	ENSMMUG00000010011	ENSRNOG00000043512	ENSMUSG00000014294
					ENSRNOG00000017571	
NDUFA3	O95167	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 3, 9kDa	ENSG00000170906	-	ENSRNOG00000030778	ENSMUSG00000035674
			ENSG00000272487		ENSRNOG00000029868	
					ENSRNOG00000014224	
					ENSRNOG00000050514	
NDUFA4	O00483	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4, 9kDa	ENSG00000189043	ENSMMUG00000020894	ENSRNOG00000005512	ENSMUSG00000029632
				ENSMMUG00000030210		
				ENSMMUG00000029771		
NDUFA4L2	Q9NRX3	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4-like 2	ENSG00000185633	ENSMMUG00000021786	ENSRNOG00000031851	ENSMUSG00000040280
NDUFA5	Q16718	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5	ENSG00000128609	ENSMMUG00000016341	ENSRNOG00000005698	ENSMUSG00000023089
				ENSMMUG00000017321		
				ENSMMUG00000011538		
NDUFA6	P56556	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 6, 14kDa	ENSG00000184983	ENSMMUG00000013060	ENSRNOG00000008569	ENSMUSG00000022450
NDUFA7	O95182	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 7, 14.5kDa	ENSG00000167774	ENSMMUG00000016697	ENSRNOG00000006939	ENSMUSG00000041881

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NDUFA8	P51970	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8, 19kDa	ENSG00000119421	ENSMMUG00000011450	ENSRNOG00000005668	ENSMUSG00000026895
NDUFA9	Q16795	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 9, 39kDa	ENSG00000139180	ENSMMUG00000020813	ENSRNOG00000026930	ENSMUSG00000000399
NDUFAB1	O14561	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	ENSG00000004779	ENSMMUG00000020522	ENSRNOG00000018129	ENSMUSG00000030869
				ENSMMUG00000031759	ENSRNOG00000025242	
NDUFAF1	Q9Y375	NADH dehydrogenase (ubiquinone) complex I, assembly factor 1	ENSG00000137806	-	ENSRNOG00000005006	ENSMUSG00000027305
NDUFAF2	Q8N183	NADH dehydrogenase (ubiquinone) complex I, assembly factor 2	ENSG00000164182	ENSMMUG00000021813	-	ENSMUSG00000068184
NDUFAF3	Q9BU61	NADH dehydrogenase (ubiquinone) complex I, assembly factor 3	ENSG00000178057	ENSMMUG00000005654	ENSRNOG00000020068	ENSMUSG00000070283
NDUFAF4	Q9P032	NADH dehydrogenase (ubiquinone) complex I, assembly factor 4	ENSG00000123545	ENSMMUG00000018843	ENSRNOG00000007506	ENSMUSG00000028261
					ENSRNOG00000003342	
NDUFB1	O75438	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7kDa	ENSG00000183648	ENSMMUG00000022428	-	-
NDUFB10	O96000	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 10, 22kDa	ENSG00000140990	ENSMMUG00000029268	ENSRNOG00000014568	ENSMUSG00000040048
					ENSRNOG00000031321	
NDUFB11	Q9NX14	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 11, 17.3kDa	ENSG00000147123	ENSMMUG00000017422	ENSRNOG00000008329	ENSMUSG00000031059
					ENSRNOG00000005572	ENSMUSG00000061633
NDUFB2	O95178	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 2, 8kDa	ENSG00000090266	ENSMMUG00000015284	ENSRNOG00000026616	ENSMUSG00000002416
NDUFB3	O43676	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 3, 12kDa	ENSG00000119013	ENSMMUG00000030840	-	ENSMUSG00000026032
				ENSMMUG00000030819		
				ENSMMUG00000012225		

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NDUFB4	O95168	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 4, 15kDa	ENSG00000065518	ENSMMUG00000001592	ENSRNOG00000030089 ENSRNOG00000032733 ENSRNOG00000034289 ENSRNOG00000034182 ENSRNOG00000002721	ENSMUSG00000022820
NDUFB5	O43674	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 5, 16kDa	ENSG00000136521	ENSMMUG00000009746	ENSRNOG00000011949	ENSMUSG00000027673
NDUFB6	O95139	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6, 17kDa	ENSG00000165264	ENSMMUG00000002685	ENSRNOG00000024539	ENSMUSG00000071014
NDUFB7	P17568	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 7, 18kDa	ENSG00000099795	ENSMMUG00000021612	ENSRNOG00000028717	ENSMUSG00000033938
NDUFB8	O95169	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 8, 19kDa	ENSG00000166136	ENSMMUG00000006636	ENSRNOG00000014078	ENSMUSG00000025204 ENSMUSG00000091471
NDUFB9	Q9Y6M9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9, 22kDa	ENSG00000147684	ENSMMUG00000015493	ENSRNOG00000009364	ENSMUSG00000022354
NDUFC1	O43677	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1, 6kDa	ENSG00000109390	ENSMMUG00000018182	ENSRNOG00000038218	ENSMUSG00000037152
NDUFC2	O95298	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 2, 14.5kDa	ENSG00000151366	ENSMMUG00000017471	ENSRNOG00000012383	ENSMUSG00000030647
NDUFS1	P28331	NADH dehydrogenase (ubiquinone) Fe-S protein 1, 75kDa (NADH-coenzyme Q reductase)	ENSG00000023228	ENSMMUG00000016788	ENSRNOG00000011849	ENSMUSG00000025968
NDUFS2	O75306	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49kDa (NADH-coenzyme Q reductase)	ENSG00000158864	ENSMMUG00000004520	ENSRNOG00000038372	ENSMUSG00000013593
NDUFS3	O75489	NADH dehydrogenase (ubiquinone) Fe-S protein 3, 30kDa (NADH-coenzyme Q reductase)	ENSG00000213619	ENSMMUG00000004286	ENSRNOG00000009155	ENSMUSG00000005510

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NDUFS4	O43181	NADH dehydrogenase (ubiquinone) Fe-S protein 4, 18kDa (NADH-coenzyme Q reductase)	ENSG00000164258	ENSMMUG00000004861	ENSRNOG00000011383	ENSMUSG00000021764
NDUFS5	O43920	NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)	ENSG00000168653	ENSMMUG00000021561	ENSRNOG00000026646	ENSMUSG00000028648
				ENSMMUG00000010773	ENSRNOG00000030585	
				ENSMMUG00000030508	ENSRNOG00000029339	
NDUFS6	O75380	NADH dehydrogenase (ubiquinone) Fe-S protein 6, 13kDa (NADH-coenzyme Q reductase)	ENSG00000145494	ENSMMUG00000004447	ENSRNOG00000023387	ENSMUSG00000021606
					ENSRNOG00000018068	
					ENSRNOG00000049394	
NDUFS7	O75251	NADH dehydrogenase (ubiquinone) Fe-S protein 7, 20kDa (NADH-coenzyme Q reductase)	ENSG00000115286	ENSMMUG00000003045	ENSRNOG00000024568	ENSMUSG00000020153
NDUFS8	O00217	NADH dehydrogenase (ubiquinone) Fe-S protein 8, 23kDa (NADH-coenzyme Q reductase)	ENSG00000110717	ENSMMUG00000023189	ENSRNOG00000017446	ENSMUSG00000059734
NDUFV1	P49821	NADH dehydrogenase (ubiquinone) flavoprotein 1, 51kDa	ENSG00000167792	ENSMMUG00000018526	ENSRNOG00000018117	ENSMUSG00000037916
NDUFV2	P19404	NADH dehydrogenase (ubiquinone) flavoprotein 2, 24kDa	ENSG00000178127	ENSMMUG00000001354	ENSRNOG00000042503	ENSMUSG00000024099
				ENSMMUG00000022095		
NDUFV3	P56181	NADH dehydrogenase (ubiquinone) flavoprotein 3, 10kDa	ENSG00000160194	ENSMMUG00000001553	ENSRNOG00000027593	ENSMUSG00000024038
					ENSRNOG00000001182	
NPC1L1	Q9UHC9	NPC1-like 1	ENSG00000015520	ENSMMUG00000000782	ENSRNOG00000049955	ENSMUSG00000020447
NPR1	P16066	natriuretic peptide receptor A/guanylate cyclase A (atrionatriuretic peptide receptor A)	ENSG00000169418	ENSMMUG00000008153	ENSRNOG00000014684	ENSMUSG00000027931
NR1H4	Q96RI1	nuclear receptor subfamily 1, group H, member 4	ENSG00000012504	ENSMMUG00000003571	ENSRNOG00000007197	ENSMUSG00000047638

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NR3C1	P04150	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	ENSG00000113580	ENSMMUG00000000421	ENSRNOG00000014096	ENSMUSG00000024431
NR3C2	P08235	nuclear receptor subfamily 3, group C, member 2	ENSG00000151623	ENSMMUG00000014955	ENSRNOG00000034007	ENSMUSG00000031618
ODC1	P11926	ornithine decarboxylase 1	ENSG00000115758	ENSMMUG00000030822 ENSMMUG00000020441	ENSRNOG00000005424	ENSMUSG00000011179
OPRD1	P41143	opioid receptor, delta 1	ENSG00000116329	ENSMMUG00000009389	ENSRNOG00000010531	ENSMUSG00000050511
OPRK1	P41145	opioid receptor, kappa 1	ENSG00000082556	ENSMMUG00000018372	ENSRNOG00000007647	ENSMUSG00000025905
OPRM1	P35372	opioid receptor, mu 1	ENSG00000112038	ENSMMUG00000022691	ENSRNOG00000018191	ENSMUSG00000000766
OXTR	P30559	oxytocin receptor	ENSG00000180914	ENSMMUG00000009703	ENSRNOG00000005806	ENSMUSG00000049112
P2RY12	Q9H244	purinergic receptor P2Y, G-protein coupled, 12	ENSG00000169313	ENSMMUG00000001105	ENSRNOG00000013902	ENSMUSG00000036353
PAH	P00439	phenylalanine hydroxylase	ENSG00000171759	ENSMMUG00000011908	ENSRNOG00000004302	ENSMUSG00000020051
PDE10A	Q9Y233	phosphodiesterase 10A	ENSG00000112541	ENSMMUG00000023614	ENSRNOG00000011310	ENSMUSG00000023868
PDE1A	P54750	phosphodiesterase 1A, calmodulin-dependent	ENSG00000115252	ENSMMUG00000001163	ENSRNOG00000006154	ENSMUSG00000059173
PDE1B	Q01064	phosphodiesterase 1B, calmodulin-dependent	ENSG00000123360	ENSMMUG00000000111	ENSRNOG00000036828	ENSMUSG00000022489
PDE1C	Q14123	phosphodiesterase 1C, calmodulin-dependent 70kDa	ENSG00000154678	ENSMMUG00000001625	ENSRNOG00000012337	ENSMUSG00000004347
PDE2A	O00408	phosphodiesterase 2A, cGMP-stimulated	ENSG00000186642	ENSMMUG00000015051	ENSRNOG00000019560	ENSMUSG00000030653
PDE3A	Q14432	phosphodiesterase 3A, cGMP-inhibited	ENSG00000172572	ENSMMUG00000008269	ENSRNOG00000025042	ENSMUSG00000041741
PDE3B	Q13370	phosphodiesterase 3B, cGMP-inhibited	ENSG00000261923 ENSG00000152270	ENSMMUG00000000759	ENSRNOG00000011417	ENSMUSG00000030671
PDE4A	P27815	phosphodiesterase 4A, cAMP-specific	ENSG00000065989	ENSMMUG00000016122	ENSRNOG00000020828	ENSMUSG00000032177
PDE4B	Q07343	phosphodiesterase 4B, cAMP-specific	ENSG00000184588	ENSMMUG00000007490	ENSRNOG00000005905	ENSMUSG00000028525
PDE4C	Q08493	phosphodiesterase 4C, cAMP-specific	ENSG00000105650	ENSMMUG00000005028	ENSRNOG00000019518	ENSMUSG00000031842
PDE4D	Q08499	phosphodiesterase 4D, cAMP-specific	ENSG00000113448	ENSMMUG00000012132	ENSRNOG00000042536	ENSMUSG00000021699
PDE5A	O76074	phosphodiesterase 5A, cGMP-specific	ENSG00000138735	ENSMMUG00000007305	ENSRNOG00000014443	ENSMUSG00000053965
PDE6A	P16499	phosphodiesterase 6A, cGMP-specific, rod, alpha	ENSG00000132915	ENSMMUG00000004752	ENSRNOG00000017816	ENSMUSG00000024575
PDE6B	P35913	phosphodiesterase 6B, cGMP-specific, rod, beta	ENSG00000133256	ENSMMUG00000003282	ENSRNOG00000000065	ENSMUSG00000029491
PDE6C	P51160	phosphodiesterase 6C, cGMP-specific, cone, alpha prime	ENSG00000095464	ENSMMUG00000014713	ENSRNOG00000016593	ENSMUSG00000024992

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
PDE6D	O43924	phosphodiesterase 6D, cGMP-specific, rod, delta	ENSG00000156973	ENSMMUG00000019565	ENSRNOG00000018610	ENSMUSG00000026239
PDE6G	P18545	phosphodiesterase 6G, cGMP-specific, rod, gamma	ENSG00000185527	ENSMMUG00000001588	ENSRNOG00000046962	ENSMUSG00000025386
PDE6H	Q13956	phosphodiesterase 6H, cGMP-specific, cone, gamma	ENSG00000139053	ENSMMUG00000001009	ENSRNOG00000005947	ENSMUSG00000064330
PDE7A	Q13946	phosphodiesterase 7A	ENSG00000205268	ENSMMUG00000011262	ENSRNOG00000013048	ENSMUSG00000069094
PDE7B	Q9NP56	phosphodiesterase 7B	ENSG00000171408	ENSMMUG00000003708	ENSRNOG00000013436	ENSMUSG00000019990
PDE8A	O60658	phosphodiesterase 8A	ENSG00000073417	ENSMMUG00000019838	ENSRNOG00000019005	ENSMUSG00000025584
PDE8B	O95263	phosphodiesterase 8B	ENSG00000113231	ENSMMUG00000021358	ENSRNOG00000010280	ENSMUSG00000021684
PDE9A	O76083	phosphodiesterase 9A	ENSG00000160191	ENSMMUG00000022107	ENSRNOG00000001174	ENSMUSG00000041119
PDGFRA	P16234	platelet-derived growth factor receptor, alpha polypeptide	ENSG00000134853	ENSMMUG00000015707	ENSRNOG00000002244	ENSMUSG00000029231
PDGFRB	P09619	platelet-derived growth factor receptor, beta polypeptide	ENSG00000113721	ENSMMUG00000011557	ENSRNOG00000018461	ENSMUSG00000024620
PGA5	P0DJD9	pepsinogen 5, group I (pepsinogen A)	ENSG00000256713	ENSMMUG00000029681	-	-
				ENSMMUG00000011493		
				ENSMMUG00000013985		
PGF	P49763	placental growth factor	ENSG00000119630	ENSMMUG00000002909	ENSRNOG00000005650	ENSMUSG00000004791
PGR	P06401	progesterone receptor	ENSG00000082175	ENSMMUG00000011846	ENSRNOG00000006831	ENSMUSG00000031870
PLAT	P00750	plasminogen activator, tissue	ENSG00000104368	ENSMMUG00000014432	ENSRNOG00000019018	ENSMUSG00000031538
PLG	P00747	plasminogen	ENSG00000122194	ENSMMUG00000016945	ENSRNOG00000017223	ENSMUSG00000059481
PNLIP	P16233	pancreatic lipase	ENSG00000175535	ENSMMUG00000000148	ENSRNOG00000017725	ENSMUSG00000046008
POLA1	P09884	polymerase (DNA directed), alpha 1, catalytic subunit	ENSG00000101868	ENSMMUG00000018433	ENSRNOG00000013322	ENSMUSG00000006678
POLD1	P28340	polymerase (DNA directed), delta 1, catalytic subunit	ENSG00000062822	ENSMMUG00000021525	ENSRNOG00000019681	ENSMUSG00000038644
POLE	Q07864	polymerase (DNA directed), epsilon, catalytic subunit	ENSG00000177084	ENSMMUG00000031453	ENSRNOG00000037449	ENSMUSG00000007080
				ENSMMUG00000015463		
PPARA	Q07869	peroxisome proliferator-activated receptor alpha	ENSG00000186951	ENSMMUG00000016258	ENSRNOG00000021463	ENSMUSG00000022383
PPARG	P37231	peroxisome proliferator-activated receptor gamma	ENSG00000132170	ENSMMUG00000007191	ENSRNOG00000008839	ENSMUSG00000000440
PPAT	Q06203	phosphoribosyl pyrophosphate amidotransferase	ENSG00000128059	ENSMMUG00000016362	ENSRNOG00000002128	ENSMUSG00000029246
PPIA	P62937	peptidylprolyl isomerase A (cyclophilin A)	ENSG00000196262	ENSMMUG00000030052	ENSRNOG00000007183	ENSMUSG00000071866
					ENSRNOG00000022492	ENSMUSG00000062933
					ENSRNOG00000006423	ENSMUSG00000055795
PSMB5	P28074	proteasome (prosome, macropain) subunit, beta type, 5	ENSG00000100804	ENSMMUG00000006007	ENSRNOG00000013386	ENSMUSG00000022193

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
PTGER1	P34995	prostaglandin E receptor 1 (subtype EP1), 42kDa	ENSG00000160951	-	ENSRNOG00000004094	ENSMUSG00000019464
PTGER2	P43116	prostaglandin E receptor 2 (subtype EP2), 53kDa	ENSG00000125384	ENSMMUG00000000023	ENSRNOG00000050968	ENSMUSG00000037759
PTGER3	P43115	prostaglandin E receptor 3 (subtype EP3)	ENSG00000050628	ENSMMUG00000017309	ENSRNOG00000010325	ENSMUSG00000040016
PTGER4	P35408	prostaglandin E receptor 4 (subtype EP4)	ENSG00000171522	ENSMMUG00000016167	ENSRNOG00000013240	ENSMUSG00000039942
PTGFR	P43088	prostaglandin F receptor (FP)	ENSG00000122420	ENSMMUG00000010594	ENSRNOG00000046468	ENSMUSG00000028036
PTGIR	P43119	prostaglandin I2 (prostacyclin) receptor (IP)	ENSG00000160013	ENSMMUG00000017922	ENSRNOG00000016756	ENSMUSG00000043017
PTGS1	P23219	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)	ENSG00000095303	ENSMMUG00000012886	ENSRNOG00000007415	ENSMUSG00000047250
PTGS2	P35354	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	ENSG00000073756	ENSMMUG00000016857	ENSRNOG00000002525	ENSMUSG00000032487
PTH1R	Q03431	parathyroid hormone 1 receptor	ENSG00000160801	ENSMMUG00000014742	ENSRNOG00000020948	ENSMUSG00000032492
PTK6	Q13882	protein tyrosine kinase 6	ENSG00000101213	ENSMMUG00000006065	ENSRNOG00000012987	ENSMUSG00000038751
RAF1	P04049	v-raf-1 murine leukemia viral oncogene homolog 1	ENSG00000132155	ENSMMUG00000007450 ENSMMUG00000009202	ENSRNOG00000010153	ENSMUSG00000000441
RARA	P10276	retinoic acid receptor, alpha	ENSG00000131759	ENSMMUG00000012486	ENSRNOG00000009972	ENSMUSG00000037992
RARB	P10826	retinoic acid receptor, beta	ENSG00000077092	ENSMMUG00000007367	ENSRNOG00000024061	ENSMUSG00000017491
RARG	P13631	retinoic acid receptor, gamma	ENSG00000172819	ENSMMUG00000022017	ENSRNOG00000012499	ENSMUSG00000001288
REN	P00797	renin	ENSG00000143839	ENSMMUG00000008464	ENSRNOG00000002937	ENSMUSG00000070645
RET	P07949	ret proto-oncogene	ENSG00000165731	ENSMMUG00000015400	ENSRNOG00000014751	ENSMUSG00000030110
RRM1	P23921	ribonucleotide reductase M1	ENSG00000167325	ENSMMUG00000003962	ENSRNOG00000045752	ENSMUSG00000030978
RXRA	P19793	retinoid X receptor, alpha	ENSG00000186350 ENSG00000269571	ENSMMUG00000008183	ENSRNOG00000009446	ENSMUSG00000015846
RXRB	P28702	retinoid X receptor, beta	ENSG00000227322 ENSG00000235712 ENSG00000204231 ENSG00000228333 ENSG00000206289 ENSG00000231321	ENSMMUG00000003858	ENSRNOG00000000464	ENSMUSG00000039656
RXRG	P48443	retinoid X receptor, gamma	ENSG00000143171	ENSMMUG00000002892	ENSRNOG00000004537	ENSMUSG00000015843
RYR1	P21817	ryanodine receptor 1 (skeletal)	ENSG00000196218	ENSMMUG00000028770 ENSMMUG00000002897 ENSMMUG00000022576 ENSMMUG00000028768	ENSRNOG00000020557	ENSMUSG00000030592
RYR3	Q15413	ryanodine receptor 3	ENSG00000198838	ENSMMUG00000029784 ENSMMUG00000005270	ENSRNOG00000006645	ENSMUSG00000057378

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
S1PR1	P21453	sphingosine-1-phosphate receptor 1	ENSG00000170989	ENSMMUG00000017778	ENSRNOG00000013683	ENSMUSG00000045092
S1PR2	O95136	sphingosine-1-phosphate receptor 2	ENSG00000267534	ENSMMUG00000006586	ENSRNOG00000020653	ENSMUSG00000043895
S1PR3	Q99500	sphingosine-1-phosphate receptor 3	ENSG00000213694	ENSMMUG00000020356	ENSRNOG00000014524	ENSMUSG00000067586
S1PR4	O95977	sphingosine-1-phosphate receptor 4	ENSG00000125910	ENSMMUG00000017035	ENSRNOG00000005370	ENSMUSG00000044199
S1PR5	Q9H228	sphingosine-1-phosphate receptor 5	ENSG00000180739	ENSMMUG00000008686	ENSRNOG00000020901	ENSMUSG00000045087
SCN10A	Q9Y5Y9	sodium channel, voltage-gated, type X, alpha subunit	ENSG00000185313	ENSMMUG00000015827	ENSRNOG00000032473	ENSMUSG00000034533
SCN11A	Q9UI33	sodium channel, voltage-gated, type XI, alpha subunit	ENSG00000168356	ENSMMUG00000019962	ENSRNOG00000032884	ENSMUSG00000034115
SCN1A	P35498	sodium channel, voltage-gated, type I, alpha subunit	ENSG00000144285	ENSMMUG00000008407	ENSRNOG00000005989	ENSMUSG00000064329
SCN2A	Q99250	sodium channel, voltage-gated, type II, alpha subunit	ENSG00000136531	ENSMMUG00000002712	ENSRNOG00000005018	ENSMUSG00000075318
SCN3A	Q9NY46	sodium channel, voltage-gated, type III, alpha subunit	ENSG00000153253	ENSMMUG00000018840	ENSRNOG00000005007	ENSMUSG00000057182
SCN4A	P35499	sodium channel, voltage-gated, type IV, alpha subunit	ENSG00000007314	ENSMMUG00000000580	ENSRNOG00000012134	ENSMUSG00000001027
SCN5A	Q14524	sodium channel, voltage-gated, type V, alpha subunit	ENSG00000183873	ENSMMUG00000015825	ENSRNOG00000015049	ENSMUSG00000032511
SCN7A	Q01118	sodium channel, voltage-gated, type VII, alpha subunit	ENSG00000136546	ENSMMUG00000002297	ENSRNOG00000029342	ENSMUSG00000034810
SCN8A	Q9UQD0	sodium channel, voltage gated, type VIII, alpha subunit	ENSG00000196876	ENSMMUG00000000415	ENSRNOG00000005309	ENSMUSG00000023033
SCN9A	Q15858	sodium channel, voltage-gated, type IX, alpha subunit	ENSG00000169432	ENSMMUG00000018510	ENSRNOG00000006639	ENSMUSG00000075316
SCNN1A	P37088	sodium channel, non-voltage-gated 1 alpha subunit	ENSG00000111319	ENSMMUG00000020212	ENSRNOG00000019368	ENSMUSG00000030340
SCTR	P47872	secretin receptor	ENSG00000080293	ENSMMUG00000011128	ENSRNOG00000049766	ENSMUSG00000026387
SERPINC1	P01008	serpin peptidase inhibitor, clade C (antithrombin), member 1	ENSG00000117601	ENSMMUG00000000629	ENSRNOG00000002783	ENSMUSG00000026715
SIGMAR1	Q99720	sigma non-opioid intracellular receptor 1	ENSG00000147955	ENSMMUG00000020774	ENSRNOG00000014604	ENSMUSG00000036078
SLC12A1	Q13621	solute carrier family 12 (sodium/potassium/chloride transporters), member 1	ENSG00000074803	ENSMMUG00000016896	ENSRNOG00000005367	ENSMUSG00000027202
SLC12A3	P55017	solute carrier family 12 (sodium/chloride transporters), member 3	ENSG00000070915	ENSMMUG00000004812	ENSRNOG00000018607	ENSMUSG00000031766

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
SLC18A2	Q05940	solute carrier family 18 (vesicular monoamine), member 2	ENSG00000165646	ENSMMUG00000014817	ENSRNOG00000008890	ENSMUSG00000025094
SLC22A11	Q9NSA0	solute carrier family 22 (organic anion/urate transporter), member 11	ENSG00000168065	ENSMMUG00000015372	-	-
SLC22A12	Q96S37	solute carrier family 22 (organic anion/urate transporter), member 12	ENSG00000197891	ENSMMUG00000022594	ENSRNOG00000021108	ENSMUSG00000061742
SLC22A6	Q4U2R8	solute carrier family 22 (organic anion transporter), member 6	ENSG00000197901	ENSMMUG00000016227	ENSRNOG00000018215	ENSMUSG00000024650
SLC22A8	Q8TCC7	solute carrier family 22 (organic anion transporter), member 8	ENSG00000149452	ENSMMUG00000016225	ENSRNOG00000018086	ENSMUSG00000063796
SLC29A1	Q99808	solute carrier family 29 (nucleoside transporters), member 1	ENSG00000112759	ENSMMUG00000022685	ENSRNOG00000019752	ENSMUSG00000023942
SLC5A2	P31639	solute carrier family 5 (sodium/glucose cotransporter), member 2	ENSG00000140675	ENSMMUG00000000897	ENSRNOG00000020039	ENSMUSG00000030781
SLC6A1	P30531	solute carrier family 6 (neurotransmitter transporter, GABA), member 1	ENSG00000157103	ENSMMUG00000005917	ENSRNOG00000006527	ENSMUSG00000030310
SLC6A2	P23975	solute carrier family 6 (neurotransmitter transporter, noradrenalin), member 2	ENSG00000103546	ENSMMUG00000020238	ENSRNOG00000016311	ENSMUSG00000055368
SLC6A3	Q01959	solute carrier family 6 (neurotransmitter transporter, dopamine), member 3	ENSG00000142319	ENSMMUG00000005198	ENSRNOG00000017302	ENSMUSG00000021609
SLC6A4	P31645	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4	ENSG00000108576	ENSMMUG00000003515	ENSRNOG00000003476	ENSMUSG00000020838
SMO	Q99835	smoothened, frizzled family receptor	ENSG00000128602	ENSMMUG00000012110	ENSRNOG00000008332	ENSMUSG00000001761
SNAP25	P60880	synaptosomal-associated protein, 25kDa	ENSG00000132639	ENSMMUG00000011064	ENSRNOG00000006037	ENSMUSG00000027273
SRC	P12931	v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)	ENSG00000197122	ENSMMUG00000020131	ENSRNOG00000009495	ENSMUSG00000027646
SRD5A1	P18405	steroid-5-alpha-reductase, alpha polypeptide 1 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 1)	ENSG00000145545	ENSMMUG00000013186	ENSRNOG00000017601	ENSMUSG00000021594
SRD5A3	Q9H8P0	steroid 5 alpha-reductase 3	ENSG00000128039	ENSMMUG00000010113	ENSRNOG00000002216	ENSMUSG00000029233

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
SRMS	Q9H3Y6	src-related kinase lacking C-terminal regulatory tyrosine and N-terminal myristylation sites	ENSG00000125508	ENSMMUG00000006067	ENSRNOG00000013003	ENSMUSG00000027579
SSTR1	P30872	somatostatin receptor 1	ENSG00000139874	ENSMMUG00000018563	ENSRNOG00000048145	ENSMUSG00000035431
SSTR2	P30874	somatostatin receptor 2	ENSG00000180616	ENSMMUG00000020947	ENSRNOG00000002793	ENSMUSG00000047904
SSTR3	P32745	somatostatin receptor 3	ENSG00000183473	ENSMMUG00000018853	ENSRNOG00000007332	ENSMUSG00000044933
SSTR4	P31391	somatostatin receptor 4	ENSG00000132671	ENSMMUG00000016815	ENSRNOG00000004641	ENSMUSG00000037014
SSTR5	P35346	somatostatin receptor 5	ENSG00000162009	-	ENSRNOG00000018834	ENSMUSG00000050824
SV2A	Q7L0J3	synaptic vesicle glycoprotein 2A	ENSG00000159164 ENSG00000265307	ENSMMUG00000007257	ENSRNOG00000021182	ENSMUSG00000038486
TACR1	P25103	tachykinin receptor 1	ENSG00000115353	ENSMMUG00000001846	ENSRNOG00000005853	ENSMUSG00000030043
TEK	Q02763	TEK tyrosine kinase, endothelial	ENSG00000120156	ENSMMUG00000006080	ENSRNOG00000008587	ENSMUSG00000006386
TH	P07101	tyrosine hydroxylase	ENSG00000180176	ENSMMUG00000003912	ENSRNOG00000020410	ENSMUSG00000000214
THRA	P10827	thyroid hormone receptor, alpha	ENSG00000126351	ENSMMUG00000002931	ENSRNOG00000009066	ENSMUSG00000058756
THRB	P10828	thyroid hormone receptor, beta	ENSG00000151090	ENSMMUG00000000067	ENSRNOG00000006649	ENSMUSG00000021779
TLR7	Q9NYK1	toll-like receptor 7	ENSG00000196664	ENSMMUG00000014254	ENSRNOG00000004249	ENSMUSG00000044583
TLR9	Q9NR96	toll-like receptor 9	ENSG00000239732	ENSMMUG00000012263	ENSRNOG00000048161	ENSMUSG00000045322
TNF	P01375	tumor necrosis factor	ENSG00000206439 ENSG00000204490 ENSG00000230108 ENSG00000223952 ENSG00000228321 ENSG00000228849 ENSG00000232810	ENSMMUG00000008845	ENSRNOG00000000837	ENSMUSG00000024401
TNFRSF8	P28908	tumor necrosis factor receptor superfamily, member 8	ENSG00000120949	ENSMMUG00000008340	ENSRNOG00000016782	ENSMUSG00000028602
TNFSF11	O14788	tumor necrosis factor (ligand) superfamily, member 11	ENSG00000120659	ENSMMUG00000017876	ENSRNOG00000049547 ENSRNOG00000009559	ENSMUSG00000022015
TNFSF13B	Q9Y275	tumor necrosis factor (ligand) superfamily, member 13b	ENSG00000102524	ENSMMUG00000008040	-	ENSMUSG00000031497
TOP1	P11387	topoisomerase (DNA) I	ENSG00000198900	ENSMMUG00000011112	-	ENSMUSG00000070544
TOP1MT	Q969P6	topoisomerase (DNA) I, mitochondrial	ENSG00000184428	ENSMMUG00000022933	ENSRNOG00000007500	ENSMUSG00000000934
TOP2A	P11388	topoisomerase (DNA) II alpha 170kDa	ENSG00000131747	ENSMMUG00000008941	ENSRNOG00000032404	ENSMUSG00000020914
TOP2B	Q02880	topoisomerase (DNA) II beta 180kDa	ENSG00000077097	ENSMMUG00000007371	ENSRNOG00000048795	ENSMUSG00000017485
TPO	P07202	thyroid peroxidase	ENSG00000115705	ENSMMUG00000009662	ENSRNOG00000004646	ENSMUSG00000020673
TRHR	P34981	thyrotropin-releasing hormone receptor	ENSG00000174417	ENSMMUG00000008734	ENSRNOG00000005048	ENSMUSG00000038760

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
TRPA1	O75762	transient receptor potential cation channel, subfamily A, member 1	ENSG00000104321	ENSMMUG00000010748	ENSRNOG00000007354	ENSMUSG00000032769
TRPM8	Q7Z2W7	transient receptor potential cation channel, subfamily M, member 8	ENSG00000144481	ENSMMUG00000021224	ENSRNOG00000019035	ENSMUSG00000036251
TRPV1	Q8NER1	transient receptor potential cation channel, subfamily V, member 1	ENSG00000196689	ENSMMUG00000000711	ENSRNOG00000019486	ENSMUSG00000005952
TSHR	P16473	thyroid stimulating hormone receptor	ENSG00000165409	ENSMMUG00000019713	ENSRNOG00000003972	ENSMUSG00000020963
TUBB	P07437	tubulin, beta class I	ENSG00000224156 ENSG00000232575 ENSG00000235067 ENSG00000196230 ENSG00000232421 ENSG00000183311 ENSG00000227739 ENSG00000229684	ENSMMUG00000023684	ENSRNOG00000000821	ENSMUSG00000001525
TUBB1	Q9H4B7	tubulin, beta 1 class VI	ENSG00000101162	ENSMMUG00000013187	ENSRNOG00000046151	ENSMUSG00000016255
TUBB2A	Q13885	tubulin, beta 2A class IIa	ENSG00000137267	ENSMMUG00000030028	ENSRNOG00000017558	ENSMUSG00000058672
TUBB2B	Q9BVA1	tubulin, beta 2B class IIb	ENSG00000137285	ENSMMUG00000023785	ENSRNOG00000017445	ENSMUSG00000045136
TUBB3	Q13509	tubulin, beta 3 class III	ENSG00000258947	ENSMMUG00000029059	ENSRNOG00000017209	ENSMUSG00000062380
TUBB4A	P04350	tubulin, beta 4A class IVa	ENSG00000104833	-	ENSRNOG00000047505	ENSMUSG00000062591
TUBB4B	P68371	tubulin, beta 4B class IVb	ENSG00000188229	ENSMMUG00000012164	ENSRNOG00000010170	ENSMUSG00000036752
TUBB6	Q9BUF5	tubulin, beta 6 class V	ENSG00000176014	ENSMMUG00000032572	ENSRNOG00000018371	ENSMUSG00000001473
TUBB8	Q3ZCM7	tubulin, beta 8 class VIII	ENSG00000261456	-	-	-
TXNRD1	Q16881	thioredoxin reductase 1	ENSG00000198431	ENSMMUG00000012091	ENSRNOG00000009088	ENSMUSG00000020250
TYK2	P29597	tyrosine kinase 2	ENSG00000105397	ENSMMUG00000003110	ENSRNOG00000032948	ENSMUSG00000032175
TYMS	P04818	thymidylate synthetase	ENSG00000176890	ENSMMUG00000004046	ENSRNOG00000037225	ENSMUSG00000025747
TYR	P14679	tyrosinase	ENSG00000077498	ENSMMUG00000011115	ENSRNOG00000016421	ENSMUSG00000004651
UGCG	Q16739	UDP-glucose ceramide glucosyltransferase	ENSG00000148154	ENSMMUG00000021526	ENSRNOG00000015644	ENSMUSG00000028381
VAMP2	P63027	vesicle-associated membrane protein 2 (synaptobrevin 2)	ENSG00000220205	ENSMMUG00000021801	ENSRNOG00000006989	ENSMUSG00000020894
VDR	P11473	vitamin D (1,25-dihydroxyvitamin D3) receptor	ENSG00000111424	ENSMMUG00000006727	ENSRNOG00000008574	ENSMUSG00000022479
VEGFA	P15692	vascular endothelial growth factor A	ENSG00000112715	ENSMMUG00000004577	ENSRNOG00000019598	ENSMUSG00000023951
VKORC1	Q9BQB6	vitamin K epoxide reductase complex, subunit 1	ENSG00000167397	ENSMMUG00000016031	ENSRNOG00000050828	ENSMUSG00000030804 ENSMUSG00000096145
XDH	P47989	xanthine dehydrogenase	ENSG00000158125	ENSMMUG00000000561	ENSRNOG00000007081	ENSMUSG00000024066

Continued on next page

Table 3 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
YES1	P07947	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	ENSG00000176105	ENSMMUG00000004054	ENSRNOG00000037227	ENSMUSG00000014932

Table 4: List of human metabolic enzymes for all FDA approved drugs. Human orthologues of pharmacological relevant genes of *Macaca mulata*, *Rat norvegicus*, and *Mus musculus* were extracted from ENSEMBL Compara version 72 [144; 145]

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ADH1A	P07327	alcohol dehydrogenase 1A (class I), alpha polypeptide	ENSG00000187758	-	ENSRNOG00000012464	ENSMUSG00000074207
ADH1B	P00325	alcohol dehydrogenase 1B (class I), beta polypeptide	ENSG00000196616	ENSMMUG00000018738	ENSRNOG00000012464	ENSMUSG00000074207
ADH4	P08319	alcohol dehydrogenase 4 (class II), pi polypeptide	ENSG00000198099	ENSMMUG00000011059	ENSRNOG00000046357	ENSMUSG00000037797
ADH5	P11766	alcohol dehydrogenase 5 (class III), chi polypeptide	ENSG00000197894	ENSMMUG00000003552	ENSRNOG00000033854	ENSMUSG00000028138
ADH6	P28332	alcohol dehydrogenase 6 (class V)	ENSG00000172955	ENSMMUG00000014993 ENSMMUG00000015966	ENSRNOG00000012436	-
ADH7	P40394	alcohol dehydrogenase 7 (class IV), mu or sigma polypeptide	ENSG00000196344	ENSMMUG00000011253	ENSRNOG00000032959	ENSMUSG00000055301
ADHFE1	Q8IWW8	alcohol dehydrogenase, iron containing, 1	ENSG00000147576	ENSMMUG00000000887	ENSRNOG00000007069	ENSMUSG00000025911
AHR	P35869	aryl hydrocarbon receptor	ENSG00000106546	ENSMMUG00000012203	ENSRNOG00000004342	ENSMUSG00000019256
AKR1A1	P14550	aldo-keto reductase family 1, member A1 (aldehyde reductase)	ENSG00000117448	ENSMMUG00000020085	ENSRNOG00000016727	ENSMUSG00000028692
AKR1B1	P15121	aldo-keto reductase family 1, member B1 (aldose reductase)	ENSG00000085662	ENSMMUG00000000917	ENSRNOG00000023285	ENSMUSG00000001642
AKR1B10	O60218	aldo-keto reductase family 1, member B10 (aldose reductase)	ENSG00000198074	ENSMMUG00000012324 ENSMMUG00000006587 ENSMMUG00000013098 ENSMMUG00000007557	ENSRNOG00000009513 ENSRNOG00000033953 ENSRNOG00000026887 ENSRNOG00000049765 ENSRNOG00000042697 ENSRNOG00000009875 ENSRNOG00000027433 ENSRNOG00000009734 ENSRNOG00000028754 ENSRNOG00000046247 ENSRNOG00000033469 ENSRNOG00000029005	ENSMUSG00000029762 ENSMUSG00000061758 ENSMUSG00000052131
AKR1C1	Q04828	aldo-keto reductase family 1, member C1 (dihydrodiol dehydrogenase 1; 20-alpha (3-alpha)-hydroxysteroid dehydrogenase)	ENSG00000187134	ENSMMUG00000007932 ENSMMUG00000023282	-	-

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
AKR1C2	P52895	aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile acid binding protein; 3-alpha hydroxysteroid dehydrogenase, type III)	ENSG00000151632	ENSMMUG00000007932	-	-
AKR1C3	P42330	aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II)	ENSG00000196139	ENSMMUG00000023282 ENSMMUG00000023280	-	-
AKR1C4	P17516	aldo-keto reductase family 1, member C4 (chlordecone reductase; 3-alpha hydroxysteroid dehydrogenase, type I; dihydrodiol dehydrogenase 4)	ENSG00000198610	ENSMMUG00000000958	-	-
AKR1D1	P51857	aldo-keto reductase family 1, member D1 (delta 4-3-ketosteroid-5-beta-reductase)	ENSG00000122787	ENSMMUG00000005422	ENSRNOG00000013004	ENSMUSG00000038641
AKR7A2	O43488	aldo-keto reductase family 7, member A2 (aflatoxin aldehyde reductase)	ENSG00000053371	-	-	ENSMUSG00000028743
AKR7A3	O95154	aldo-keto reductase family 7, member A3 (aflatoxin aldehyde reductase)	ENSG00000162482	ENSMMUG00000006891	-	-
ALDH16A1	Q8IZ83	aldehyde dehydrogenase 16 family, member A1	ENSG00000161618	ENSMMUG00000022117	ENSRNOG00000020623	ENSMUSG00000007833
ALDH18A1	P54886	aldehyde dehydrogenase 18 family, member A1	ENSG00000059573	ENSMMUG00000003102	ENSRNOG00000015267	ENSMUSG00000025007
ALDH1A1	P00352	aldehyde dehydrogenase 1 family, member A1	ENSG00000165092	ENSMMUG00000013303	ENSRNOG00000047305 ENSRNOG00000017878 ENSRNOG00000017619	ENSMUSG00000053279 ENSMUSG00000024747
ALDH1A2	O94788	aldehyde dehydrogenase 1 family, member A2	ENSG00000128918	ENSMMUG00000019504	ENSRNOG00000016042	ENSMUSG00000013584
ALDH1A3	P47895	aldehyde dehydrogenase 1 family, member A3	ENSG00000184254	ENSMMUG00000006659	ENSRNOG00000013028	ENSMUSG00000015134
ALDH1B1	P30837	aldehyde dehydrogenase 1 family, member B1	ENSG00000137124	ENSMMUG00000013057	ENSRNOG00000011497	ENSMUSG00000035561
ALDH1L1	O75891	aldehyde dehydrogenase 1 family, member L1	ENSG00000144908	ENSMMUG00000012488	ENSRNOG00000047023	ENSMUSG00000030088
ALDH1L2	Q3SY69	aldehyde dehydrogenase 1 family, member L2	ENSG00000136010	ENSMMUG00000005797	ENSRNOG00000008586	ENSMUSG00000020256

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ALDH2	P05091	aldehyde dehydrogenase 2 family (mitochondrial)	ENSG00000111275	ENSMMUG00000008336	ENSRNOG00000001344	ENSMUSG00000029455
ALDH3A1	P30838	aldehyde dehydrogenase 3 family, member A1	ENSG00000108602	ENSMMUG00000001328	ENSRNOG00000002331	ENSMUSG00000019102
ALDH3A2	P51648	aldehyde dehydrogenase 3 family, member A2	ENSG00000072210	ENSMMUG00000014130	ENSRNOG00000002342	ENSMUSG00000010025
ALDH3B2	P48448	aldehyde dehydrogenase 3 family, member B2	ENSG00000132746	ENSMMUG00000018533	ENSRNOG00000017872	ENSMUSG00000075296
						ENSMUSG00000037263
ALDH4A1	P30038	aldehyde dehydrogenase 4 family, member A1	ENSG00000159423	ENSMMUG00000004751	ENSRNOG00000042920	ENSMUSG00000028737
ALDH5A1	P51649	aldehyde dehydrogenase 5 family, member A1	ENSG00000112294	ENSMMUG00000019368	ENSRNOG00000023538	ENSMUSG00000035936
ALDH6A1	Q02252	aldehyde dehydrogenase 6 family, member A1	ENSG00000119711	ENSMMUG00000012684	ENSRNOG00000011419	ENSMUSG00000021238
ALDH7A1	P49419	aldehyde dehydrogenase 7 family, member A1	ENSG00000164904	ENSMMUG00000002489	ENSRNOG00000014645	ENSMUSG00000053644
ALDH8A1	Q9H2A2	aldehyde dehydrogenase 8 family, member A1	ENSG00000118514	ENSMMUG00000029428	ENSRNOG00000014907	ENSMUSG00000037542
ALDH9A1	P49189	aldehyde dehydrogenase 9 family, member A1	ENSG00000143149	ENSMMUG00000015946	ENSRNOG00000004027	ENSMUSG00000026687
AOC1	P19801	amiloride binding protein 1 (amine oxidase (copper-containing))	ENSG00000002726	ENSMMUG00000020435	ENSRNOG00000008714	ENSMUSG00000029811
					ENSRNOG00000008575	ENSMUSG00000029813
					ENSRNOG00000032432	ENSMUSG00000068536
					ENSRNOG00000023612	ENSMUSG00000039215
AOC3	Q16853	amine oxidase, copper containing 3 (vascular adhesion protein 1)	ENSG00000131471	ENSMMUG00000001871	ENSRNOG00000036806	ENSMUSG00000019326
AOX1	Q06278	aldehyde oxidase 1	ENSG00000138356	ENSMMUG00000021931	ENSRNOG00000015354	ENSMUSG00000063558
ARNT	P27540	aryl hydrocarbon receptor nuclear translocator	ENSG00000143437	ENSMMUG00000000541	ENSRNOG00000031174	ENSMUSG00000015522
ARSA	P15289	arylsulfatase A	ENSG00000100299	ENSMMUG00000016274	ENSRNOG00000012953	ENSMUSG00000022620
AS3MT	Q9HBK9	arsenic (+3 oxidation state) methyltransferase	ENSG00000214435	ENSMMUG00000007662	ENSRNOG00000020081	ENSMUSG00000003559
ATP7A	Q04656	ATPase, Cu++ transporting, alpha polypeptide	ENSG00000165240	ENSMMUG00000012871	ENSRNOG00000002515	ENSMUSG00000033792
ATP7B	P35670	ATPase, Cu++ transporting, beta polypeptide	ENSG00000123191	ENSMMUG00000016520	ENSRNOG00000012878	ENSMUSG00000006567
BAAT	Q14032	bile acid CoA: amino acid N-acyltransferase (glycine N-choloyltransferase)	ENSG00000136881	ENSMMUG00000023793	ENSRNOG00000007395	ENSMUSG00000039653
CAT	P04040	catalase	ENSG00000121691	ENSMMUG00000002429	ENSRNOG00000008364	ENSMUSG00000027187
CBR1	P16152	carbonyl reductase 1	ENSG00000159228	ENSMMUG00000020565	-	ENSMUSG00000051483

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CBR3	O75828	carbonyl reductase 3	ENSG00000159231	ENSMMUG00000010338	ENSRNOG00000001701	ENSMUSG00000022947
CDA	P32320	cytidine deaminase	ENSG00000158825	ENSMMUG00000001498	ENSRNOG00000015677	ENSMUSG00000028755
					ENSRNOG00000049700	
CES1	P23141	carboxylesterase 1	ENSG00000198848	ENSMMUG00000007651	ENSRNOG00000015438	ENSMUSG00000057074
					ENSRNOG00000015317	ENSMUSG00000056973
					ENSRNOG00000039725	ENSMUSG00000061959
					ENSRNOG00000018292	ENSMUSG00000078964
					ENSRNOG00000015519	ENSMUSG00000057400
						ENSMUSG00000071047
						ENSMUSG00000031725
CES2	O00748	carboxylesterase 2	ENSG00000172831	ENSMMUG00000022825	ENSRNOG00000012034	ENSMUSG00000031886
				ENSMMUG00000017445	ENSRNOG00000048823	ENSMUSG00000062826
					ENSRNOG00000013808	ENSMUSG00000091813
					ENSRNOG00000011635	ENSMUSG00000050097
					ENSRNOG00000048039	ENSMUSG00000031877
					ENSRNOG00000047661	ENSMUSG00000061825
					ENSRNOG00000036571	ENSMUSG00000031884
					ENSRNOG00000029744	ENSMUSG00000055730
					ENSRNOG00000039079	
					ENSRNOG00000050702	
					ENSRNOG00000011330	
CES3	Q6UWW8	carboxylesterase 3	ENSG00000172828	-	-	ENSMUSG00000062181
						ENSMUSG00000069922
CFTR	P13569	cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7)	ENSG00000001626	ENSMMUG00000011269	ENSRNOG00000008284	ENSMUSG00000041301
CHST1	O43916	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	ENSG00000175264	ENSMMUG00000000874	ENSRNOG00000007989	ENSMUSG00000027221
CHST10	O43529	carbohydrate sulfotransferase 10	ENSG00000115526	ENSMMUG00000021645	ENSRNOG00000012815	ENSMUSG00000026080
CHST11	Q9NPF2	carbohydrate (chondroitin 4) sulfotransferase 11	ENSG00000171310	ENSMMUG00000017127	ENSRNOG00000008885	ENSMUSG00000034612
CHST12	Q9NRB3	carbohydrate (chondroitin 4) sulfotransferase 12	ENSG00000136213	ENSMMUG00000006475	ENSRNOG00000001252	ENSMUSG00000036599
CHST13	Q8NET6	carbohydrate (chondroitin 4) sulfotransferase 13	ENSG00000180767	ENSMMUG00000004201	ENSRNOG00000025930	ENSMUSG00000056643
CHST2	Q9Y4C5	carbohydrate (N-acetylglucosamine-6-O) sulfotransferase 2	ENSG00000175040	ENSMMUG00000002779	ENSRNOG00000047734	ENSMUSG00000033350
CHST3	Q7LGC8	carbohydrate (chondroitin 6) sulfotransferase 3	ENSG00000122863	ENSMMUG00000014783	ENSRNOG00000000572	ENSMUSG00000057337
CHST4	Q8NCG5	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4	ENSG00000140835	ENSMMUG00000004391	ENSRNOG00000016953	ENSMUSG00000035930

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CHST5	Q9GZS9	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 5	ENSG00000135702	-	ENSRNOG00000021521	ENSMUSG00000031952
CHST6	Q9GZX3	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 6	ENSG00000183196	ENSMMUG00000001563	ENSRNOG00000021521	ENSMUSG00000031952
CHST7	Q9NS84	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 7	ENSG00000147119	ENSMMUG00000019663	ENSRNOG00000004258	ENSMUSG00000037347
CHST8	Q9H2A9	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 8	ENSG00000124302	ENSMMUG00000018024	ENSRNOG00000022919	ENSMUSG00000060402
CHST9	Q7L1S5	carbohydrate (N-acetylgalactosamine 4-0) sulfotransferase 9	ENSG00000154080	ENSMMUG00000011801	ENSRNOG00000015867	ENSMUSG00000047161
COMT	P21964	catechol-O-methyltransferase	ENSG00000093010	ENSMMUG00000017152	ENSRNOG00000001889	ENSMUSG00000000326
CYB5A	P00167	cytochrome b5 type A (microsomal)	ENSG00000166347	ENSMMUG00000013607	ENSRNOG00000014757	ENSMUSG00000024646
CYB5R3	P00387	cytochrome b5 reductase 3	ENSG00000100243	ENSMMUG00000010198	ENSRNOG00000015205 ENSRNOG00000009592	ENSMUSG00000018042
CYP11A1	P05108	cytochrome P450, family 11, subfamily A, polypeptide 1	ENSG00000140459	ENSMMUG00000003239	ENSRNOG00000008074	ENSMUSG00000032323
CYP11B1	P15538	cytochrome P450, family 11, subfamily B, polypeptide 1	ENSG00000160882	ENSMMUG00000013631	ENSRNOG00000046367	ENSMUSG00000022589
CYP11B2	P19099	cytochrome P450, family 11, subfamily B, polypeptide 2	ENSG00000179142	ENSMMUG00000013629 ENSMMUG00000000959	ENSRNOG00000030111 ENSRNOG00000049205 ENSRNOG00000046367 ENSRNOG00000030111 ENSRNOG00000049205	ENSMUSG00000075604 ENSMUSG00000022589 ENSMUSG00000075604
CYP17A1	P05093	cytochrome P450, family 17, subfamily A, polypeptide 1	ENSG00000148795	ENSMMUG00000017601	ENSRNOG00000020035	ENSMUSG00000003555
CYP19A1	P11511	cytochrome P450, family 19, subfamily A, polypeptide 1	ENSG00000137869	ENSMMUG00000002553	ENSRNOG00000000196	ENSMUSG00000032274
CYP1A1	P04798	cytochrome P450, family 1, subfamily A, polypeptide 1	ENSG00000140465	ENSMMUG00000022481	ENSRNOG00000019500	ENSMUSG00000032315
CYP1A2	P05177	cytochrome P450, family 1, subfamily A, polypeptide 2	ENSG00000140505	ENSMMUG00000021832	ENSRNOG00000016173	ENSMUSG00000032310
CYP1B1	Q16678	cytochrome P450, family 1, subfamily B, polypeptide 1	ENSG00000138061	ENSMMUG00000013945	ENSRNOG00000040287 ENSRNOG00000047433	ENSMUSG00000024087
CYP20A1	Q6UW02	cytochrome P450, family 20, subfamily A, polypeptide 1	ENSG00000119004	ENSMMUG00000029750	ENSRNOG00000017631	ENSMUSG00000049439

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CYP21A2	P08686	cytochrome P450, family 21, subfamily A, polypeptide 2	ENSG00000235134 ENSG00000198457	-	-	-
CYP24A1	Q07973	cytochrome P450, family 24, subfamily A, polypeptide 1	ENSG00000019186	ENSMMUG00000014860	ENSRNOG00000013062	ENSMUSG00000038567
CYP26A1	O43174	cytochrome P450, family 26, subfamily A, polypeptide 1	ENSG00000095596	-	ENSRNOG00000016750	ENSMUSG00000024987
CYP26C1	Q6V0L0	cytochrome P450, family 26, subfamily C, polypeptide 1	ENSG00000187553	ENSMMUG00000010881 ENSMMUG00000029373 ENSMMUG00000029167	ENSRNOG00000030698	ENSMUSG00000062432
CYP27A1	Q02318	cytochrome P450, family 27, subfamily A, polypeptide 1	ENSG00000135929	ENSMMUG00000006758	ENSRNOG00000017188	ENSMUSG00000026170
CYP27B1	O15528	cytochrome P450, family 27, subfamily B, polypeptide 1	ENSG00000111012	ENSMMUG00000022144	ENSRNOG00000049170 ENSRNOG00000046214	ENSMUSG00000006724
CYP2A13	Q16696	cytochrome P450, family 2, subfamily A, polypeptide 13	ENSG00000197838	ENSMMUG00000006394 ENSMMUG00000021068	ENSRNOG00000020928 ENSRNOG00000020817 ENSRNOG00000028900	ENSMUSG00000005547 ENSMUSG00000074254 ENSMUSG00000060407 ENSMUSG00000091867
CYP2A6	P11509	cytochrome P450, family 2, subfamily A, polypeptide 6	ENSG00000255974	ENSMMUG00000006394 ENSMMUG00000021068	ENSRNOG00000020928 ENSRNOG00000020817 ENSRNOG00000028900	ENSMUSG00000005547 ENSMUSG00000074254 ENSMUSG00000060407 ENSMUSG00000091867
CYP2A7	P20853	cytochrome P450, family 2, subfamily A, polypeptide 7	ENSG00000198077	ENSMMUG00000006394 ENSMMUG00000021068	ENSRNOG00000020928 ENSRNOG00000020817 ENSRNOG00000028900	ENSMUSG00000005547 ENSMUSG00000074254 ENSMUSG00000060407 ENSMUSG00000091867
CYP2B6	P20813	cytochrome P450, family 2, subfamily B, polypeptide 6	ENSG00000197408	ENSMMUG00000018363	ENSRNOG00000020801 ENSRNOG00000033164 ENSRNOG00000031529 ENSRNOG00000050823 ENSRNOG00000020775 ENSRNOG00000043320 ENSRNOG00000032004 ENSRNOG00000033680	ENSMUSG00000030483 ENSMUSG00000040650 ENSMUSG00000066704 ENSMUSG00000040583 ENSMUSG00000040660
CYP2C18	P33260	cytochrome P450, family 2, subfamily C, polypeptide 18	ENSG00000108242	ENSMMUG00000018330	ENSRNOG00000012896 ENSRNOG00000001466 ENSRNOG00000013241	ENSMUSG00000025002

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CYP2C19	P33261	cytochrome P450, family 2, subfamily C, polypeptide 19	ENSG00000165841	-	ENSRNOG00000047945	ENSMUSG00000003053
					ENSRNOG00000030005	ENSMUSG00000067225
					ENSRNOG00000024016	ENSMUSG00000067231
					ENSRNOG00000049464	ENSMUSG00000067229
					ENSRNOG00000021405	ENSMUSG00000042248
					ENSRNOG00000030661	ENSMUSG00000054827
					ENSRNOG00000046080	ENSMUSG00000025003
						ENSMUSG00000032808
						ENSMUSG00000074882
						ENSMUSG00000025004
						ENSMUSG00000062624
						ENSMUSG00000092008
CYP2C8	P10632	cytochrome P450, family 2, subfamily C, polypeptide 8	ENSG00000138115	ENSMMUG00000000124	ENSRNOG00000047945	ENSMUSG00000003053
					ENSRNOG00000030005	ENSMUSG00000067225
					ENSRNOG00000024016	ENSMUSG00000067231
					ENSRNOG00000049464	ENSMUSG00000067229
					ENSRNOG00000021405	ENSMUSG00000042248
					ENSRNOG00000030661	ENSMUSG00000054827
					ENSRNOG00000046080	ENSMUSG00000025003
						ENSMUSG00000032808
						ENSMUSG00000074882
						ENSMUSG00000025004
						ENSMUSG00000062624
						ENSMUSG00000092008
CYP2C9	P11712	cytochrome P450, family 2, subfamily C, polypeptide 9	ENSG00000138109	ENSMMUG00000015542	ENSRNOG00000047945	ENSMUSG00000003053
				ENSMMUG00000004006	ENSRNOG00000030005	ENSMUSG00000067225
					ENSRNOG00000024016	ENSMUSG00000067231
					ENSRNOG00000049464	ENSMUSG00000067229
					ENSRNOG00000021405	ENSMUSG00000042248
					ENSRNOG00000030661	ENSMUSG00000054827
					ENSRNOG00000046080	ENSMUSG00000025003
						ENSMUSG00000032808
						ENSMUSG00000074882
						ENSMUSG00000025004
						ENSMUSG00000062624
						ENSMUSG00000092008
CYP2D6	P10635	cytochrome P450, family 2, subfamily D, polypeptide 6	ENSG00000100197	ENSMMUG00000017962	ENSRNOG00000032261	ENSMUSG00000061740
					ENSRNOG00000029179	ENSMUSG00000022445
					ENSRNOG00000008988	ENSMUSG00000094559

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
					ENSRNOG00000029128	ENSMUSG00000094806
					ENSRNOG00000030182	ENSMUSG00000068085
						ENSMUSG00000096852
						ENSMUSG00000068086
						ENSMUSG00000058613
						ENSMUSG00000068083
CYP2E1	P05181	cytochrome P450, family 2, subfamily E, polypeptide 1	ENSG00000130649	-	ENSRNOG00000012458	ENSMUSG00000025479
CYP2F1	P24903	cytochrome P450, family 2, subfamily F, polypeptide 1	ENSG00000197446	-	ENSRNOG00000032805	ENSMUSG00000052974
CYP2J2	P51589	cytochrome P450, family 2, subfamily J, polypeptide 2	ENSG00000134716	ENSMMUG00000015972	ENSRNOG00000030466	ENSMUSG00000052520
					ENSRNOG00000031004	ENSMUSG00000052914
					ENSRNOG00000042224	ENSMUSG00000081225
						ENSMUSG00000015224
						ENSMUSG00000081362
						ENSMUSG00000028571
						ENSMUSG00000082932
						ENSMUSG00000066097
CYP2R1	Q6VVX0	cytochrome P450, family 2, subfamily R, polypeptide 1	ENSG00000186104	ENSMMUG00000000761	ENSRNOG00000011367	ENSMUSG00000030670
CYP2S1	Q96SQ9	cytochrome P450, family 2, subfamily S, polypeptide 1	ENSG00000167600	-	ENSRNOG00000020743	ENSMUSG00000040703
CYP39A1	Q9NYL5	cytochrome P450, family 39, subfamily A, polypeptide 1	ENSG00000146233	ENSMMUG00000019247	ENSRNOG00000010519	ENSMUSG00000023963
CYP3A4	P08684	cytochrome P450, family 3, subfamily A, polypeptide 4	ENSG00000160868	-	ENSRNOG00000001379	ENSMUSG00000029727
					ENSRNOG00000032560	ENSMUSG00000038656
					ENSRNOG00000047745	ENSMUSG00000070419
					ENSRNOG00000000978	ENSMUSG00000056035
					ENSRNOG00000046643	ENSMUSG00000054417
					ENSRNOG00000049143	ENSMUSG00000029630
					ENSRNOG00000000969	ENSMUSG00000061292
						ENSMUSG00000075552
						ENSMUSG00000075551
CYP3A43	Q9HB55	cytochrome P450, family 3, subfamily A, polypeptide 43	ENSG00000021461	ENSMMUG00000022490	ENSRNOG00000001379	ENSMUSG00000029727
					ENSRNOG00000032560	ENSMUSG00000038656
					ENSRNOG00000047745	ENSMUSG00000070419
					ENSRNOG00000000978	ENSMUSG00000056035
					ENSRNOG00000046643	ENSMUSG00000054417
					ENSRNOG00000049143	ENSMUSG00000029630
					ENSRNOG00000000969	ENSMUSG00000061292

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
						ENSMUSG00000075552
						ENSMUSG00000075551
CYP3A5	P20815	cytochrome P450, family 3, subfamily A, polypeptide 5	ENSG00000106258	-	ENSRNOG00000001379	ENSMUSG00000029727
					ENSRNOG00000032560	ENSMUSG00000038656
					ENSRNOG00000047745	ENSMUSG00000070419
					ENSRNOG00000000978	ENSMUSG00000056035
					ENSRNOG00000046643	ENSMUSG00000054417
					ENSRNOG00000049143	ENSMUSG00000029630
					ENSRNOG00000000969	ENSMUSG00000061292
						ENSMUSG00000075552
						ENSMUSG00000075551
CYP3A7	P24462	cytochrome P450, family 3, subfamily A, polypeptide 7	ENSG00000160870	ENSMMUG00000022730	ENSRNOG00000001379	ENSMUSG00000029727
					ENSRNOG00000032560	ENSMUSG00000038656
					ENSRNOG00000047745	ENSMUSG00000070419
					ENSRNOG00000000978	ENSMUSG00000056035
					ENSRNOG00000046643	ENSMUSG00000054417
					ENSRNOG00000049143	ENSMUSG00000029630
					ENSRNOG00000000969	ENSMUSG00000061292
						ENSMUSG00000075552
						ENSMUSG00000075551
CYP46A1	Q9Y6A2	cytochrome P450, family 46, subfamily A, polypeptide 1	ENSG00000036530	ENSMMUG00000000100	ENSRNOG00000007147	ENSMUSG00000021259
CYP4A11	Q02928	cytochrome P450, family 4, subfamily A, polypeptide 11	ENSG00000187048	ENSMMUG00000032143	ENSRNOG00000030154	ENSMUSG00000028715
					ENSRNOG00000008842	ENSMUSG00000066071
					ENSRNOG00000009597	ENSMUSG00000078597
						ENSMUSG00000063929
						ENSMUSG00000066072
						ENSMUSG00000028712
CYP4B1	P13584	cytochrome P450, family 4, subfamily B, polypeptide 1	ENSG00000142973	ENSMMUG00000019873	ENSRNOG00000009741	ENSMUSG00000028713
CYP4F11	Q9HBI6	cytochrome P450, family 4, subfamily F, polypeptide 11	ENSG00000171903	ENSMMUG00000006873	-	-
CYP4F12	Q9HCS2	cytochrome P450, family 4, subfamily F, polypeptide 12	ENSG00000186204	ENSMMUG00000010734	-	-
CYP4F2	P78329	cytochrome P450, family 4, subfamily F, polypeptide 2	ENSG00000186115	ENSMMUG00000006872	-	-
CYP4F3	Q08477	cytochrome P450, family 4, subfamily F, polypeptide 3	ENSG00000186529	-	ENSRNOG00000015751	-
CYP4F8	P98187	cytochrome P450, family 4, subfamily F, polypeptide 8	ENSG00000186526	-	-	-

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
CYP4Z1	Q86W10	cytochrome P450, family 4, subfamily Z, polypeptide 1	ENSG00000186160	-	-	-
CYP51A1	Q16850	cytochrome P450, family 51, subfamily A, polypeptide 1	ENSG00000001630	-	ENSRNOG00000007234	ENSMUSG00000001467
CYP7A1	P22680	cytochrome P450, family 7, subfamily A, polypeptide 1	ENSG00000167910	ENSMMUG00000006543	ENSRNOG00000009488	ENSMUSG00000028240
CYP7B1	O75881	cytochrome P450, family 7, subfamily B, polypeptide 1	ENSG00000172817	ENSMMUG00000004411	ENSRNOG00000009730	ENSMUSG00000039519
CYP8B1	Q9UNU6	cytochrome P450, family 8, subfamily B, polypeptide 1	ENSG00000180432	-	ENSRNOG00000019481	ENSMUSG00000050445
DDO	Q99489	D-aspartate oxidase	ENSG00000203797	ENSMMUG00000008185	ENSRNOG00000000582	ENSMUSG00000063428
DHRS1	Q96LJ7	dehydrogenase/reductase (SDR family) member 1	ENSG00000157379	ENSMMUG00000010051	ENSRNOG00000020264	ENSMUSG00000002332
DHRS12	A0PJE2	dehydrogenase/reductase (SDR family) member 12	ENSG00000102796	ENSMMUG00000009420	-	-
DHRS13	Q6UX07	dehydrogenase/reductase (SDR family) member 13	ENSG00000167536	ENSMMUG00000021553	ENSRNOG00000009673	ENSMUSG00000020834
DHRS2	Q13268	dehydrogenase/reductase (SDR family) member 2	ENSG00000100867	ENSMMUG00000021849	ENSRNOG00000018177	ENSMUSG00000022209
DHRS3	O75911	dehydrogenase/reductase (SDR family) member 3	ENSG00000162496	ENSMMUG00000005681	ENSRNOG00000015736	ENSMUSG00000066026
DHRS4	Q9BTZ2	dehydrogenase/reductase (SDR family) member 4	ENSG00000157326	ENSMMUG00000007465	ENSRNOG00000018239	ENSMUSG00000022210
DHRS4L2	Q6PKH6	dehydrogenase/reductase (SDR family) member 4 like 2	ENSG00000187630	ENSMMUG00000007465	ENSRNOG00000018239	ENSMUSG00000022210
DHRS7	Q9Y394	dehydrogenase/reductase (SDR family) member 7	ENSG00000100612	ENSMMUG00000002540	ENSRNOG00000025648	ENSMUSG00000021094
					ENSRNOG00000005589 ENSRNOG00000047172	
DHRS7B	Q6IAN0	dehydrogenase/reductase (SDR family) member 7B	ENSG00000109016	ENSMMUG00000013194	ENSRNOG00000005360	ENSMUSG00000042569
DHRS7C	A6NNS2	dehydrogenase/reductase (SDR family) member 7C	ENSG00000184544	ENSMMUG00000011968	ENSRNOG00000026548	ENSMUSG00000033044
DHRS9	Q9BPW9	dehydrogenase/reductase (SDR family) member 9	ENSG00000073737	ENSMMUG00000013835	ENSRNOG00000006773	ENSMUSG00000027068
DHRSX	Q8N5I4	dehydrogenase/reductase (SDR family) X-linked	ENSG00000169084	ENSMMUG00000011897	ENSRNOG00000028323	ENSMUSG00000063897
DPEP1	P16444	dipeptidase 1 (renal)	ENSG00000015413	ENSMMUG00000014910	ENSRNOG00000015880	ENSMUSG00000019278
DPYD	Q12882	dihydropyrimidine dehydrogenase	ENSG00000188641	ENSMMUG00000020726	ENSRNOG00000049685 ENSRNOG00000017105	ENSMUSG00000033308
EPHX1	P07099	epoxide hydrolase 1, microsomal (xenobiotic)	ENSG00000143819	ENSMMUG00000005782	ENSRNOG00000037928	ENSMUSG00000038776
EPHX2	P34913	epoxide hydrolase 2, cytoplasmic	ENSG00000120915	ENSMMUG00000021515	ENSRNOG00000017286	ENSMUSG00000022040
EPX	P11678	eosinophil peroxidase	ENSG00000121053	ENSMMUG00000011973	ENSRNOG00000008707	ENSMUSG00000052234

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
FMO1	Q01740	flavin containing monooxygenase 1	ENSG00000010932	ENSMMUG00000005069	ENSRNOG00000034191	ENSMUSG00000040181
FMO2	Q99518	flavin containing monooxygenase 2 (non-functional)	ENSG00000094963	ENSMMUG00000019727	ENSRNOG00000003510	ENSMUSG00000040170
FMO3	P31513	flavin containing monooxygenase 3	ENSG00000007933	ENSMMUG00000007952	ENSRNOG00000003620	ENSMUSG00000026691
FMO4	P31512	flavin containing monooxygenase 4	ENSG00000076258	ENSMMUG00000014230	ENSRNOG00000003400	ENSMUSG00000026692
FMO5	P49326	flavin containing monooxygenase 5	ENSG00000131781	ENSMMUG00000000476	ENSRNOG00000018076	ENSMUSG00000028088
FMO6P	O60774	flavin containing monooxygenase 6 pseudogene	ENSG00000117507	ENSMMUG00000007956	ENSRNOG00000049851	ENSMUSG00000095576
GAMT	Q14353	guanidinoacetate N-methyltransferase	ENSG00000130005	ENSMMUG00000003046	ENSRNOG00000024577	ENSMUSG00000020150
GNMT	Q14749	glycine N-methyltransferase	ENSG00000124713	ENSMMUG00000001393	ENSRNOG00000016349 ENSRNOG00000048623	ENSMUSG00000002769
GPX1	P07203	glutathione peroxidase 1	ENSG00000233276	ENSMMUG00000012963	ENSRNOG00000048812	ENSMUSG00000063856
GPX2	P18283	glutathione peroxidase 2 (gastrointestinal)	ENSG00000176153	ENSMMUG00000021987	-	ENSMUSG00000042808
GPX3	P22352	glutathione peroxidase 3 (plasma)	ENSG00000211445	-	-	ENSMUSG00000018339
GPX4	P36969	glutathione peroxidase 4	ENSG00000167468	ENSMMUG00000028946	ENSRNOG00000013604	ENSMUSG00000075706
GPX5	O75715	glutathione peroxidase 5 (epididymal androgen-related protein)	ENSG00000224586	ENSMMUG00000004618	ENSRNOG00000018596	ENSMUSG00000004344
GPX6	P59796	glutathione peroxidase 6 (olfactory)	ENSG00000198704	ENSMMUG00000004617	ENSRNOG00000018473	ENSMUSG00000004341
GPX7	Q96SL4	glutathione peroxidase 7	ENSG00000116157	ENSMMUG00000019676	ENSRNOG00000009751	ENSMUSG00000028597
GSR	P00390	glutathione reductase	ENSG00000104687	-	ENSRNOG00000014915	ENSMUSG00000031584
GSS	P48637	glutathione synthetase	ENSG00000100983	ENSMMUG00000017940	ENSRNOG00000018964	ENSMUSG00000027610
GSTA1	P08263	glutathione S-transferase alpha 1	ENSG00000243955	ENSMMUG00000017137	ENSRNOG00000033402 ENSRNOG00000013484 ENSRNOG00000029711 ENSRNOG00000050464	-
GSTA2	P09210	glutathione S-transferase alpha 2	ENSG00000244067	ENSMMUG00000017137	ENSRNOG00000033402 ENSRNOG00000013484 ENSRNOG00000029711 ENSRNOG00000050464	-
GSTA3	Q16772	glutathione S-transferase alpha 3	ENSG00000174156	-	ENSRNOG00000033402 ENSRNOG00000013484 ENSRNOG00000029711 ENSRNOG00000050464	-
GSTA4	O15217	glutathione S-transferase alpha 4	ENSG00000170899	ENSMMUG00000014899	-	-
GSTA5	Q7RTV2	glutathione S-transferase alpha 5	ENSG00000182793	-	ENSRNOG00000033402	-

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
					ENSRNOG00000013484	
					ENSRNOG00000029711	
					ENSRNOG00000050464	
GSTCD	Q8NEC7	glutathione S-transferase, C-terminal domain containing	ENSG00000138780	ENSMMUG00000018811	ENSRNOG00000011669	ENSMUSG00000028018
GSTK1	Q9Y2Q3	glutathione S-transferase kappa 1	ENSG00000197448	ENSMMUG00000012628	ENSRNOG00000016484	ENSMUSG00000029864
GSTM1	P09488	glutathione S-transferase mu 1	ENSG00000134184	ENSMMUG00000020851	-	ENSMUSG00000091561
						ENSMUSG00000040562
						ENSMUSG00000068762
						ENSMUSG00000004038
						ENSMUSG00000058135
GSTM2	P28161	glutathione S-transferase mu 2 (muscle)	ENSG00000213366	-	ENSRNOG00000018937	ENSMUSG00000004035
GSTM3	P21266	glutathione S-transferase mu 3 (brain)	ENSG00000134202	ENSMMUG00000008029	ENSRNOG00000049743	ENSMUSG00000004032
					ENSRNOG00000047034	
GSTM4	Q03013	glutathione S-transferase mu 4	ENSG00000168765	ENSMMUG00000019630	ENSRNOG00000019221	ENSMUSG00000027890
						ENSMUSG00000095238
GSTM5	P46439	glutathione S-transferase mu 5	ENSG00000134201	-	-	ENSMUSG00000091561
						ENSMUSG00000040562
						ENSMUSG00000068762
						ENSMUSG00000004038
						ENSMUSG00000058135
GSTO1	P78417	glutathione S-transferase omega 1	ENSG00000148834	ENSMMUG00000032415	ENSRNOG00000028746	ENSMUSG00000025068
				ENSMMUG00000015334		
GSTO2	Q9H4Y5	glutathione S-transferase omega 2	ENSG00000065621	ENSMMUG00000015337	ENSRNOG00000049964	ENSMUSG00000025069
					ENSRNOG00000012801	
GSTP1	P09211	glutathione S-transferase pi 1	ENSG00000084207	ENSMMUG00000018524	ENSRNOG00000018287	-
GSTT1	P30711	glutathione S-transferase theta 1	ENSG00000184674	ENSMMUG00000020989	ENSRNOG00000049771	ENSMUSG00000001663
GSTT2	P0CG29	glutathione S-transferase theta 2	ENSG00000099984	ENSMMUG00000018530	ENSRNOG00000028188	ENSMUSG00000033318
				ENSMMUG00000023053		
GSTZ1	O43708	glutathione transferase zeta 1	ENSG00000100577	ENSMMUG00000001097	ENSRNOG00000047708	ENSMUSG00000021033
HAGH	Q16775	hydroxyacylglutathione hydrolase	ENSG00000063854	ENSMMUG00000000737	ENSRNOG00000014743	ENSMUSG00000024158
HNF4A	P41235	hepatocyte nuclear factor 4, alpha	ENSG00000101076	ENSMMUG00000006464	ENSRNOG00000008895	ENSMUSG00000017950
HNMT	P50135	histamine N-methyltransferase	ENSG00000150540	ENSMMUG00000006918	ENSRNOG00000005223	ENSMUSG00000026986
				ENSMMUG00000007849		
HSD11B1	P28845	hydroxysteroid (11-beta) dehydrogenase 1	ENSG00000117594	ENSMMUG00000001014	-	ENSMUSG00000016194
HSD17B11	Q8NBQ5	hydroxysteroid (17-beta) dehydrogenase 11	ENSG00000198189	ENSMMUG00000019041	ENSRNOG00000002210	ENSMUSG00000029311

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
HSD17B14	Q9BPX1	hydroxysteroid (17-beta) dehydrogenase 14	ENSG00000087076	ENSMMUG00000028736	ENSRNOG00000020949	ENSMUSG00000030825
IAPP	P10997	islet amyloid polypeptide	ENSG00000121351	ENSMMUG00000032087	ENSRNOG00000012417	ENSMUSG00000041681
INMT	O95050	indolethylamine N-methyltransferase	ENSG00000241644	ENSMMUG00000011760	-	-
KCNJ11	Q14654	potassium inwardly-rectifying channel, subfamily J, member 11	ENSG00000187486	ENSMMUG00000023580	ENSRNOG00000021128	-
LPO	P22079	lactoperoxidase	ENSG00000167419	ENSMMUG00000002264	ENSRNOG00000008422	ENSMUSG00000009356
MAOA	P21397	monoamine oxidase A	ENSG00000189221	ENSMMUG00000031960 ENSMMUG00000031961 ENSMMUG00000019489	ENSRNOG00000002848	ENSMUSG00000025037
MAOB	P27338	monoamine oxidase B	ENSG00000069535	ENSMMUG00000019492	ENSRNOG00000029778	ENSMUSG00000040147
MAT1A	Q00266	methionine adenosyltransferase I, alpha	ENSG00000151224	ENSMMUG00000018634	ENSRNOG00000011351	ENSMUSG00000037798
METAP1	P53582	methionyl aminopeptidase 1	ENSG00000164024	ENSMMUG00000011062	ENSRNOG00000012947	ENSMUSG00000005813
MGST1	P10620	microsomal glutathione S-transferase 1	ENSG00000008394	ENSMMUG00000002359	ENSRNOG00000007743	ENSMUSG00000008540
MGST2	Q99735	microsomal glutathione S-transferase 2	ENSG00000085871	ENSMMUG00000021999	ENSRNOG00000013296	ENSMUSG00000074604
MGST3	O14880	microsomal glutathione S-transferase 3	ENSG00000143198	ENSMMUG00000015948	ENSRNOG00000003606 ENSRNOG00000004245	ENSMUSG00000026688
MPO	P05164	myeloperoxidase	ENSG00000005381	ENSMMUG00000002266	ENSRNOG00000008310	ENSMUSG00000009350
NAT1	P18440	N-acetyltransferase 1 (arylamine N-acetyltransferase)	ENSG00000171428	ENSMMUG00000005920	ENSRNOG00000049498	ENSMUSG00000056426
NAT2	P11245	N-acetyltransferase 2 (arylamine N-acetyltransferase)	ENSG00000156006	ENSMMUG00000021680	ENSRNOG00000029977 ENSRNOG00000014055 ENSRNOG00000049498 ENSRNOG00000029977 ENSRNOG00000014055	ENSMUSG00000056426
NNMT	P40261	nicotinamide N-methyltransferase	ENSG00000166741	ENSMMUG00000007086	ENSRNOG00000005930	ENSMUSG00000032271
NOS1	P29475	nitric oxide synthase 1 (neuronal)	ENSG00000089250	ENSMMUG00000014027	ENSRNOG00000001130	ENSMUSG00000029361
NOS2	P35228	nitric oxide synthase 2, inducible	ENSG00000007171	ENSMMUG00000010083	ENSRNOG00000048887 ENSRNOG00000011023 ENSRNOG00000049980	ENSMUSG00000020826
NOS3	P29474	nitric oxide synthase 3 (endothelial cell)	ENSG00000164867	ENSMMUG00000020437	ENSRNOG00000009348	ENSMUSG00000028978
NQO1	P15559	NAD(P)H dehydrogenase, quinone 1	ENSG00000181019	ENSMMUG00000019684	ENSRNOG00000012772	ENSMUSG00000003849
NQO2	P16083	NAD(P)H dehydrogenase, quinone 2	ENSG00000124588	ENSMMUG00000008487	ENSRNOG00000017820	ENSMUSG00000046949

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
NR1I2	O75469	nuclear receptor subfamily 1, group I, member 2	ENSG00000144852	ENSMMUG00000001976	ENSRNOG00000002906	ENSMUSG00000022809
NR1I3	Q14994	nuclear receptor subfamily 1, group I, member 3	ENSG00000143257	ENSMMUG00000018128	ENSRNOG00000003260	ENSMUSG00000005677
PAOX	Q6QHF9	polyamine oxidase (exo-N4-amino)	ENSG00000148832	ENSMMUG00000006326	ENSRNOG00000018838	ENSMUSG00000025464
PDE3A	Q14432	phosphodiesterase 3A, cGMP-inhibited	ENSG00000172572	ENSMMUG00000008269	ENSRNOG00000025042	ENSMUSG00000041741
PDE3B	Q13370	phosphodiesterase 3B, cGMP-inhibited	ENSG00000152270	ENSMMUG00000000759	ENSRNOG00000011417	ENSMUSG00000030671
PEMT	Q9UBM1	phosphatidylethanolamine N-methyltransferase	ENSG00000133027	ENSMMUG00000006495	ENSRNOG00000043216	ENSMUSG00000000301
PLGLB1	Q02325	plasminogen-like B1	ENSG00000183281	-	-	-
PLGLB2	Q02325	plasminogen-like B2	ENSG00000125551	-	-	-
PNMT	P11086	phenylethanolamine N-methyltransferase	ENSG00000141744	ENSMMUG00000001346	ENSRNOG00000046057	ENSMUSG00000038216
POMT1	Q9Y6A1	protein-O-mannosyltransferase 1	ENSG00000130714	ENSMMUG00000009975	ENSRNOG00000010477	ENSMUSG00000039254
PON1	P27169	paraoxonase 1	ENSG00000005421	ENSMMUG00000008011	ENSRNOG00000008902	ENSMUSG00000002588
PON2	Q15165	paraoxonase 2	ENSG00000105854	ENSMMUG00000008014	ENSRNOG00000009112	ENSMUSG00000032667
PON3	Q15166	paraoxonase 3	ENSG00000105852	ENSMMUG00000008013	ENSRNOG00000009096	ENSMUSG00000029759
POR	P16435	P450 (cytochrome) oxidoreductase	ENSG00000127948	ENSMMUG00000020086	ENSRNOG00000001442	ENSMUSG00000005514
PPARA	Q07869	peroxisome proliferator-activated receptor alpha	ENSG00000186951	ENSMMUG00000016258	ENSRNOG00000021463	ENSMUSG00000022383
PPARD	Q03181	peroxisome proliferator-activated receptor delta	ENSG00000112033	ENSMMUG00000015904	ENSRNOG00000000503	ENSMUSG00000002250
PPARG	P37231	peroxisome proliferator-activated receptor gamma	ENSG00000132170	ENSMMUG00000007191	ENSRNOG00000008839	ENSMUSG00000000440
PTGS1	P23219	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)	ENSG00000095303	ENSMMUG00000012886	ENSRNOG00000007415	ENSMUSG00000047250
PTGS2	P35354	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	ENSG00000073756	ENSMMUG00000016857	ENSRNOG00000002525	ENSMUSG00000032487
RXRA	P19793	retinoid X receptor, alpha	ENSG00000186350	ENSMMUG00000008183	ENSRNOG00000009446	ENSMUSG00000015846
SERPINA7	P05543	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 7	ENSG00000123561	ENSMMUG00000003266	ENSRNOG00000011081	ENSMUSG00000031271
SOD1	P00441	superoxide dismutase 1, soluble	ENSG00000142168	ENSMMUG00000001711	ENSRNOG00000002115	ENSMUSG00000022982
SOD2	P04179	superoxide dismutase 2, mitochondrial	ENSG00000112096	ENSMMUG00000005885	ENSRNOG00000019048	ENSMUSG00000006818
SOD3	P08294	superoxide dismutase 3, extracellular	ENSG00000109610	ENSMMUG00000002612	ENSRNOG00000003869	ENSMUSG00000072941
SULF1	Q8IWU6	sulfatase 1	ENSG00000137573	ENSMMUG00000023274	ENSRNOG00000009037	ENSMUSG00000016918

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
SULT1A1	P50225	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 1	ENSG00000196502	ENSMMUG00000022334 ENSMMUG00000022330 ENSMMUG00000011372	ENSRNOG00000019342	ENSMUSG00000030711
SULT1A2	P50226	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 2	ENSG00000197165	ENSMMUG00000022334 ENSMMUG00000022330 ENSMMUG00000011372	ENSRNOG00000019342	ENSMUSG00000030711
SULT1A3	P50224	Sulfotransferase 1A3/1A4	ENSG00000261052	ENSMMUG00000007332	ENSRNOG00000019342	ENSMUSG00000030711
SULT1A4	P50224	sulfotransferase family, cytosolic, 1A, phenol-preferring, member 4	ENSG00000213648	ENSMMUG00000007332	ENSRNOG00000019342	ENSMUSG00000030711
SULT1B1	O43704	sulfotransferase family, cytosolic, 1B, member 1	ENSG00000173597	ENSMMUG00000006538	ENSRNOG00000001967	ENSMUSG00000029269
SULT1C2	O00338	sulfotransferase family, cytosolic, 1C, member 2	ENSG00000198203	ENSMMUG00000003696	ENSRNOG00000047369 ENSRNOG00000040195 ENSRNOG00000031833 ENSRNOG00000042111 ENSRNOG00000040215 ENSRNOG00000050885	ENSMUSG00000023122
SULT1C4	O75897	sulfotransferase family, cytosolic, 1C, member 4	ENSG00000198075	ENSMMUG00000023779	-	-
SULT1E1	P49888	sulfotransferase family 1E, estrogen-preferring, member 1	ENSG00000109193	ENSMMUG00000004667	ENSRNOG00000001957	ENSMUSG00000029272
SULT2A1	Q06520	sulfotransferase family, cytosolic, 2A, dehydroepiandrosterone (DHEA)-preferring, member 1	ENSG00000105398	ENSMMUG00000000047	ENSRNOG00000047986 ENSRNOG00000042113 ENSRNOG00000046091 ENSRNOG00000042400	ENSMUSG00000030378 ENSMUSG00000094156 ENSMUSG00000070810 ENSMUSG00000078799 ENSMUSG00000074377 ENSMUSG00000074375 ENSMUSG00000070811 ENSMUSG00000078798
SULT2B1	O00204	sulfotransferase family, cytosolic, 2B, member 1	ENSG00000088002	ENSMMUG00000022155	ENSRNOG00000021046	ENSMUSG00000003271
SULT4A1	Q9BR01	sulfotransferase family 4A, member 1	ENSG00000130540	ENSMMUG00000005252	ENSRNOG00000046975	ENSMUSG00000018865
TPMT	P51580	thiopurine S-methyltransferase	ENSG00000137364	ENSMMUG00000016627	ENSRNOG00000016468	ENSMUSG00000021376
TPO	P07202	thyroid peroxidase	ENSG00000115705	ENSMMUG00000009662	ENSRNOG00000004646	ENSMUSG00000020673
TST	Q16762	thiosulfate sulfurtransferase (rhodanese)	ENSG00000128311	ENSMMUG00000012349	ENSRNOG00000000186	ENSMUSG00000044986
UGT1A1	P22309	UDP glucuronosyltransferase 1 family, polypeptide A1	ENSG00000241635	-	ENSRNOG00000018740	ENSMUSG00000089960

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
UGT1A10	Q9HAW8	UDP glucuronosyltransferase 1 family, polypeptide A10	ENSG00000242515	-	ENSRNOG00000048327	ENSMUSG00000089675
					ENSRNOG00000046998	ENSMUSG00000090124
					ENSRNOG00000050037	ENSMUSG00000090165
						ENSMUSG00000090175
UGT1A3	P35503	UDP glucuronosyltransferase 1 family, polypeptide A3	ENSG00000243135	-	-	ENSMUSG00000089943
						ENSMUSG00000090171
UGT1A4	P22310	UDP glucuronosyltransferase 1 family, polypeptide A4	ENSG00000244474	ENSMMUG00000021216	-	ENSMUSG00000089943
						ENSMUSG00000090171
UGT1A5	P35504	UDP glucuronosyltransferase 1 family, polypeptide A5	ENSG00000240224	ENSMMUG00000021216	-	ENSMUSG00000089943
						ENSMUSG00000090171
UGT1A6	P19224	UDP glucuronosyltransferase 1 family, polypeptide A6	ENSG00000167165	-	-	ENSMUSG00000054545
						ENSMUSG00000090145
UGT1A7	Q9HAW7	UDP glucuronosyltransferase 1 family, polypeptide A7	ENSG00000244122	-	ENSRNOG00000048327	ENSMUSG00000089675
					ENSRNOG00000046998	ENSMUSG00000090124
					ENSRNOG00000050037	ENSMUSG00000090165
						ENSMUSG00000090175
UGT1A8	Q9HAW9	UDP glucuronosyltransferase 1 family, polypeptide A8	ENSG00000242366	-	ENSRNOG00000048327	ENSMUSG00000089675
					ENSRNOG00000046998	ENSMUSG00000090124
					ENSRNOG00000050037	ENSMUSG00000090165
						ENSMUSG00000090175
UGT1A9	O60656	UDP glucuronosyltransferase 1 family, polypeptide A9	ENSG00000241119	-	ENSRNOG00000048327	ENSMUSG00000089675
					ENSRNOG00000046998	ENSMUSG00000090124
					ENSRNOG00000050037	ENSMUSG00000090165
						ENSMUSG00000090175
UGT2A1	Q9Y4X1	UDP glucuronosyltransferase 2 family, polypeptide A1, complex locus	ENSG00000173610	-	ENSRNOG00000001973	ENSMUSG00000029268
UGT2B10	P36537	UDP glucuronosyltransferase 2 family, polypeptide B10	ENSG00000109181	-	-	-
UGT2B11	O75310	UDP glucuronosyltransferase 2 family, polypeptide B11	ENSG00000213759	-	-	-
UGT2B15	P54855	UDP glucuronosyltransferase 2 family, polypeptide B15	ENSG00000196620	ENSMMUG00000023715	-	-
UGT2B17	O75795	UDP glucuronosyltransferase 2 family, polypeptide B17	ENSG00000197888	-	-	-

Continued on next page

Table 4 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
UGT2B28	Q9BY64	UDP glucuronosyltransferase 2 family, polypeptide B28	ENSG00000135226	-	-	-
UGT2B4	P06133	UDP glucuronosyltransferase 2 family, polypeptide B4	ENSG00000156096	ENSMMUG00000014773	-	-
				ENSMMUG00000012355		
UGT2B7	P16662	UDP glucuronosyltransferase 2 family, polypeptide B7	ENSG00000171234	ENSMMUG00000000337	-	-
UGT8	Q16880	UDP glycosyltransferase 8	ENSG00000174607	ENSMMUG00000014456	ENSRNOG00000009345	ENSMUSG00000032854
XDH	P47989	xanthine dehydrogenase	ENSG00000158125	ENSMMUG00000000561	ENSRNOG00000007081	ENSMUSG00000024066

Table 5: List of human transporters for all FDA approved drugs. Human orthologues of pharmacological relevant genes of *Macaca mulata*, *Rat norvegicus*, and *Mus musculus* were extracted from ENSEMBL Compara version 72 [144; 145]

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ABCA1	O95477	ATP-binding cassette, sub-family A (ABC1), member 1	ENSG00000165029	ENSMMUG00000020608	ENSRNOG00000018126	ENSMUSG00000015243
ABCA4	P78363	ATP-binding cassette, sub-family A (ABC1), member 4	ENSG00000198691	ENSMMUG00000013858	ENSRNOG00000012892	ENSMUSG00000028125
ABCB1	P08183	ATP-binding cassette, sub-family B (MDR/TAP), member 1	ENSG00000085563	ENSMMUG00000006764	ENSRNOG00000008012	ENSMUSG00000028970
					ENSRNOG00000022719	ENSMUSG00000040584
ABCB11	O95342	ATP-binding cassette, sub-family B (MDR/TAP), member 11	ENSG00000073734	ENSMMUG00000013832	-	ENSMUSG00000027048
ABCB4	P21439	ATP-binding cassette, sub-family B (MDR/TAP), member 4	ENSG00000005471	ENSMMUG00000010788	ENSRNOG00000029207	ENSMUSG00000042476
ABCB5	Q2M3G0	ATP-binding cassette, sub-family B (MDR/TAP), member 5	ENSG00000004846	ENSMMUG00000007821	ENSRNOG00000006398	ENSMUSG00000072791
ABCB6	Q9NP58	ATP-binding cassette, sub-family B (MDR/TAP), member 6	ENSG00000115657	ENSMMUG00000019616	ENSRNOG00000018697	ENSMUSG00000026198
ABCB7	O75027	ATP-binding cassette, sub-family B (MDR/TAP), member 7	ENSG00000131269	ENSMMUG00000015619	ENSRNOG00000002790	ENSMUSG00000031333
ABCB8	Q9NUT2	ATP-binding cassette, sub-family B (MDR/TAP), member 8	ENSG00000197150	ENSMMUG00000020442	ENSRNOG00000008557	ENSMUSG00000028973
ABCC1	P33527	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	ENSG00000103222	ENSMMUG00000029223	ENSRNOG00000032748	ENSMUSG00000023088
				ENSMMUG00000005914		
ABCC10	Q5T3U5	ATP-binding cassette, sub-family C (CFTR/MRP), member 10	ENSG00000124574	ENSMMUG00000029660	ENSRNOG00000018863	ENSMUSG00000032842
ABCC11	Q96J66	ATP-binding cassette, sub-family C (CFTR/MRP), member 11	ENSG00000121270	ENSMMUG00000008290	-	-
ABCC12	Q96J65	ATP-binding cassette, sub-family C (CFTR/MRP), member 12	ENSG00000140798	ENSMMUG00000002201	ENSRNOG00000015666	ENSMUSG00000036872
ABCC2	Q92887	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	ENSG00000023839	ENSMMUG00000023764	-	ENSMUSG00000025194
ABCC3	O15438	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	ENSG00000108846	ENSMMUG00000003414	ENSRNOG00000002948	ENSMUSG00000020865
ABCC4	O15439	ATP-binding cassette, sub-family C (CFTR/MRP), member 4	ENSG00000125257	ENSMMUG00000021807	ENSRNOG00000010064	ENSMUSG00000032849
					ENSRNOG00000027663 ENSRNOG00000043078	
ABCC5	O15440	ATP-binding cassette, sub-family C (CFTR/MRP), member 5	ENSG00000114770	ENSMMUG00000005515	ENSRNOG00000029178	ENSMUSG00000022822

Continued on next page

Table 5 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
ABCC6	O95255	ATP-binding cassette, sub-family C (CFTR/MRP), member 6	ENSG00000091262	ENSMMUG00000002527	ENSRNOG00000028781	ENSMUSG00000030834
ABCC8	Q09428	ATP-binding cassette, sub-family C (CFTR/MRP), member 8	ENSG00000006071	ENSMMUG00000023581	ENSRNOG00000021130	ENSMUSG00000040136
ABCC9	O60706	ATP-binding cassette, sub-family C (CFTR/MRP), member 9	ENSG00000069431	ENSMMUG00000020806	ENSRNOG00000036960	ENSMUSG00000030249
ABCG1	P45844	ATP-binding cassette, sub-family G (WHITE), member 1	ENSG00000160179	ENSMMUG00000018058	ENSRNOG00000001158	ENSMUSG00000024030
ABCG2	Q9UNQ0	ATP-binding cassette, sub-family G (WHITE), member 2	ENSG00000118777	ENSMMUG00000008797	ENSRNOG00000007041	ENSMUSG00000029299
					ENSRNOG00000022032 ENSRNOG00000030216 ENSRNOG00000031458 ENSRNOG00000039890	ENSMUSG00000029802
SLC10A1	Q14973	solute carrier family 10 (sodium/bile acid cotransporter family), member 1	ENSG00000100652	ENSMMUG00000022409	ENSRNOG00000005794	ENSMUSG00000021135
SLC10A2	Q12908	solute carrier family 10 (sodium/bile acid cotransporter family), member 2	ENSG00000125255	ENSMMUG00000007664	ENSRNOG00000037753	ENSMUSG00000023073
SLC13A1	Q9BZW2	solute carrier family 13 (sodium/sulfate symporters), member 1	ENSG00000081800	ENSMMUG00000016770	ENSRNOG00000008121	ENSMUSG00000029700
SLC13A2	Q13183	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 2	ENSG00000007216	-	ENSRNOG00000010337	ENSMUSG00000001095
SLC13A3	Q8WWT9	solute carrier family 13 (sodium-dependent dicarboxylate transporter), member 3	ENSG00000158296	ENSMMUG00000016924	ENSRNOG00000019118	ENSMUSG00000018459
SLC15A1	P46059	solute carrier family 15 (oligopeptide transporter), member 1	ENSG00000088386	ENSMMUG00000001402	ENSRNOG00000011598	ENSMUSG00000025557
SLC15A2	Q16348	solute carrier family 15 (H+/peptide transporter), member 2	ENSG00000163406	ENSMMUG00000001049	ENSRNOG00000002305	ENSMUSG00000022899
SLC16A1	P53985	solute carrier family 16, member 1 (monocarboxylic acid transporter 1)	ENSG00000155380	ENSMMUG00000006171	ENSRNOG00000019996	ENSMUSG00000032902
SLC19A1	P41440	solute carrier family 19 (folate transporter), member 1	ENSG00000173638	ENSMMUG00000004477	ENSRNOG00000001232	ENSMUSG00000001436
SLC22A1	O15245	solute carrier family 22 (organic cation transporter), member 1	ENSG00000175003	ENSMMUG00000015662	ENSRNOG00000016337	ENSMUSG00000023829
SLC22A10	Q63ZE4	solute carrier family 22, member 10	ENSG00000184999	-	-	-

Continued on next page

Table 5 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
SLC22A11	Q9NSA0	solute carrier family 22 (organic anion/urate transporter), member 11	ENSG00000168065	ENSMMUG00000015372	-	-
SLC22A12	Q96S37	solute carrier family 22 (organic anion/urate transporter), member 12	ENSG00000197891	ENSMMUG00000022594	ENSRNOG00000021108	ENSMUSG00000061742
SLC22A13	Q9Y226	solute carrier family 22 (organic anion transporter), member 13	ENSG00000172940	ENSMMUG00000006480	ENSRNOG00000033660	ENSMUSG00000074028
SLC22A14	Q9Y267	solute carrier family 22, member 14	ENSG00000144671	ENSMMUG00000004842	ENSRNOG00000042660 ENSRNOG00000042259	ENSMUSG00000070280
SLC22A15	Q8IZD6	solute carrier family 22, member 15	ENSG00000163393	ENSMMUG00000011860	ENSRNOG00000033051	ENSMUSG00000033147
SLC22A16	Q86VW1	solute carrier family 22 (organic cation/carnitine transporter), member 16	ENSG00000004809	ENSMMUG00000008187	-	ENSMUSG00000019834
SLC22A17	Q8WUG5	solute carrier family 22, member 17	ENSG00000092096	ENSMMUG00000014748	ENSRNOG00000016414	ENSMUSG00000022199
SLC22A18	Q96BI1	solute carrier family 22, member 18	ENSG00000110628	ENSMMUG00000007751	ENSRNOG00000020562	ENSMUSG00000000154
SLC22A18AS	Q8N1D0	solute carrier family 22 (organic cation transporter), member 18 antisense	ENSG00000254827	-	-	-
SLC22A2	O15244	solute carrier family 22 (organic cation transporter), member 2	ENSG00000112499	ENSMMUG00000029371 ENSMMUG00000005715	ENSRNOG00000016625	ENSMUSG00000040966
SLC22A3	O75751	solute carrier family 22 (extraneuronal monoamine transporter), member 3	ENSG00000146477	ENSMMUG00000005717	ENSRNOG00000022946	ENSMUSG00000023828
SLC22A4	Q9H015	solute carrier family 22 (organic cation/ergothioneine transporter), member 4	ENSG00000197208	ENSMMUG00000009301	-	ENSMUSG00000020334
SLC22A5	O76082	solute carrier family 22 (organic cation/carnitine transporter), member 5	ENSG00000197375	ENSMMUG00000007087	ENSRNOG00000008432	ENSMUSG00000018900 ENSMUSG00000063652
SLC22A6	Q4U2R8	solute carrier family 22 (organic anion transporter), member 6	ENSG00000197901	ENSMMUG00000016227	ENSRNOG00000018215	ENSMUSG00000024650
SLC22A7	Q9Y694	solute carrier family 22 (organic anion transporter), member 7	ENSG00000137204	-	ENSRNOG00000018420	ENSMUSG00000067144
SLC22A8	Q8TCC7	solute carrier family 22 (organic anion transporter), member 8	ENSG00000149452	ENSMMUG00000016225	ENSRNOG00000018086	ENSMUSG00000063796
SLC22A9	Q8IVM8	solute carrier family 22 (organic anion transporter), member 9	ENSG00000149742	-	ENSRNOG00000017964	ENSMUSG00000024757

Continued on next page

Table 5 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
					ENSRNOG00000021214	ENSMUSG00000052562
					ENSRNOG00000029998	ENSMUSG00000053303
					ENSRNOG00000031353	ENSMUSG00000063590
						ENSMUSG00000067656
						ENSMUSG00000075044
SLC27A1	Q6PCB7	solute carrier family 27 (fatty acid transporter), member 1	ENSG00000130304	ENSMMUG00000000194	ENSRNOG00000018170	ENSMUSG00000031808
SLC28A1	O00337	solute carrier family 28 (sodium-coupled nucleoside transporter), member 1	ENSG00000156222	ENSMMUG00000016772	ENSRNOG00000018940	ENSMUSG00000025726
SLC28A2	O43868	solute carrier family 28 (sodium-coupled nucleoside transporter), member 2	ENSG00000137860	ENSMMUG00000001284	ENSRNOG00000018204	ENSMUSG00000027219
					ENSRNOG00000028668	ENSMUSG00000079071
SLC28A3	Q9HAS3	solute carrier family 28 (sodium-coupled nucleoside transporter), member 3	ENSG00000197506	ENSMMUG00000008685	ENSRNOG00000018853	ENSMUSG00000021553
SLC29A1	Q99808	solute carrier family 29 (nucleoside transporters), member 1	ENSG00000112759	ENSMMUG00000022685	ENSRNOG00000019752	ENSMUSG00000023942
SLC29A2	Q14542	solute carrier family 29 (nucleoside transporters), member 2	ENSG00000174669	ENSMMUG00000015984	ENSRNOG00000020025	ENSMUSG00000024891
SLC2A4	P14672	solute carrier family 2 (facilitated glucose transporter), member 4	ENSG00000181856	ENSMMUG00000022709	ENSRNOG00000017226	ENSMUSG00000018566
SLC2A5	P22732	solute carrier family 2 (facilitated glucose/fructose transporter), member 5	ENSG00000142583	ENSMMUG00000016307	ENSRNOG00000017693	ENSMUSG00000028976
SLC5A6	Q9Y289	solute carrier family 5 (sodium-dependent vitamin transporter), member 6	ENSG00000138074	ENSMMUG00000010257	ENSRNOG00000006556	ENSMUSG00000006641
SLC6A6	P31641	solute carrier family 6 (neurotransmitter transporter, taurine), member 6	ENSG00000131389	ENSMMUG00000006340	-	-
SLC7A5	Q01650	solute carrier family 7 (amino acid transporter light chain, L system), member 5	ENSG00000103257	ENSMMUG00000001595	ENSRNOG00000018824	ENSMUSG00000040010
SLC7A7	Q9UM01	solute carrier family 7 (amino acid transporter light chain, y+L system), member 7	ENSG00000155465	ENSMMUG00000000413	ENSRNOG00000010296	ENSMUSG00000000958
SLC7A8	Q9UHI5	solute carrier family 7 (amino acid transporter light chain, L system), member 8	ENSG00000092068	ENSMMUG00000019653	ENSRNOG00000014311	ENSMUSG00000022180

Continued on next page

Table 5 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
SLCO1A2	P46721	solute carrier organic anion transporter family, member 1A2	ENSG00000084453	ENSMMUG00000008584	ENSRNOG00000010388	ENSMUSG00000030237
					ENSRNOG00000030894	ENSMUSG00000041698
					ENSRNOG00000031249	ENSMUSG00000063975
					ENSRNOG00000036984	ENSMUSG00000079262
						ENSMUSG00000079263
						ENSMUSG00000084927
SLCO1B1	Q9Y6L6	solute carrier organic anion transporter family, member 1B1	ENSG00000134538	ENSMMUG00000008583	ENSRNOG00000030538	ENSMUSG00000030236
SLCO1B3	Q9NPD5	solute carrier organic anion transporter family, member 1B3	ENSG00000111700	ENSMMUG00000020443	ENSRNOG00000030538	ENSMUSG00000030236
SLCO1C1	Q9NYB5	solute carrier organic anion transporter family, member 1C1	ENSG00000139155	ENSMMUG00000008271	ENSRNOG00000009740	ENSMUSG00000030235
SLCO2A1	Q92959	solute carrier organic anion transporter family, member 2A1	ENSG00000174640	ENSMMUG00000012605	ENSRNOG00000009005	ENSMUSG00000032548
SLCO2B1	O94956	solute carrier organic anion transporter family, member 2B1	ENSG00000137491	ENSMMUG00000009126	ENSRNOG00000017976	ENSMUSG00000030737
SLCO3A1	Q9UIG8	solute carrier organic anion transporter family, member 3A1	ENSG00000176463	ENSMMUG00000001740	ENSRNOG00000032798	ENSMUSG00000025790
SLCO4A1	Q96BD0	solute carrier organic anion transporter family, member 4A1	ENSG00000101187	ENSMMUG00000016848	ENSRNOG00000021870	ENSMUSG00000038963
SLCO4C1	Q6ZQN7	solute carrier organic anion transporter family, member 4C1	ENSG00000173930	ENSMMUG00000007773	ENSRNOG00000022711	ENSMUSG00000040693
SLCO5A1	Q9H2Y9	solute carrier organic anion transporter family, member 5A1	ENSG00000137571	ENSMMUG00000008783	ENSRNOG00000008966	ENSMUSG00000025938
SLCO6A1	Q86UG4	solute carrier organic anion transporter family, member 6A1	ENSG00000205359	ENSMMUG00000029885	ENSRNOG00000013792	ENSMUSG00000026331
					ENSRNOG00000019252	ENSMUSG00000026336
					ENSRNOG00000031146	ENSMUSG00000045463
TAP1	Q03518	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	ENSG00000232367	ENSMMUG00000029790	ENSRNOG00000000457	ENSMUSG00000037321
			ENSG00000230705			
			ENSG00000227816			
			ENSG00000226173			
			ENSG00000224212			
			ENSG00000206297			
			ENSG00000168394			
TAP2	Q03519	transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	ENSG00000228582	ENSMMUG00000011293	ENSRNOG00000000455	ENSMUSG00000024339
			ENSG00000237599			
			ENSG00000225967			
			ENSG00000206235			

Continued on next page

Table 5 – *Continued from previous page*

Gene symbol	Uniprot ID	Description	Human ENSEMBL ID	Monkey ENSEMBL ID	Rat ENSEMBL ID	Mouse ENSEMBL ID
			ENSG00000223481			
			ENSG00000206299			
			ENSG00000232326			
			ENSG00000204267			

Appendix C

Table 6: List of human gene expression datasets with information on age and gender. (*) Processed data. Childhood: 0-18 years; Adulthood: 18-65 years; Elderhood: >65 years.

Tissue	Region	Experiment	Array	Number samples	Gender	Childhood data	Adulthood data	Elderhood data
Brain	Amygdala	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	76	M, F	Y	Y	Y
	Caudate nucleus	E-GEOD-11512	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	13	M, F	Y	Y	N
		E-GEOD-3790	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	32	M, F	N	Y	Y
		E-AFMX-6	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	32	M, F	N	Y	Y
	Cerebellum	E-GEOD-35974	Affymetrix GeneChip Human Gene 1.0 ST Array	49	M, F	N	Y	Y
		E-GEOD-3790	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	26	M, F	N	Y	Y
		E-AFMX-6	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	26	M, F	N	Y	Y
	Cerebral cortex	*E-GEOD-44768	Rosetta/Merck Human 44k 1.1 microarray	100	M, F	N	Y	Y
		*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	71	M, F	Y	Y	Y
	Frontal cortex	E-GEOD-22569	Affymetrix GeneChip Human Gene 1.0 ST Array	25	M, F	Y	Y	Y
		*E-GEOD-1572	Affymetrix GeneChip Human Genome U95Av2	30	M, F	N	Y	Y
		E-GEOD-3790	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	28	M, F	N	Y	Y
		E-AFMX-6	Affymetrix GeneChip Human Genome HG-U133A and HG-U133B	16	M, F	N	Y	Y
	Hippocampus	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	417	M, F	Y	Y	Y
		*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	82	M, F	Y	Y	Y
	Mediodorsal nucleus of the thalamus	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	73	M, F	Y	Y	Y
	Occipital cortex	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	83	M, F	Y	Y	Y
	Parietal cortex	E-GEOD-35977	Affymetrix GeneChip Human Gene 1.0 ST Array	48	M, F	N	Y	Y
	Prefrontal cortex	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	157	M, F	Y	Y	Y
		E-GEOD-17757	Affymetrix GeneChip Human Gene 1.0 ST Array	25	M, F	Y	Y	Y
		E-GEOD-13564	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	44	M, F	Y	Y	N
		E-GEOD-11512	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	44	M, F	Y	Y	Y
	Striatum	*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	75	M, F	Y	Y	Y

Continued on next page

Table 6 – *Continued from previous page*

Tissue	Region	Experiment	Array	Number samples	Gender	Childhood data	Adulthood data	Elderhood data
	Temporal cortex	*E-GEOD-37721	Illumina HumanHT-12 v3.0 Expression BeadChip	30	M, F	Y	Y	N
		*E-GEOD-25219	Affymetrix GeneChip Human Exon 1.0 ST	247	M, F	Y	Y	Y
Liver	-	*SAGE	Rosetta/Merck Human 44k 1.1 microarray	427	M, F	Y	Y	Y
	-	E-MTAB-950	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	10	M, F	N	Y	Y
	-	E-MEXP-3291	Affymetrix GeneChip Human Gene 1.0 ST Array	19	M, F	N	Y	Y
	-	*E-GEOD-28893	Illumina HumanHT-6 v2.0 Expression BeadChip	60	M, F	Y	Y	Y
	-	E-GEOD-25935	Agilent Whole Human Genome Microarray 4x44K	322	M, F	Y	Y	Y
Kidney	-	E-GEOD-9489	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	13	M, F	N	Y	Y
	-	*E-GEOD-40435	Illumina HumanHT-12 V4.0 expression beadchip	101	M, F	N	Y	Y
	-	E-GEOD-11024	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	12	M, F	Y	Y	Y
Heart	-	E-GEOD-22253	Affymetrix GeneChip Human Gene 1.0 ST Array	107	M, F	Y	Y	Y
	-	E-GEOD-21610	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	8	M, F	Y	Y	N
Lungs	-	E-GEOD-47460	Agilent Whole Human Genome Microarray 4x44K	89	M, F	N	Y	Y
	-	*E-GEOD-32863	Illumina HumanHT-6 v3.0 Expression BeadChip	58	M, F	N	Y	Y
	-	E-GEOD-32537	Affymetrix GeneChip Human Gene 1.0 ST Array	50	M, F	Y	Y	Y
	-	*E-GEOD-1643	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	40	M, F	N	Y	Y
	-	E-GEOD-11784	Affymetrix GeneChip Human Genome U133 Plus 2.0 Array	63	M, F	N	Y	Y

Table 7: List of macaque gene expression datasets with information on age and gender. Childhood: 0-4 years; Adulthood: 4-20 years; Elderhood: >20 years.

Tissue	Region	Experiment	Array	Number samples	Gender	Childhood data	Adulthood data	Elderhood data
Brain	Hippocampus (HIP)	E-GEOD-13824	Affymetrix GeneChip Rhesus Macaque Genome Array	24	M	Y	Y	N
Liver	-	E-GEOD-22624	Affymetrix GeneChip Rhesus Macaque Genome Array	5	M	N	Y	N

Table 8: List of rat gene expression datasets with information on age and gender. (*) Processed data. Childhood: 0-2.5 months; Adulthood: 2.5-20 months; Elderhood: >20 months. NA: not available.

Tissue	Region	Experiment	Array	Number samples	Gender	Strain	Childhood data	Adulthood data	Elderhood data
Brain	All	E-GEOD-5666	Affymetrix GeneChip Rat Expression Array RAE230A	20	M	F344	N	Y	Y
	Cerebral cortex	E-MEXP-1656	Affymetrix GeneChip Rat Genome 230 2.0	12	M	NA	Y	N	N
		E-GEOD-34272	RNA-seq	7	M	BN rats (Substrain BN/SsNOlaHSD)	N	Y	Y
	Frontal cortex	*E-GEOD-36722	Agilent-028279 SurePrint G3 Rat GE 8x60K Microarray	8	M	LOU/C/Jall	N	Y	Y
		E-GEOD-24515	Affymetrix GeneChip Rat Genome 230 2.0	12	M	Fischer 344	N	Y	Y
	Hippocampus	E-MEXP-553	Affymetrix GeneChip Rat Genome 230 2.0	8	M	Long-Evans	N	Y	Y
		E-MEXP-1656	Affymetrix GeneChip Rat Genome 230 2.0	12	M	NA	Y	N	N
		E-GEOD-854	Affymetrix GeneChip Rat Genome U34A	29	M	Fischer 344	N	Y	Y
		*E-GEOD-36722	Agilent-028279 SurePrint G3 Rat GE 8x60K Microarray	8	M	LOU/C/Jall	N	Y	Y
Liver	-	E-GEOD-27625	Affymetrix GeneChip Rat Genome 230 2.0	15	M	Fischer 344	N	Y	Y
	-	E-MTAB-800	Affymetrix GeneChip Rat Genome 230 2.0	12	M	Crj:CD(SD)IGS	Y	Y	N
	-	E-GEOD-7654	INRA-INSA Rat Operon V3 oligo array	12	M	LEC (Long Evans Cinnamon-like)	Y	Y	N
	-	E-GEOD-32156	Affymetrix GeneChip Rat Genome 230 2.0	10	M	Wistar/Han	Y	Y	N
	-	E-GEOD-17384	Affymetrix GeneChip Rat Genome 230 2.0	8	M	Long Evans Cinnamon (LEC/Crj)	Y	Y	N
Kidney	-	*E-GEOD-27725	Illumina RatRef-12 v1 Expression BeadChip	31	M, F	Fawn-Hooded Hypertensive	Y	Y	N
	-	E-MTAB-613	Illumina RatRef-12 v1 Expression BeadChip	21	M	Wistar Hannover Taconic	Y	Y	N
	-	E-MTAB-800	Affymetrix GeneChip Rat Genome 230 2.0	12	M	Crj:CD(SD)IGS	Y	Y	N
Heart	-	E-GEOD-24110	Affymetrix GeneChip Rat Genome 230 2.0	8	M	Dahl salt sensitive (Harlan SS/Hsd)	N	Y	N

Table 9: List of mouse gene expression datasets with information on age and gender. (*) Processed data. Childhood: 0-2 months; Adulthood: 2-12 months; Elderhood: >12 months. NA: not available.

Tissue	Region	Experiment	Array	Number samples	Gender	Strain	Childhood data	Adulthood data	Elderhood data
Brain	All	E-MEXP-1005	Affymetrix GeneChip Mouse Expression Array MOE430A	9	F	NA	N	Y	N
		*E-GEOD-16675	Affymetrix GeneChip Mouse Genome 430 2.0	12	M	C57BL/6J	Y	Y	N
	Brainstem	E-GEOD-20547	Affymetrix GeneChip Mouse Genome 430 2.0	6	M, F	NA	N	Y	Y
	Cerebellum	E-GEOD-9914	Affymetrix GeneChip Mouse Genome 430 2.0	6	F	Sca1154Q/2Q	Y	Y	N
		E-GEOD-3248	Affymetrix GeneChip Murine Genome U74Av2	23	M, F	NES144	N	Y	Y
		E-GEOD-20547	Affymetrix GeneChip Mouse Genome 430 2.0	8	M, F	NA	N	Y	Y
		E-GEOD-11596	CNIO Mouse chip 27.6K 3.0	4	M	Mecp2tm1.1Bird/J with C57BL/6J	Y	Y	N
	Cortex	E-GEOD-11596	CNIO Mouse chip 27.6K 3.0	4	M	Mecp2tm1.1Bird/J with C57BL/6J	Y	Y	N
	Hippocampus	E-MEXP-962	Affymetrix GeneChip Mouse Genome 430 2.0	9	M	NA	N	Y	N
		E-GEOD-35210	Phalanx Mouse OneArray	5	F	129SV with C57BL6	N	Y	Y
		E-GEOD-1482	Affymetrix GeneChip Murine Genome U74Av2	15	M	NF+/+	Y	N	N
	Hypothalamus	E-GEOD-21278	Affymetrix GeneChip Mouse Genome 430 2.0	24	M, F	C57BL/6J	Y	N	N
	Midbrain	E-GEOD-11596	CNIO Mouse chip 27.6K 3.0	4	M	Mecp2tm1.1Bird/J with C57BL/6J	Y	Y	N
	Prefrontal cortex	E-GEOD-4675	Affymetrix GeneChip Mouse Genome 430 2.0	23	M	C57BL/6J	Y	Y	N
	Striatum	E-GEOD-20547	Affymetrix GeneChip Mouse Genome 430 2.0	6	M, F	NA	N	Y	Y
Liver	-	E-MEXP-839	Affymetrix GeneChip Mouse Genome 430 2.0	13	M	C57BL/6J	N	Y	Y
	-	E-MEXP-153	Affymetrix GeneChip Mouse Expression Array MOE430A	11	M	NA	N	Y	Y
	-	E-MEXP-1504	Affymetrix GeneChip Mouse Genome 430 2.0	6	F	C57BL/6J	N	Y	Y
	-	E-GEOD-6998	Affymetrix GeneChip Mouse Genome 430 2.0	12	F	NA	Y	N	N
	-	E-GEOD-33467	Affymetrix GeneChip Mouse Gene 1.0 ST Array	30	F	BALB/c	Y	Y	N

Continued on next page

Table 9 – *Continued from previous page*

Tissue	Region	Experiment	Array	Number samples	Gender	Strain	Childhood data	Adulthood data	Elderhood data
	-	E-GEOD-21716	Affymetrix GeneChip Mouse Genome 430 2.0	16	M	C57BL/6J	N	Y	Y
		E-GEOD-21224	Affymetrix GeneChip Mouse Genome 430 2.0	16	M	C57BL/6J	Y	Y	N
	-	E-BAIR-10	Affymetrix GeneChip Mouse Genome 430 2.0	18	M	C57BL/6O	N	Y	Y
Kidney	-	E-MEXP-1504	Affymetrix GeneChip Mouse Genome 430 2.0	6	F	C57BL/6J	N	Y	Y
	-	E-GEOD-38754	Affymetrix GeneChip Mouse Genome 430 2.0	15	M	C57BL/6J	Y	Y	N
Heart	-	E-GEOD-75	Affymetrix GeneChip Murine Genome U74A	24	M, F, mix	FVB	Y	Y	N
	-	E-GEOD-38754	Affymetrix GeneChip Mouse Genome 430 2.0	10	M	C57BL/6J	Y	N	N
	-	*E-MTAB-500-5001	Illumina MouseWG 6 v1.0 expression beadchip	6	M	NA	N	Y	N
Lungs	-	E-GEOD-6591	Affymetrix GeneChip Mouse Genome 430 2.0	8	M	C57BL/6J	N	Y	Y
	-	E-GEOD-38754	Affymetrix GeneChip Mouse Genome 430 2.0	15	M	C57BL/6J	Y	N	N
	-	E-MEXP-1504	Affymetrix GeneChip Mouse Genome 430 2.0	6	F	C57BL/6J	N	Y	Y

Table 10: List of gene expression experiments which meet the requirements from section 4.1.1 in chapter 4, with the exception of the tissue type. These experiments have, therefore, not been analysed.

Tissue	*Homo sapiens*	*Macaca mulatta*	*Rattus norvegius*	*Mus musculus*
Blood	E-TABM-305		E-GEOD-27023	E-GEOD-41879
	E-GEOD-22098			
	E-GEOD-29676			
	E-GEOD-19743			
	E-GEOD-17952			
	E-GEOD-21311			
	E-GEOD-12288			
	E-GEOD-10586			
	E-GEOD-10922			
	E-GEOD-12408			
	E-MTAB-145			
	E-MTAB-1039			
	E-GEOD-35846			
	E-GEOD-45986			
Other	E-GEOD-27342		E-GEOD-45892	E-GEOD-11186
	E-GEOD-11504		E-MEXP-3174	E-GEOD-6952
	E-TABM-325			E-GEOD-3351
	E-MEXP-1618			E-GEOD-5334
	E-GEOD-33335			E-GEOD-7502
	E-GEOD-40645			E-GEOD-6916
	E-GEOD-41762			E-GEOD-19532
	E-GEOD-31243			E-MEXP-1717
				E-MEXP-453
				E-GEOD-27017
				E-GEOD-6453
				E-CBIL-12
				E-BAIR-9
				E-GEOD-30295
				E-GEOD-34723
				E-GEOD-39699
				E-GEOD-39769
				E-TABM-531
				E-MEXP-3369
				E-GEOD-46200

Table 11: Expression patterns of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, in relevant human tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; PF - Prefrontal cortex; FC - Frontal cortex; PaC - Parietal cortex; OC - Occipital cortex; TC - Temporal cortex; CN - Caudate nucleus; MD - Mediodorsal nucleus of the thalamus; HIP - Hippocampus.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	GABRA1	Safety, Efficacy	-	24/24	Brain	N01AB, N01AX, N03AA, N05BA, N05CD
	GABRA2		Decrease	1/24 (1/2)	Brain (CC)	
	GABRA3		Decrease	1/14 (1/3)	Brain (PC)	
	GABRA4		Increase	1/22 (1/2)	Brain (CC)	
	GABRA5		**Decrease; Increase**	**4/24 (2/3, 1/2, 1/1); 1/24 (1/2)**	Brain (PF, PaC, OC); Brain (CC)	
	GABRA6		Increase	1/23 (1/2)	Brain (CC)	
	GABRB1		-	24/24	Brain	
	GABRB2		-	24/24	Brain	
	GABRB3		Decrease	1/24 (1/4)	Brain (FC)	
	GABRD		Increase	1/15 (1/2)	Brain (CC)	
	GABRE		-	15/15	Brain	
	GABRG1		**Decrease; Increase**	**2/17 (1/1, 1/2); 1/17 (1/3)**	Brain (OC, TC); Brain (CN)	
	GABRG2		**Increase**	**2/24 (1/2, 1/4)**	Brain (PaC, FC)	
	GABRG3		Decrease	1/23 (1/2)	Brain (CC)	
	GABRP		-	24/24	Brain	
	GABRQ		-	15/15	Brain	
	KCNK10	Safety	-	22/22	Brain	N01AB
	KCNK18		-	14/14	Brain	
	KCNK2		**Decrease**	**2/24 (1/2, 1/3)**	Brain (CC, CN)	
	KCNK3		-	25/25	Brain	
	KCNK9		-	16/16	Brain	
	CACNA1A	Safety, Efficacy	-	15/15	Brain	N03AX
	CACNA1B		-	24/24	Brain	
	CACNA1C		-	24/24	Brain	
	CACNA1D		-	24/24	Brain	
	CACNA1E		-	24/24	Brain	

Continued on next page

Table 11 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	CACNA1F		-	15/15	Brain	
	CACNA1G		**Decrease**	**1/14 (1/3)**	Brain (PF)	
	CACNA1H		-	15/15	Brain	
	CACNA1I		-	14/14	Brain	
	CACNA1S		-	23/23	Brain	
	CACNA2D1		**Decrease**	**2/23 (2/2)**	Brain (CC)	
	CACNA2D2		-	24/24	Brain	
	CACNA2D3		-	23/23	Brain	
	CACNA2D4		-	16/16	Brain	
	CACNB1		-	15/15	Brain	
	CACNB2		-	24/24	Brain	
	CACNB3		-	24/24	Brain	
	CACNB4		-	25/25	Brain	
	CACNG1		-	14/14	Brain	
	CACNG2		-	14/14	Brain	
	CACNG3		-	15/15	Brain	
	CACNG4		-	13/13	Brain	
	CACNG5		-	13/13	Brain	
	CACNG6		-	8/8	Brain	
	CACNG7		-	8/8	Brain	
	CACNG8		-	8/8	Brain	
	GRIN1	Control	-	24/24	Brain	N01AX, N03AX
	GRIN2A		**Decrease**	**3/16 (1/3, 1/4, 1/3)**	Brain(PF, FC, CN)	
	GRIN2B		Decrease	2/22 (1/3, 1/2)	Brain (PF, CC)	
	GRIN2C		-	15/15	Brain	
	GRIN2D		-	14/14	Brain	
	GRIN3A		**Decrease**	**7/17 (1/2, 1/3, 1/2, 1/1, 1/1, 1/1, 1/4)**	Brain (TC, PF, PaC, OC, MD, HIP, FC)	
	GRIN3B		-	7/7	Brain	
	HTR3A	Safety, Efficacy, Control	-	24/24	Brain	A04AA
	KCNH2	Control	-	2/2	Heart	C07AA
	SCN10A	Safety, Control	-	22/22	Brain	N03AF
	SCN11A		-	22/22	Brain	
	SCN1A		**Increase**	**5/23 (1/2, 1/3, 1/2, 1/1, 1/4)**	Brain (TC, PF, PaC, OC, FC)	
	SCN2A		Decrease	1/24 (1/4)	Brain (FC)	
	SCN3A		**Decrease**	**4/23 (2/3, 1/1, 1/1)**	Brain (PF, OC, MD)	
	SCN4A		-	14/14	Brain	
	SCN5A		-	23/23	Brain	
	SCN7A		-	24/24	Brain	
	SCN8A		-	23/23	Brain	

Continued on next page

Table 11 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	SCN9A		-	23/23	Brain	
Membrane receptor	DRD2	Safety, Efficacy, Control	-	23/23	Brain	A03FA, N05AB
	HTR2A	Safety, Efficacy, Control	**Increase; Decrease**	**2/24 (1/2, 1/2); 1/24 (1/4)**	Brain (TC, PaC); Brain (FC)	N06AX
	HTR2C	Safety, Efficacy, Control	Increase	1/15 (1/3)	Brain (CN)	N06AX
	HTR1D	Efficacy	-	22/22	Brain	N02CA, N02CC
	HTR1B	Efficacy	**Increase**	**2/23 (1/1/, 1/1)**	Brain (OC, MD)	N02CC
	DRD3	Control	-	23/23	Brain	N05AH, N05AX
	DRD4	Control	-	24/24	Brain	
	CHRM3	Safety, Efficacy, Control	-; -; Increase	4/4; 3/3; 1/5	Liver; Kidney; Lungs	A03BA, G04BD, R03BB
Transporter	SLC6A4	Safety, Efficacy, Control	-	15/15	Brain	N06AA, N06AB, N06BA, N06AX
	SLC6A2	Safety, Efficacy, Control	-	14/14	Brain	
	SLC6A3	Safety, Control	-	23/23	Brain	N06BA
	SLC12A3	Efficacy	-	2/2	Heart	C03AX
	SLC18A2	Control	-	23/23	Brain	N06BA
Enzyme	ALDH5A1	Safety	-	24/24	Brain	N03AG
	CSF1R	Safety	?	-	No relevant tissue to analyse	L01XE
	EGFR	Safety, Control	?	-	No relevant tissue to analyse	L01XE
	ABAT	Efficacy	-	15/15	Brain	N03AG
	MT-ND1	Efficacy	-	3/3	Liver	A10BA
	MT-ND2		-	3/3	Liver	
	MT-ND3		-	3/3	Liver	
	MT-ND4		-	3/3	Liver	
	MT-ND4L		-	3/3	Liver	
	MT-ND5		-	3/3	Liver	
	MT-ND6		-	3/3	Liver	
	NDUFA1		-	5/5	Liver	
	NDUFA10		-	5/5	Liver	
	NDUFA11		-	4/4	Liver	
	NDUFA12		-	5/5	Liver	

Continued on next page

Table 11 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	NDUFA13		-	5/5	Liver	
	NDUFA2		-	5/5	Liver	
	NDUFA3		-	5/5	Liver	
	NDUFA4		-	5/5	Liver	
	NDUFA4L2		-	4/4	Liver	
	NDUFA5		-	5/5	Liver	
	NDUFA6		-	5/5	Liver	
	NDUFA7		-	5/5	Liver	
	NDUFA8		-	5/5	Liver	
	NDUFA9		-	5/5	Liver	
	NDUFAB1		-	5/5	Liver	
	NDUFAF1		-	5/5	Liver	
	NDUFAF2		-	5/5	Liver	
	NDUFAF3		-	4/4	Liver	
	NDUFAF4		-	5/5	Liver	
	NDUFB1		-	5/5	Liver	
	NDUFB10		-	5/5	Liver	
	NDUFB11		-	4/4	Liver	
	NDUFB2		-	4/4	Liver	
	NDUFB3		-	5/5	Liver	
	NDUFB4		-	5/5	Liver	
	NDUFB5		-	5/5	Liver	
	NDUFB6		-	5/5	Liver	
	NDUFB7		-	5/5	Liver	
	NDUFB8		-	5/5	Liver	
	NDUFC1		-	5/5	Liver	
	NDUFC2		-	5/5	Liver	
	NDUFS2		-	5/5	Liver	
	NDUFS3		-	5/5	Liver	
	NDUFS4		-	5/5	Liver	
	NDUFS5		-	5/5	Liver	
	NDUFS6		-	5/5	Liver	
	NDUFS7		-	4/4	Liver	
	NDUFS8		-	5/5	Liver	
	NDUFV1		-	5/5	Liver	
	NDUFV2		-	5/5	Liver	
	NDUFV3		-	5/5	Liver	
	PTGS2	Efficacy, Control	?; -; -; -;	-; 2/2; 3/3; 24/24	Muscle; Heart; kidney; Brain	M01AB, C01EB, G02CC, M02AA, M01AX
	ERBB2	Control	?	-	No relevant tissue to analyse	L01XE

Continued on next page

Table 11 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Transcription factor	NR3C1	Safety	-; -; -; Decrease	5/5; 2/2; 24/24; 1/5	Liver; Heart; Brain; Lungs	A07EA, C05AA, H02AB, R01AD, R03BA
	PGR	Safety	Increase	1/5	Lungs	R01AD, R03BA
	RARB	Safety	-	24/24	Brain	D10AD
	NR1H4	Efficacy	-	5/5	Liver	A05AA
	VDR	Control	-; -; -;	5/5; 23/23; 3/3	Liver; Brain; Kidney	A11CC, D05AX, H05BX

Table 12: Expression patterns of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, in relevant human tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; PF - Prefrontal cortex; FC - Frontal cortex; PaC - Parietal cortex; OC - Occipital cortex; TC - Temporal cortex; CN - Caudate nucleus; MD - Mediodorsal nucleus of the thalamus; HIP - Hippocampus.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	GABRA1	Safety, Efficacy	-	24/24	Brain	N01AB, N01AX, N03AA, N05BA, N05CD
	GABRA2		Decrease	1/24 (1/2)	Brain (CC)	
	GABRA3		Decrease	1/14 (1/3)	Brain (PC)	
	GABRA4		Increase	1/22 (1/2)	Brain (CC)	
	GABRA5		**Decrease; Increase**	**4/24 (2/3, 1/2, 1/1); 1/24 (1/2)**	Brain (PF, PaC, OC); Brain (CC)	
	GABRA6		Increase	1/23 (1/2)	Brain (CC)	
	GABRB1		-	24/24	Brain	
	GABRB2		-	24/24	Brain	
	GABRB3		Decrease	1/24 (1/4)	Brain (FC)	
	GABRD		Increase	1/15 (1/2)	Brain (CC)	
	GABRE		-	15/15	Brain	
	GABRG1		**Decrease; Increase**	**2/17 (1/1, 1/2); 1/17 (1/3)**	Brain (OC, TC); Brain (CN)	
	GABRG2		**Increase**	**2/24 (1/2, 1/4)**	Brain (PaC, FC)	
	GABRG3		Decrease	1/23 (1/2)	Brain (CC)	
	GABRP		-	24/24	Brain	
	GABRQ		-	15/15	Brain	
	CACNA1C	Safety, Efficacy	-	2/2	Heart	C08CA, C08DA, C08DB
	CACNA1D		-	2/2	Heart	
	CACNA1F		-	2/2	Heart	
	CACNA1S		-	2/2	Heart	
	ABCC8	Safety	-	3/3	Liver	A10BB
	KCNJ11		-	4/4	Liver	
	CHRNA1	Efficacy, Control	?	-	No relevant tissue to analyse	M03AC, M03AA,
	CHRNB1		?	-	No relevant tissue to analyse	

Continued on next page

Table 12 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	CHRND		?	-	No relevant tissue to analyse	
	CHRNE		?	-	No relevant tissue to analyse	
	CHRNG		?	-	No relevant tissue to analyse	
	HTR3A	Efficacy	-	24/24	Brain	A03FA
	SCNN1A	Efficacy	-	2/2; 3/3	Heart; kidney	C03DB
	HRH1	Safety	-; Decrease	23/23; 1/5	Brain; Lungs	N05BB, R06AB, R06AD, R06AE
	OPRM1	Safety, Efficacy	-; -	4/4; 20/20	Liver; Brain	A07DA, N01AH, N02AB
	CHRM3	Safety	-; Increase	4/4; 1/5	Liver; Lungs	A03BA, R03BB
	ADRB1	Efficacy	-; Decrease	2/2; 1/5	Heart; Lungs	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA
	ADRB2		-; Decrease	2/2; 1/5	Heart; Lungs	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA, G04BD
	ADRB3	Efficacy, Control	-; -; -	2/2; 5/5; 3/3	Heart; Lungs; Kidney	A03FA, N05AB, C04AE
	DRD2	Safety, Efficacy	-	23/23	Brain	N06AX
	HTR2A	Control	**Increase; Decrease**	**2/24 (1/2, 1/2); 1/24 (1/4)**	Brain (TC, PaC); Brain (FC)	
	HTR2C	Control	Increase	1/15 (1/3)	Brain (CN)	N06AX
	SLC6A2	Safety, Efficacy, Control	-	14/14	Brain	N06AA, N06AB, N06AX
	SLC6A4	Safety, Efficacy, Control	-	15/15	Brain	N06AA, N06AB, N06AX
	SLC12A1	Safety, Efficacy	-	2/2	Heart	C03CA
	ATP1A1	Safety	-	2/2	Heart	C01AA

Continued on next page

Table 12 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	ATP1A2		-	2/2	Heart	
	ATP1A3		-	2/2	Heart	
	ATP1A4		-	2/2	Heart	
	SLC29A1	Efficacy	-	2/2	Heart	B01AC
Enzyme	PTGS1	Safety	-; Increase	4/4; 1/24 (1/1)	Liver; Brain (HIP)	A01AD, N02BA, M01AB
	PTGS2		-; -	5/5; 24/24	Liver; Brain (HIP)	
	PDE3A	Safety	Increase	1/5	Lungs	
	PDE3B		-	5/5	Lungs	R03DA
	ACHE	Efficacy	?	-	No relevant tissue to analyse	Unclassified
	PDE5A	Efficacy	-	2/2	Heart	B01AC
	ACE	Control	-	2/2	Heart	C09AA
	F2	Control	-	2/2	Heart	B01AE
	HMGCR	Control	-	3/3	Liver	C10AA
Transcription factor	NR3C1	Safety	-	24/24	Brain	D07AA, H02BX, D10AA, H02AB
	PPARG	Safety	-	5/5	Liver	A10BG

Table 13: Expression patterns of the orthologues of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, in relevant rat tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; FC - Frontal cortex; HIP - Hippocampus; all - All brain.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	Gabra1	Safety, Efficacy	Increase	1/7 (1/2)	Brain (CC)	N01AB, N01AX, N03AA, N05BA, N05CD
	Gabra2		-	5/5	Brain	
	Gabra3		Increase	1/6 (1/2)	Brain (CC)	
	Gabra4		Increase	1/8 (1/2)	Brain (CC)	
	Gabra5		**Decrease; Increase**	**1/8 (1/2); 1/8 (1/2)**	Brain (CC); Brain (CC)	
	Gabra6		-	8/8	Brain	
	Gabrb1		-	8/8	Brain	
	Gabrb2		Decrease	2/8 (2/2)	Brain (CC)	
	Gabrb3		Increase	1/8 (1/2)	Brain (CC)	
	Gabrd		-	6/6	Brain	
	Gabre		Decrease	1/8 (1/2)	Brain (CC)	
	Gabrg1		**Decrease**	**1/8 (1/2)**	Brain (CC)	
	Gabrg2		**-**	**1/1**	Brain	
	Gabrg3		Increase	1/8 (1/2)	Brain (CC)	
	Gabrp		-	7/7	Brain	
	Gabrq		-	7/7	Brain	
	Glra1	Safety	Increase	1/8 (1/2)	Brain (CC)	N01AB
	Glrb		-	7/7	Brain	
	Kcnk10	Safety	Decrease	1/7 (1/2)	Brain (CC)	N01AB
	Kcnk18		-	2/2	Brain (CC)	
	Kcnk2		**Increase**	**1/7 (1/2)**	Brain (CC)	
	Kcnk3		Increase	1/8 (1/2)	Brain (CC)	
	Kcnk9		Increase	1/7 (1/2)	Brain (CC)	
	Cacna1a	Safety, Efficacy	-	7/7	Brain	N03AX
	Cacna1b		-	8/8	Brain	
	Cacna1c		-	8/8	Brain	
	Cacna1d		Decrease	1/8 (1/2)	Brain (CC)	
	Cacna1e		-	8/8	Brain	
	Cacna1f		-	7/7	Brain	
	Cacna1g		**Increase**	**1/7 (1/2)**	Brain (CC)	
	Cacna1h		Decrease	1/8 (1/2)	Brain (CC)	

Continued on next page

Table 13 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Cacna1i		-	5/5	Brain	
	Cacna1s		-	6/6	Brain	
	Cacna2d1		**Decrease**	**1/8 (1/2)**	Brain (FC)	
	Cacna2d2		Increase	1/2 (1/2)	Brain (CC)	
	Cacna2d3		-	No probe	Brain	
	Cacna2d4		-	No probe	Brain	
	Cacnb1		-	8/8	Brain	
	Cacnb2		-	8/8	Brain	
	Cacnb3		-	8/8	Brain	
	Cacnb4		-	8/8	Brain	
	Cacng1		-	8/8	Brain	
	Cacng2		-	7/7	Brain	
	Cacng3		-	7/7	Brain	
	Cacng4		-	7/7	Brain	
	Cacng5		-	7/7	Brain	
	Cacng6		-	7/7	Brain	
	Cacng7		-	7/7	Brain	
	Cacng8		-	5/5	Brain	
	Grin1	Control	Decrease	1/8 (1/2)	Brain (CC)	N01AX, N03AX
	Grin2a		**Decrease**	**1/8 (1/2)**	Brain (CC)	
	Grin2b		Decrease	1/8 (1/2)	Brain (CC)	
	Grin2c		-	8/8	Brain (CC)	
	Grin2d		Increase	1/7 (1/2)	Brain (CC)	
	Grin3a		**-**	**No probe**	Brain	
	Grin3b		-	7/7	Brain	
	Htr3a	Safety, Efficacy, Control	-	8/8	Brain	A04AA
	Kcnh2	Control	-	1/1	Heart	C07AA
	Scn10a	Safety, Control	-	8/8	Brain	N03AF
	Scn11a		-	8/8	Brain	
	Scn1a		**-**	**7/7**	Brain	
	Scn2a		-	No probe	Brain	
	Scn3a		**Decrease**	**3/8 (1/4, 2/2)**	Brain (HIP, CC)	
	Scn4a		Increase	1/8 (1/2)	Brain (CC)	
	Scn5a		Decrease	1/8 (1/2)	Brain (CC)	
	Scn7a		Decrease	1/3 (1/2)	Brain (CC)	
	Scn8a		-	7/7	-	
	Scn9a		Decrease	1/8 (1/4)	Brain (HIP)	
Membrane receptor	Drd2	Safety, Efficacy, Control	Increase	1/7 (1/2)	Brain (CC)	A03FA, N05AB

Continued on next page

Table 13 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Htr2a	Safety, Efficacy, Control	**Increase**	**1/8 (1/2)**	Brain (CC)	N06AX
	Htr2c	Safety, Efficacy, Control	Increase	1/8 (1/2)	Brain (CC)	N06AX
	Htr1d	Efficacy	Increase	1/8 (1/2)	Brain (CC)	N02CA, N02CC
	Htr1b	Efficacy	**Increase**	**1/8 (1/2)**	Brain (CC)	N02CC
	Drd3	Control	-	8/8	Brain	N05AH, N05AX
	Drd4	Control	-	8/8	Brain	
	Chrm3	Safety, Efficacy, Control	-; -	4/4; 1/1	Liver; Kidney	A03BA, G04BD, R03BB
Transporter	Slc6a4	Safety, Efficacy, Control	-	8/8	Brain	N06AA, N06AB, N06BA, N06AX
	Slc6a2	Safety, Efficacy, Control	-	8/8	Brain	
	Slc6a3	Safety, Control	-	8/8	Brain	N06BA
	Slc12a3	Efficacy	-	1/1	Heart	C03AX
	Slc18a2	Control	-	8/8	Brain	N06BA
Enzyme	Aldh5a1	Safety	Increase	1/7 (1/2)	Brain (CC)	N03AG
	Csf1r	Safety	?	-	No relevant tissue to analyse	L01XE
	Egfr	Safety, Control	?	-	No relevant tissue to analyse	L01XE
	Abat	Efficacy	-	8/8	Brain (CC)	N03AG
	Mt-Nd1	Efficacy	-	No probe	Liver	A10BA
	Mt-Nd2		-	No probe	Liver	
	Mt-Nd3		-	No probe	Liver	
	Mt-Nd4		-	No probe	Liver	
	Mt-Nd4l		-	No probe	Liver	
	Mt-Nd5		-	No probe	Liver	
	Mt-Nd6		-	No probe	Liver	
	Ndufa1		-	4/4	Liver	
	Ndufa10		-	4/4	Liver	
	Ndufa11		-	4/4	Liver	
	Ndufa12		-	5/5	Liver	
	Ndufa13		-	5/5	Liver	
	Ndufa2		-	5/5	Liver	
	Ndufa3		-	5/5	Liver	

Continued on next page

Table 13 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Ndufa4		-	5/5	Liver	
	Ndufa4l2		-	5/5	Liver	
	Ndufa5		-	5/5	Liver	
	Ndufa6		-	5/5	Liver	
	Ndufa7		-	5/5	Liver	
	Ndufa8		-	5/5	Liver	
	Ndufa9		-	5/5	Liver	
	Ndufab1		-	5/5	Liver	
	Ndufaf1		-	5/5	Liver	
	Ndufaf2		-	No probe	Liver	
	Ndufaf3		-	5/5	Liver	
	Ndufaf4		-	5/5	Liver	
	Ndufb1		-	No probe	Liver	
	Ndufb10		-	5/5	Liver	
	Ndufb11		-	5/5	Liver	
	Ndufb2		-	5/5	Liver	
	Ndufb3		-	No probe	Liver	
	Ndufb4		-	5/5	Liver	
	Ndufb5		-	5/5	Liver	
	Ndufb6		-	5/5	Liver	
	Ndufb7		-	5/5	Liver	
	Ndufb8		-	1/1	Liver	
	Ndufc1		-	4/4	Liver	
	Ndufc2		-	5/5	Liver	
	Ndufs2		-	4/4	Liver	
	Ndufs3		Decrease	1/5	Liver	
	Ndufs4		-	5/5	Liver	
	Ndufs5		-	4/4	Liver	
	Ndufs6		-	4/4	Liver	
	Ndufs7		Decrease	1/5	Liver	
	Ndufs8		Decrease	1/5	Liver	
	Ndufv1		Decrease	1/5	Liver	
	Ndufv2		-	4/4	Liver	
	Ndufv3		-	5/5	Liver	
	Ptgs2	Efficacy, Control	?; -; -; Increase; Decrease	-; 1/1; 1/1; 1/8 (1/2); 1/8 (1/2)	Muscle; Heart; Kidney; Brain (CC); Brain (CC)	M01AB, C01EB, G02CC, M02AA, M01AX
	Erbb2	Control	?	-	No relevant tissue to analyse	L01XE

Continued on next page

Table 13 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Transcription factor	Nr3c1	Safety	Increase; Decrease; -; Increase	1/5; 1/5; -; 1/8;	Liver; Liver; Heart; Brain (CC)	A07EA, C05AA, H02AB, R01AD, R03BA
	Pgr	Safety	?	-	No relevant tissue to analyse	R01AD, R03BA
	Rarb	Safety	Decrease	1/2 (1/2)	Brain (CC)	D10AD
	Nr1h4	Efficacy	Decrease	2/5	Liver	A05AA
	Vdr	Control	Decrease; -	1/5; 8/8	Liver; Brain	A11CC, D05AX, H05BX

Table 14: Expression patterns of orthologues of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, in relevant rat tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; FC - Frontal cortex; HIP - Hippocampus; all - All brain.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	Gabra1	Safety, Efficacy	Increase	1/7 (1/2)	Brain (CC)	N01AB, N01AX, N03AA, N05BA, N05CD
	Gabra2		-	5/5	Brain	
	Gabra3		Increase	1/6 (1/2)	Brain (CC)	
	Gabra4		Increase	1/8 (1/2)	Brain (CC)	
	Gabra5		**Decrease; Increase**	**1/8 (1/2); 1/8 (1/2)**	Brain (CC); Brain (CC)	
	Gabra6		-	8/8	Brain	
	Gabrb1		-	8/8	Brain	
	Gabrb2		Decrease	2/8 (2/2)	Brain (CC)	
	Gabrb3		Increase	1/8 (1/2)	Brain (CC)	
	Gabrd		-	6/6	Brain	
	Gabre		Decrease	1/8 (1/2)	Brain (CC)	
	Gabrg1		**Decrease**	**1/8 (1/2)**	Brain (CC)	
	Gabrg2		**-**	**1/1**	Brain	
	Gabrg3		Increase	1/8 (1/2)	Brain (CC)	
	Gabrp		-	7/7	Brain	
	Gabrq		-	7/7	Brain	
	Cacna1c	Safety, Efficacy	-	1/1	Heart	C08CA, C08DA, C08DB
	Cacna1d		-	1/1	Heart	
	Cacna1f		-	1/1	Heart	
	Cacna1s		-	1/1	Heart	
	Abcc8	Safety	Decrease; Increase	2/5; 1/5	Liver	A10BB
	Kcnj11		Decrease	1/5	Liver	
	Chrna1	Efficacy, Control	Increase; Decrease	1/8 (1/2); 1/8 (1/2)	Brain (FC); Brain (CC)	M03AC, M03AA,
	Chrnb1		Decrease	1/8 (1/2)	Brain (CC)	
	Chrnd		-	8/8	Brain	
	Chrne		Decrease	1/8 (1/2)	Brain (CC)	
	Chrng		-	8/8	Brain	
	Htr3a	Efficacy	-	8/8	Brain	A03FA

Continued on next page

Table 14 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Scnn1a	Efficacy	-; Increase	1/1; 2/3	Heart; Kidney	C03DB
	Hrh1	Safety	Increase	1/8 (1/2)	Brain (CC)	N05BB, R06AB, R06AD, R06AE
	Oprm1	Safety, Efficacy	Decrease; Increase	1/5; 1/8 (1/2)	Liver; Brain (CC)	A07DA, N01AH, N02AB
	Chrm3	Safety	-	4/4	Liver	A03BA, R03BB
	Adrb1	Efficacy	-	1/1	Heart	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA
	Adrb2		-	1/1	Heart	
	Adrb3	Efficacy, Control	-; -	1/1; 3/3	Heart; Kidney	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA, G04BD
	Drd2	Safety, Efficacy	Increase	1/7 (1/2)	Brain (CC)	A03FA, N05AB, C04AE
	Htr2a	Control	**Increase**	**1/8 (1/2)**	Brain (CC)	N06AX
	Htr2c	Control	Increase	1/8 (1/2)	Brain (CC)	N06AX
	Slc6a2	Safety, Efficacy, Control	-	8/8	Brain	N06AA, N06AB, N06AX
	Slc6a4	Safety, Efficacy, Control	-	8/8	Brain	N06AA, N06AB, N06AX
	Slc12a1	Safety, Efficacy	-	1/1	Heart	C03CA
	Atp1a1	Safety	-	1/1	Heart	C01AA
	Atp1a2		-	1/1	Heart	
	Atp1a3		-	1/1	Heart	
	Atp1a4		-	1/1	Heart	
	Slc29a1	Efficacy	-	1/1	Heart	B01AC

Continued on next page

Table 14 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Enzyme	Ptgs1	Safety	Increase; -	1/5; 8/8	Liver; Brain	A01AD, N02BA, M01AB
	Ptgs2		Decrease; Increase; Decrease	1/5; 1/8 (1/2); 1/8 (1/2)	Liver; Brain (CC); Brain (CC)	
	Pde3a	Safety	?	-	No relevant tissue to analyse	
	Pde3b		?	-	No relevant tissue to analyse	R03DA
	Ache	Efficacy	?	-	No relevant tissue to analyse	Unclassified
	Pde5a	Efficacy	-	1/1	Heart	B01AC
	Ace	Control	?	No probe	Heart	C09AA
	F2	Control	-	1/1	Heart	B01AE
	Hmgcr	Control	Decrease	2/5	Liver	C10AA
Transcription factor	Nr3c1	Safety	Increase	1/8 (1/2)	Brain (CC)	D07AA, H02BX, D10AA, H02AB
	Pparg	Safety	Increase; Decrease	1/5; 1/5	Liver	A10BG

Table 15: Expression patterns of the orthologues of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in paediatric populations when compared with adult populations, in relevant mouse tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; PF - Prefrontal cortex; FC - Frontal cortex; HIP - Hippocampus; HYP - Hypothalamus; all - All brain.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	Gabra1	Safety, Efficacy	Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	N01AB, N01AX, N03AA, N05BA, N05CD
	Gabra2		Increase	3/13 (1/1, 1/3, 1/2)	Brain (HYP, HIP, all)	
	Gabra3		Increase; Decrease	2/12 (1/1, 1/2); 1/12 (1/3)	Brain (HYP, all); Brain (HIP)	
	Gabra4		Increase	2/10 (1/1, 1/2)	Brain (HYP, all)	
	Gabra5		**Increase; Decrease**	**2/7 (1/1, 1/2); 1/7 (1/3)**	Brain (HYP, all); Brain (HIP)	
	Gabra6		Increase	1/12 (1/2)	Brain (all)	
	Gabrb1		Increase; Decrease	2/12 (1/1, 1/2); 1/12 (1/3)	Brain (HYP, all); Brain (HIP)	
	Gabrb2		Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	
	Gabrb3		Increase	2/15 (1/1, 1/2)	Brain (HYP, all)	
	Gabrd		Increase; Decrease	1/12 (1/3); 1/12 (1/2)	Brain (HIP, all)	
	Gabre		Decrease	1/10 (1/2)	Brain (all)	
	Gabrg1		**Increase**	**1/12 (1/1)**	Brain (HYP)	
	Gabrg2		**Increase; Decrease**	**1/15 (1/1); 1/15 (1/2)**	Brain (HYP); Brain (all)	
	Gabrg3		Decrease	1/15 (1/3)	Brain (HIP)	
	Gabrp		Decrease	1/13 (1/2)	Brain (all)	
	Gabrq		Decrease	2/10 (1/3, 1/2)	Brain (HIP, all)	
	Glra1	Safety	-	12/12	Brain	N01AB
	Glrb		Increase; Decrease	2/12 (1/1, 1/2); 1/12 (1/3)	Brain (HYP, all); Brain (HIP)	
	Kcnk10	Safety	Decrease	1/10 (1/2)	Brain (all)	N01AB
	Kcnk18		Decrease	1/1 (1/3)	Brain (HIP)	
	Kcnk2		**Decrease, Increase**	**1/15 (1/3); 1/15 (1/1)**	Brain (HIP); Brain (HYP)	
	Kcnk3		-	12/12	Brain	
	Kcnk9		?	No probe	Brain	
	Cacna1a	Safety, Efficacy	-	15/15	Brain	N03AX
	Cacna1b		Increase; Decrease	1/15 (1/1); 1/15 (1/1)	Brain (HYP, brainstem)	
	Cacna1c		-	12/12	Brain	
	Cacna1d		-	10/10	Brain	

Continued on next page

Table 15 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Cacna1e		-	12/12	Brain	
	Cacna1f		Decrease	1/10 (1/2)	Brain (all)	
	Cacna1g		**Decrease; Increase**	**2/12 (1/3, 1/1); 1/12 (1/1)**	Brain (HIP, brainstem); Brain (HYP)	
	Cacna1h		Increase	2/15 (1/3, 1/1)	Brain (HYP, cortex)	
	Cacna1i		-	3/3	Brain	
	Cacna1s		-	15/15	Brain	
	Cacna2d1		**Increase**	**2/12 (1/1, 1/2)**	Brain (HYP, all)	
	Cacna2d2		Increase	1/10 (1/1)	Brain (HYP)	
	Cacna2d3		-	12/12	Brain	
	Cacna2d4		-	7/7	Brain	
	Cacnb1		-	10/10	Brain	
	Cacnb2		-	15/15	Brain	
	Cacnb3		-	15/15	Brain	
	Cacnb4		-	10/10	Brain	
	Cacng1		-	15/15	Brain	
	Cacng2		-	11/11	Brain	
	Cacng3		-	10/10	Brain	
	Cacng4		-	10/10	Brain	
	Cacng5		-	13/13	Brain	
	Cacng6		-	10/10	Brain	
	Cacng7		-	10/10	Brain	
	Cacng8		-	10/10	Brain	
	Grin1	Control	Increase; Decrease	1/12 (1/2); 1/12 (1/3)	Brain (all); Brain (HIP)	N01AX, N03AX
	Grin2a		**Increase; Decrease**	**1/12 (1/2); 1/12 (1/3)**	Brain (all); Brain (HIP)	
	Grin2b		-	12/12	Brain	
	Grin2c		Increase	1/12 (1/2)	Brain (all)	
	Grin2d		Decrease	2/12 (1/3, 1/2)	Brain (HIP, all)	
	Grin3a		**Increase**	**1/7 (1/1)**	Brain (HYP)	
	Grin3b		Increase; Decrease	1/12 (1/2); 1/12 (1/3)	Brain (all); Brain (HIP)	
	Htr3a	Safety, Efficacy, Control	Decrease	2/12 (1/3, 1/2)	Brain (HIP, all)	A04AA
	Kcnh2	Control	-	2/2	Heart	C07AA
	Scn10a	Safety, Control	Decrease	1/3 (1/3)	Brain (HIP)	N03AF
	Scn11a		Decrease	1/10 (1/2)	Brain (all)	
	Scn1a		**Increase**	**3/12 (1/1, 2/2)**	Brain (all)	
	Scn2a		Increase; Decrease	2/10 (1/1, 1/2); 1/10 (1/3)	Brain (HYP, all); Brain (HIP)	
	Scn3a		**Decrease; Increase**	**1/12 (1/2); 1/12 (1/1)**	Brain (all); Brain (HYP)	
	Scn4a		Decrease	1/10 (1/3)	Brain (HIP)	

Continued on next page

Table 15 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Scn5a		Increase	1/13 (1/3)	Brain (HIP)	
	Scn7a		-	12/12	Brain	
	Scn8a		Increase; Decrease	2/12 (1/1, 1/2); 2/12 (1/3, 1/1)	Brain (HYP, all); Brain (HIP, brainstem)	
	Scn9a		Decrease	1/12 (1/2)	Brain (all)	
Membrane receptor	Drd2	Safety, Efficacy, Control	Decrease	1/12 (1/3)	Brain (HIP)	A03FA, N05AB
	Htr2a	Safety, Efficacy, Control	**-**	**1/1**	Brain	N06AX
	Htr2c	Safety, Efficacy, Control	Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	N06AX
	Htr1d	Efficacy	Decrease	1/12 (1/3)	Brain (HIP)	N02CA, N02CC
	Htr1b	Efficacy	**Decrease**	**1/12 (1/2)**	Brain (all)	N02CC
	Drd3	Control	Decrease	1/12 (1/2)	Brain (all)	N05AH, N05AX
	Drd4	Control	Increase	1/12 (1/3)	Brain (HIP)	
	Chrm3	Safety, Efficacy, Control	-; -; -	6/6; 2/2; 3/3	Liver; Kidney; Lungs	A03BA, G04BD, R03BB
Transporter	Slc6a4	Safety, Efficacy, Control	Decrease	1/12 (1/2)	Brain (all)	N06AA, N06AB, N06BA, N06AX
	Slc6a2	Safety, Efficacy, Control	Decrease	2/12 (1/3, 1/2)	Brain (HIP, all)	
	Slc6a3	Safety, Control	Increase	1/12 (1/3)	Brain (HIP)	N06BA
	Slc12a3	Efficacy	-	2/2	Heart	C03AX
	Slc18a2	Control	Decrease	1/7 (1/2)	Brain (all)	N06BA
Enzyme	Aldh5a1	Safety	Increase; Decrease	1/7 (1/1); 1/7 (1/3)	Brain (HYP); Brain (HIP)	N03AG
	Csf1r	Safety	?	-	No relevant tissue to analyse	L01XE
	Egfr	Safety, Control	?	-	No relevant tissue to analyse	L01XE
	Abat	Efficacy	Increase; Decrease	2/13 (1/1, 1/2); 1/13 (1/3)	Brain (HYP, all); Brain (HIP)	N03AG
	Mt-Nd1	Efficacy	-	No probe	Liver	A10BA
	Mt-Nd2		-	No probe	Liver	
	Mt-Nd3		-	No probe	Liver	
	Mt-Nd4		-	No probe	Liver	
	Mt-Nd4l		-	No probe	Liver	
	Mt-Nd5		-	No probe	Liver	
	Mt-Nd6		-	No probe	Liver	

Continued on next page

Table 15 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Ndufa1		-	8/8	Liver	
	Ndufa10		-	8/8	Liver	
	Ndufa11		-	8/8	Liver	
	Ndufa12		-	8/8	Liver	
	Ndufa13		-	8/8	Liver	
	Ndufa2		-	8/8	Liver	
	Ndufa3		-	8/8	Liver	
	Ndufa4		-	8/8	Liver	
	Ndufa4l2		-	7/7	Liver	
	Ndufa5		-	8/8	Liver	
	Ndufa6		-	8/8	Liver	
	Ndufa7		-	8/8	Liver	
	Ndufa8		-	8/8	Liver	
	Ndufa9		-	8/8	Liver	
	Ndufab1		-	8/8	Liver	
	Ndufaf1		-	8/8	Liver	
	Ndufaf2		-	7/7	Liver	
	Ndufaf3		-	8/8	Liver	
	Ndufaf4		-	8/8	Liver	
	Ndufb1		-	No probe	Liver	
	Ndufb10		-	8/8	Liver	
	Ndufb11		-	8/8	Liver	
	Ndufb2		-	8/8	Liver	
	Ndufb3		-	8/8	Liver	
	Ndufb4		-	8/8	Liver	
	Ndufb5		-	8/8	Liver	
	Ndufb6		-	8/8	Liver	
	Ndufb7		-	8/8	Liver	
	Ndufb8		-	8/8	Liver	
	Ndufc1		-	8/8	Liver	
	Ndufc2		-	8/8	Liver	
	Ndufs2		-	8/8	Liver	
	Ndufs3		-	8/8	Liver	
	Ndufs4		-	8/8	Liver	
	Ndufs5		-	8/8	Liver	
	Ndufs6		-	8/8	Liver	
	Ndufs7		-	8/8	Liver	
	Ndufs8		-	8/8	Liver	
	Ndufv1		-	8/8	Liver	
	Ndufv2		-	8/8	Liver	
	Ndufv3		-	8/8	Liver	

Continued on next page

Table 15 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Ptgs2	Efficacy, Control	?; -; -; Increase	-; 2/2; 2/2; 2/15 (1/3, 1/2)	Muscle; Heart; Kidney; Brain (HIP, all)	M01AB, C01EB, G02CC, M02AA, M01AX
	Erbb2	Control	?	-	No relevant tissue to analyse	L01XE
Transcription factor	Nr3c1	Safety	-; -; Increase; -	8/8; 2/2; 1/15 (1/2); 3/3	Liver; Heart; Brain (all); Lungs	A07EA, C05AA, H02AB, R01AD, R03BA
	Pgr	Safety	-	3/3	Lungs	R01AD, R03BA
	Rarb	Safety	-	13/13	Brain	D10AD
	Nr1h4	Efficacy	-	8/8	Liver	A05AA
	Vdr	Control	-; -; -	8/8; 15/15; 3/3	Liver; Brain; Kidney	A11CC, D05AX, H05BX

Table 16: Expression patterns of orthologues of top drug targets associated with drugs, which have shown either severe and unseen side effects (DDR - safety), a lack of efficacy (DDR - efficacy), or no difference in response (DDR - control) in geriatric populations when compared with adult populations, in relevant mouse tissues. Strong signals in human tissues are highlighted in bold. CC - Cerebral cortex; PF - Prefrontal cortex; FC - Frontal cortex; HIP - Hippocampus; HYP - Hypothalamus; all - All brain.

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
Ion channel	Gabra1	Safety, Efficacy	Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	N01AB, N01AX, N03AA, N05BA, N05CD
	Gabra2		Increase	3/13 (1/1, 1/3, 1/2)	Brain (HYP, HIP, all)	
	Gabra3		Increase; Decrease	2/12 (1/1, 1/2); 1/12 (1/3)	Brain (HYP, all); Brain (HIP)	
	Gabra4		Increase	2/10 (1/1, 1/2)	Brain (HYP, all)	
	Gabra5		**Increase; Decrease**	**2/7 (1/1, 1/2); 1/7 (1/3)**	Brain (HYP, all); Brain (HIP)	
	Gabra6		Increase	1/12 (1/2)	Brain (all)	
	Gabrb1		Increase; Decrease	2/12 (1/1, 1/2); 1/12 (1/3)	Brain (HYP, all); Brain (HIP)	
	Gabrb2		Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	
	Gabrb3		Increase	2/15 (1/1, 1/2)	Brain (HYP, all)	
	Gabrd		Increase; Decrease	1/12 (1/3); 1/12 (1/2)	Brain (HIP, all)	
	Gabre		Decrease	1/10 (1/2)	Brain (all)	
	Gabrg1		**Increase**	**1/12 (1/1)**	Brain (HYP)	
	Gabrg2		**Increase; Decrease**	**1/15 (1/1); 1/15 (1/2)**	Brain (HYP); Brain (all)	
	Gabrg3		Decrease	1/15 (1/3)	Brain (HIP)	
	Gabrp		Decrease	1/13 (1/2)	Brain (all)	
	Gabrq		Decrease	2/10 (1/3, 1/2)	Brain (HIP, all)	
	Cacna1c	Safety, Efficacy	Increase	1/2	Heart	C08CA, C08DA, C08DB
	Cacna1d		-	1/1	Heart	
	Cacna1f		-	1/1	Heart	
	Cacna1s		-	2/2	Heart	
	Abcc8	Safety	-	8/8	Liver	A10BB
	Kcnj11		-	8/8	Liver	
	Chrna1	Efficacy, Control	?	-	No relevant tissue to analyse	M03AC, M03AA,
	Chrnb1		?	-	No relevant tissue to analyse	
	Chrnd		?	-	No relevant tissue to analyse	

Continued on next page

Table 16 – *Continued from previous page*

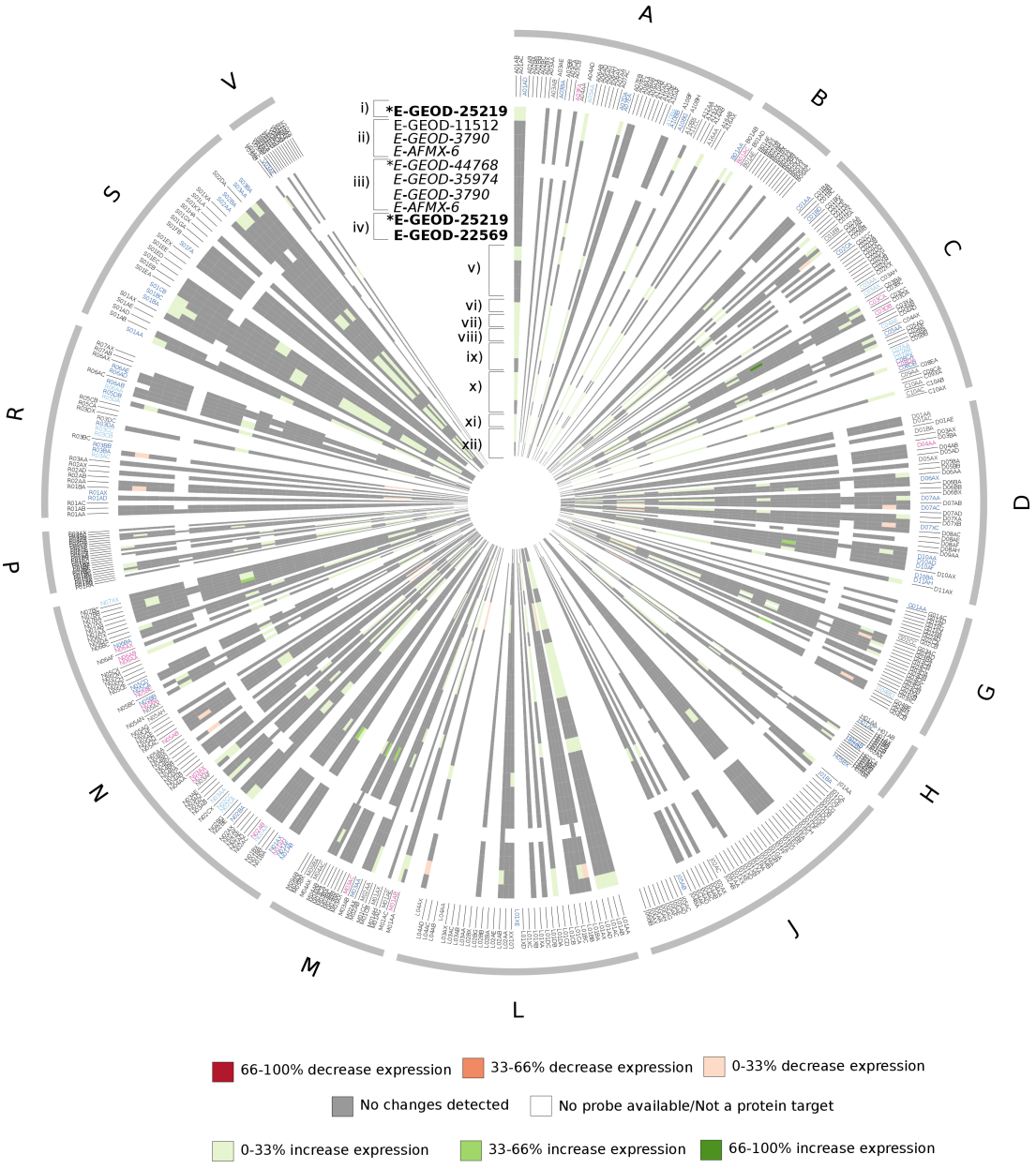
Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Chrne		?	-	No relevant tissue to analyse	
	Chrng		?	-	No relevant tissue to analyse	
	Htr3a	Efficacy	Decrease	2/12 (1/3, 1/2)	Brain (HIP, all)	A03FA
	Scnn1a	Efficacy	-; Increase	1/1; 1/2	Heart; Kidney	C03DB
	Hrh1	Safety	Increase; -	1/12 (1/2); 3/3	Brain (all); Lungs	N05BB, R06AB, R06AD, R06AE
	Oprm1	Safety, Efficacy	-; Decrease	8/8; 1/12 (1/2)	Liver; Brain (all)	A07DA, N01AH, N02AB
	Chrm3	Safety	-; -	6/6; 3/3	Liver; Lungs	A03BA, R03BB
	Adrb1	Efficacy	Decrease; -	1/2; 3/3	Heart; Lungs	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA
	Adrb2		-; Increase	2/2; 1/3	Heart; Lungs	
	Adrb3	Efficacy, Control	-; Increase; -	2/2; 1/3; 2/2	Heart; Lungs; Kidney	C07AB, C07AG, C07AA, R03AC, R03CB, R03A, C01CA, G04BD
	Drd2	Safety, Efficacy	Decrease	1/12 (1/3)	Brain (HIP)	A03FA, N05AB, C04AE
	Htr2a	Control	**-**	**1/1**	Brain	N06AX
	Htr2c	Control	Increase	2/12 (1/1, 1/2)	Brain (HYP, all)	N06AX
	Slc6a2	Safety, Efficacy, Control	Decrease	2/12 (1/3, 1/2)	Brain (HIP, all)	N06AA, N06AB, N06AX
	Slc6a4	Safety, Efficacy, Control	Decrease	1/12 (1/2)	Brain (all)	N06AA, N06AB, N06AX
	Slc12a1	Safety, Efficacy	-	2/2	Heart	C03CA
	Atp1a1	Safety	-	2/2	Heart	C01AA
	Atp1a2		-	2/2	Heart	
	Atp1a3		-	1/1	Heart	

Continued on next page

Table 16 – *Continued from previous page*

Family class	Gene Symbol	DDR category	Expression pattern	Frequency	Tissue	ATC code
	Atp1a4		-	1/1	Heart	
	Slc29a1	Efficacy	-	2/2	Heart	B01AC
Enzyme	Ptgs1	Safety	Increase; Increase	1/8; 1/15 (1/2)	Liver; Brain (all)	A01AD, N02BA, M01AB
	Ptgs2		-; Increase	8/8; 2/15 (1/3, 1/2)	Liver; Brain (HIP, all)	
	Pde3a	Safety	-	3/3	Lungs	
	Pde3b		-	3/3	Lungs	R03DA
	Ache	Efficacy	?	-	No relevant tissue to analyse	Unclassified
	Pde5a	Efficacy	Decrease	1/1	Heart	B01AC
	Ace	Control	Decrease	1/1	Heart	C09AA
	F2	Control	-	2/2	Heart	B01AE
	Hmgcr	Control	Increase	1/8	Liver	C10AA
Transcription factor	Nr3c1	Safety	Increase	1/15 (1/2)	Brain (all)	D07AA, H02BX, D10AA, H02AB
	Pparg	Safety	-	8/8	Liver	A10BG

(a)


A
B
C
D
G
H
J
L
M
N
P
R
S
V
i) *E-GEOD-25219
ii) E-GEOD-11512
E-GEOD-3790
E-AFMX-6
iii) *E-GEOD-44768
E-GEOD-35974
E-GEOD-3790
E-AFMX-6
iv) *E-GEOD-25219
E-GEOD-22569
v)
vi)
vii)
viii)
ix)
x)
xi)
xii)
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

(b)

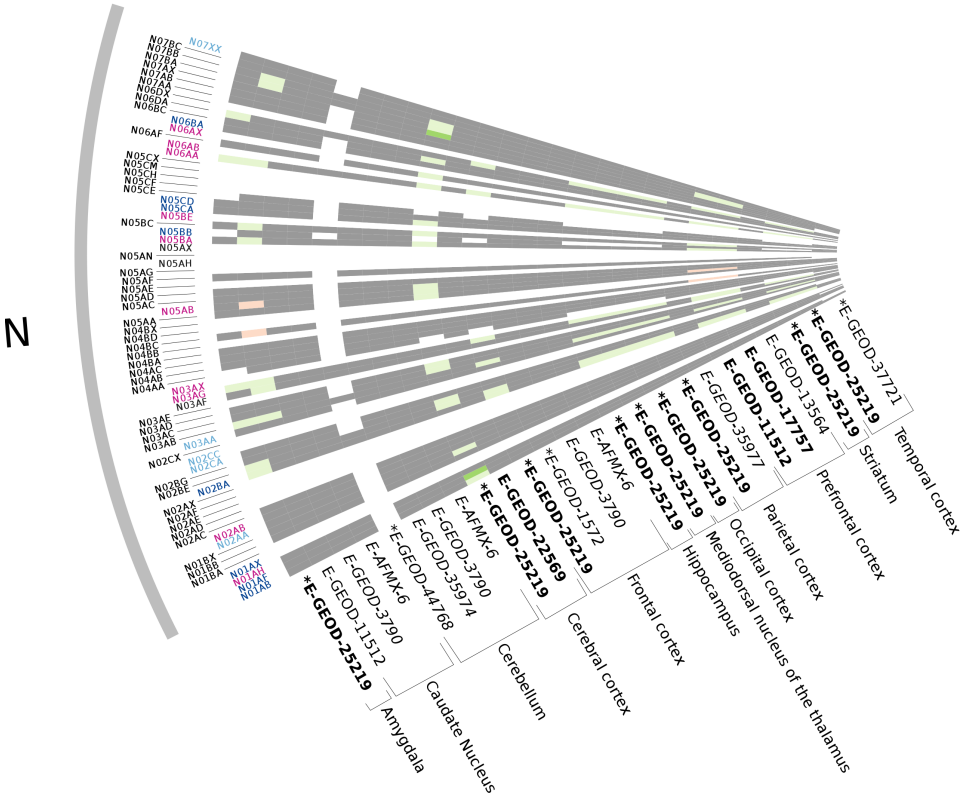
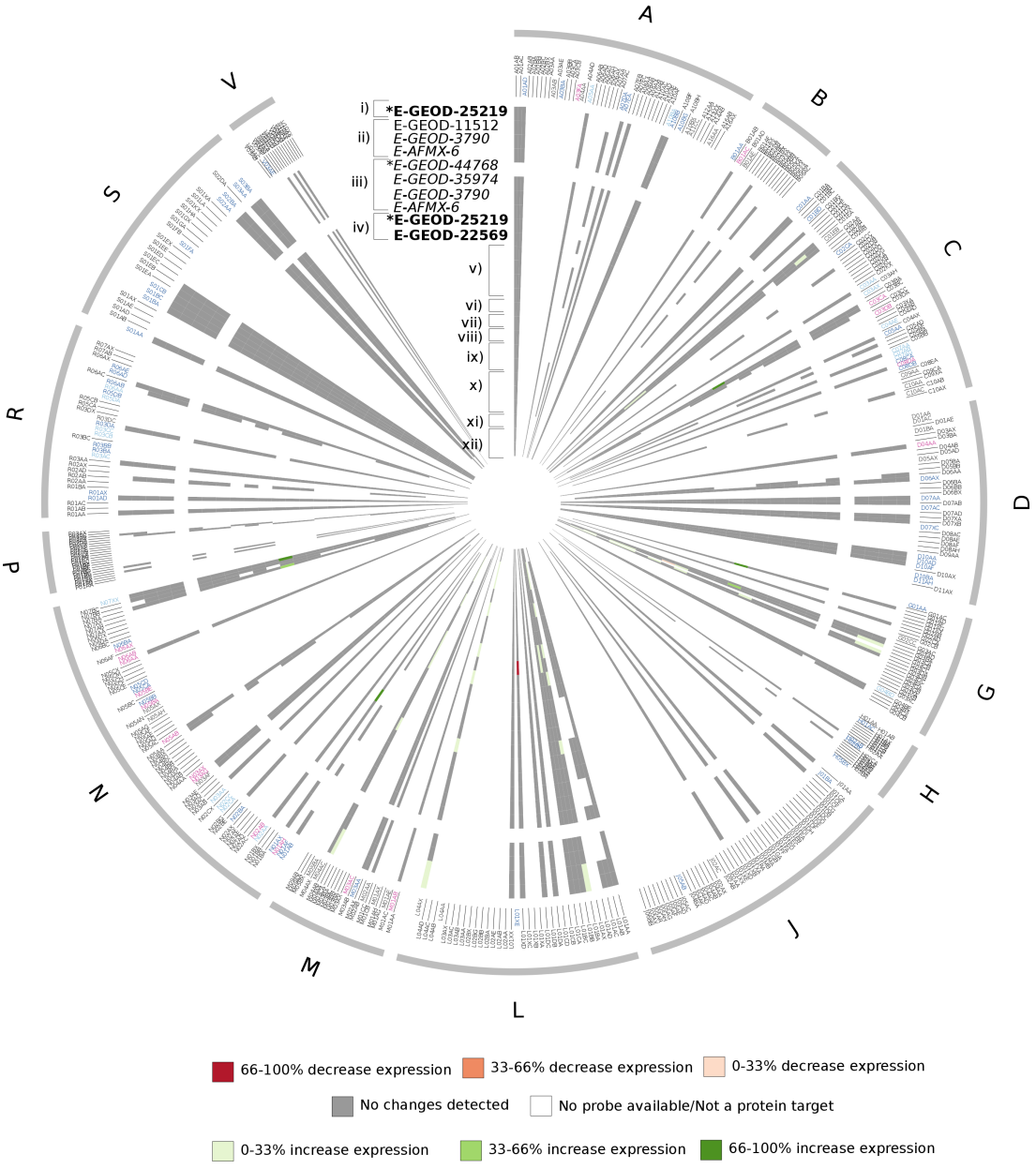


Figure 1: Overview of gene expression changes as a function of age of metabolic enzymes in (a) human brain per ATC level 4 in (b) the nervous system category (ATC code 'N'). The outer shell corresponds to the level 1 of the ATC code, scaled to the number of drugs in that category. The inner semi-rings represent the level 4 of the ATC code, coloured by the gene expression variations detected in each individual experiment, and reflect the number of drug targets associated with that ATC level 4, which have decrease its expression with age (depicted in red), or have increased with age (depicted in green). The labels of semi-rings code are coloured in dark blue, if at least one drug in that class has shown either severe and unseen side effects in paediatric or geriatric populations when compared with adult populations (DDR - safety), light blue if at least one drug in that class has shown a lack of efficacy in paediatric or geriatric populations when compared with adult populations (DDR - efficacy), pink if at least one drug has shown both reported safety and efficacy issues when transferred from one aged population to another, and grey if the drug can be safely and effectively transferred from one aged population to another (DDR - control). The experiment accession is depicted in bold if the developmental stages covered by the samples range from childhood to elderhood, italic if from adulthood to elderhood, and normal if it only covers from childhood to adulthood. Asterisks (*) indicate expression datasets without raw data. i) Amygdala, ii) Caudate nucleus, iii) Cerebellum, iv) Cerebral cortex, v) Frontal cortex, vi) Hippocampus, vii) Mediodorsal nucleus of the thalamus, viii) Occipital cortex, ix) Parietal cortex, x) Prefrontal cortex, xi) Striatum, xii) Temporal cortex.

(a)


A
B
C
D
G
H
J
L
M
N
P
R
S
V
i) *E-GEOD-25219
ii) E-GEOD-11512
E-GEOD-3790
E-AFMX-6
iii) *E-GEOD-44768
E-GEOD-35974
E-GEOD-3790
E-AFMX-6
iv) *E-GEOD-25219
E-GEOD-22569
v)
vi)
vii)
viii)
ix)
x)
xi)
xii)
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

(b)

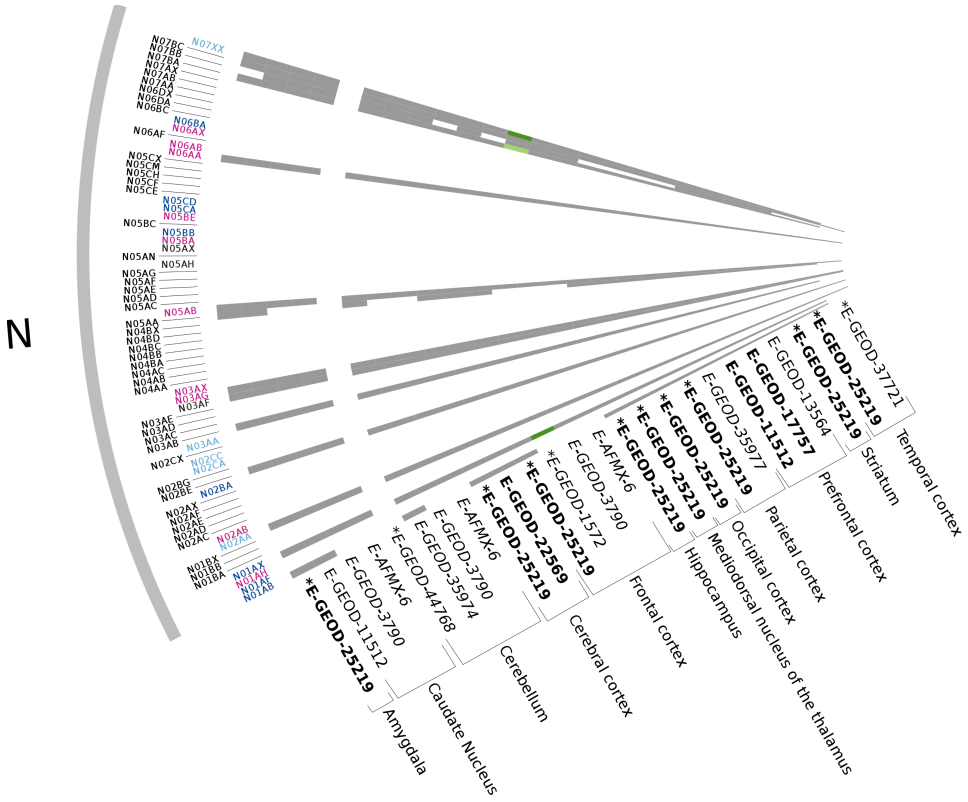


Figure 2: Overview of gene expression changes as a function of age of transporters in human brain per ATC level 4 (a) in the nervous system category (ATC code 'N') (b). The outer shell corresponds to the level 1 of the ATC code, scaled to the number of drugs in that category. See caption from figure 1 for further details. i) Amygdala, ii) Caudate nucleus, iii) Cerebellum, iv) Cerebral cortex, v) Frontal cortex, vi) Hippocampus, vii) Mediodorsal nucleus of the thalamus, viii) Occipital cortex, ix) Parietal cortex, x) Prefrontal cortex, xi) Striatum, xii) Temporal cortex.

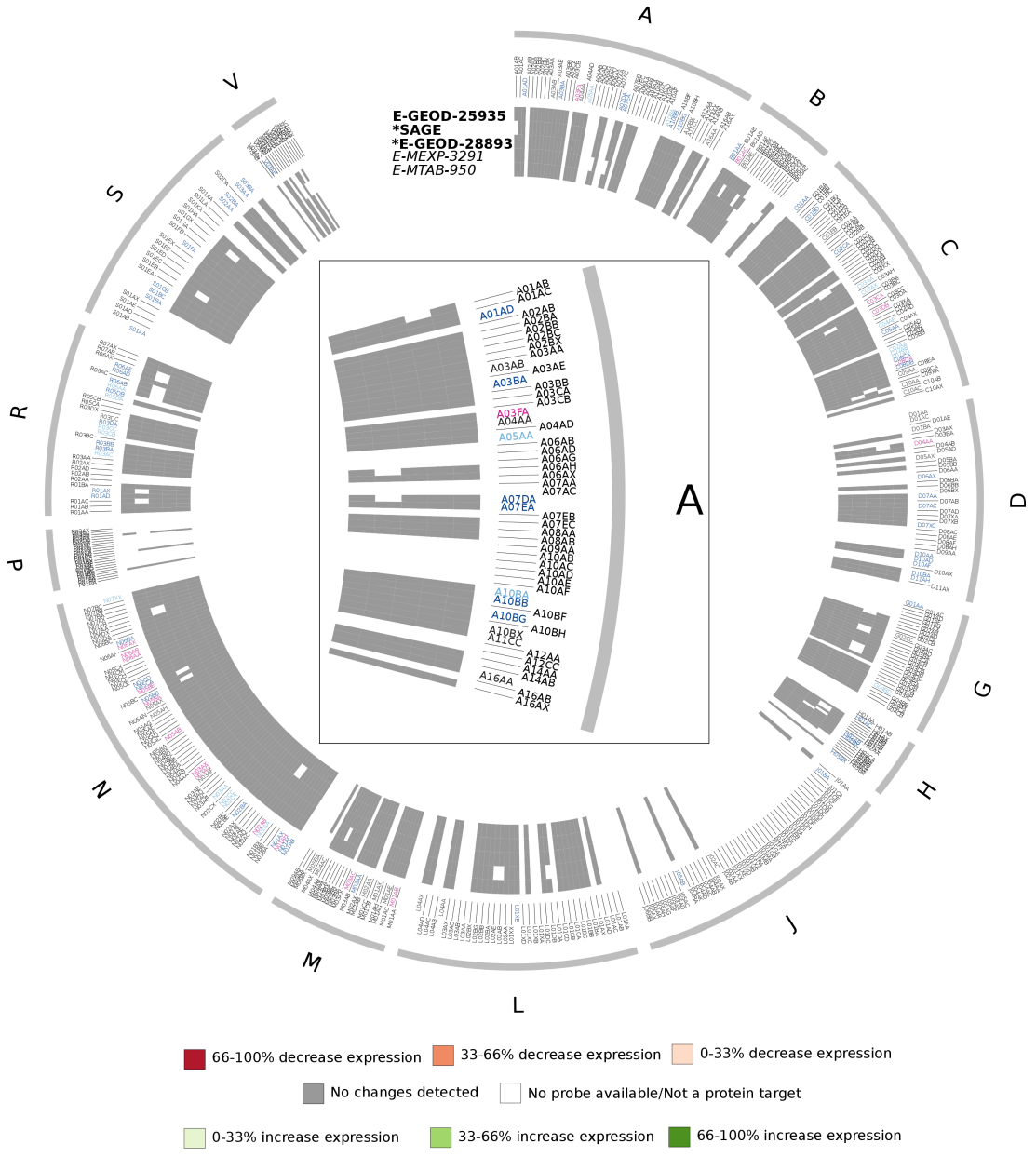


Figure 3: Overview of gene expression changes as a function of age of drug targets in human liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

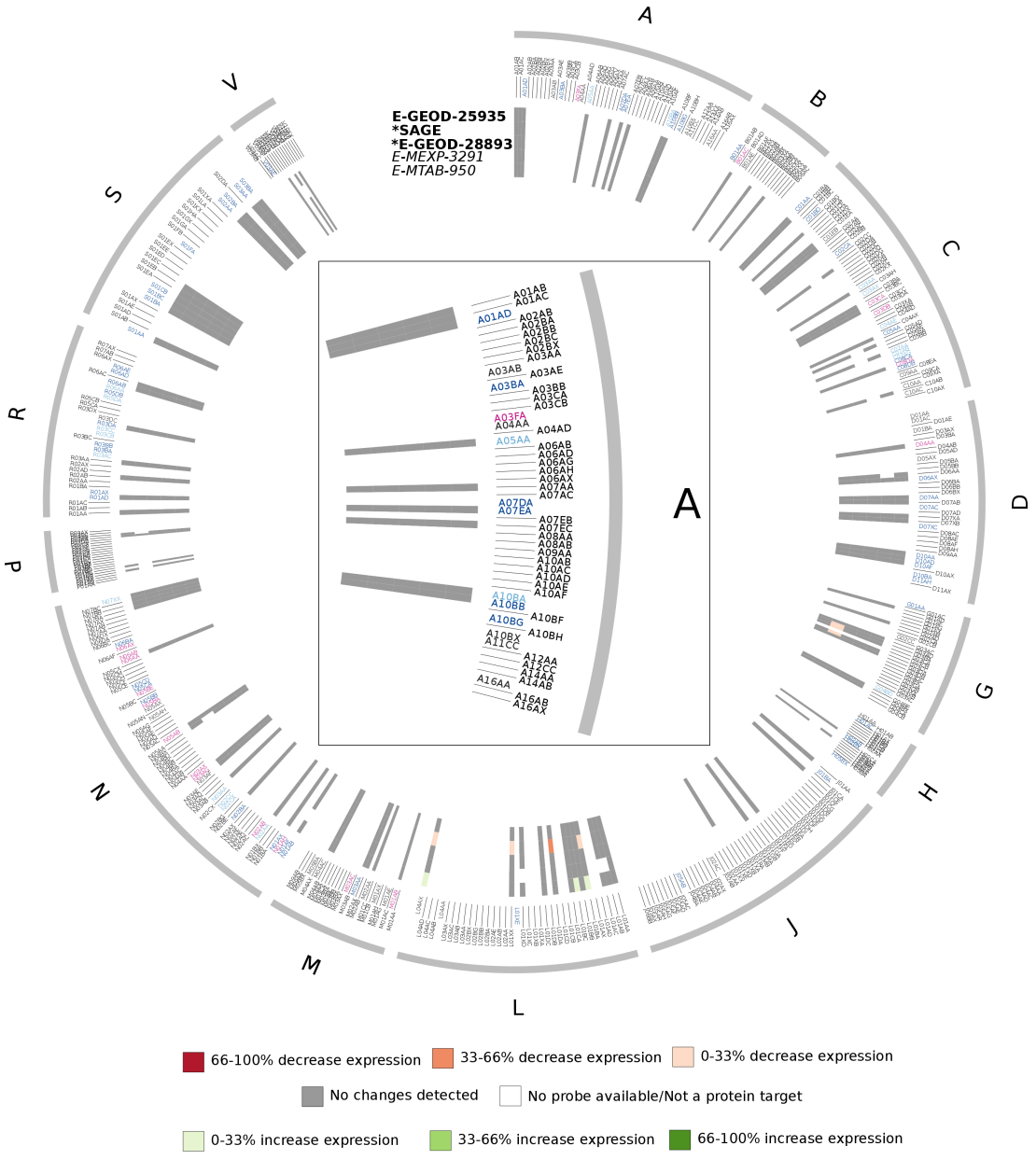


Figure 4: Overview of gene expression changes as a function of age of transporters in human liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

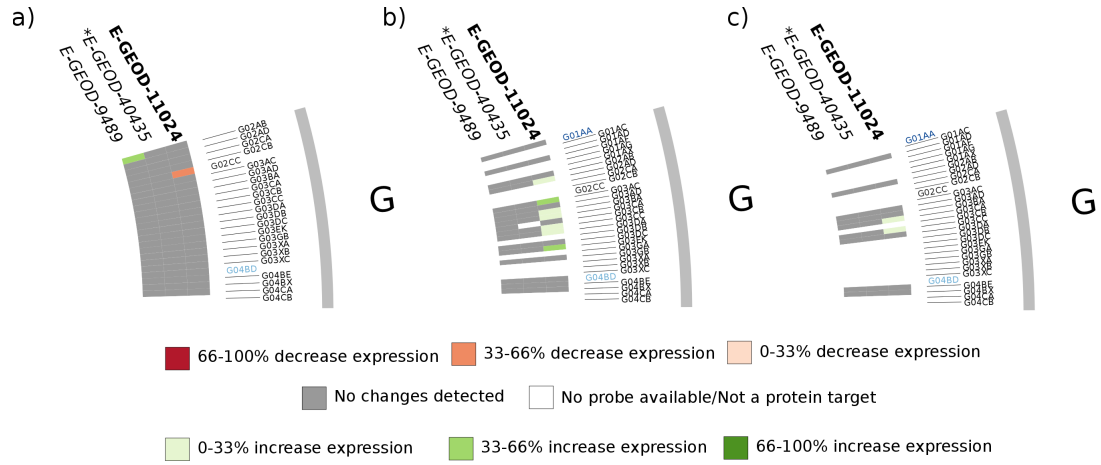


Figure 5: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in human kidney per ATC level 4 in the genito-urinary system category (ATC code 'G'). See caption from figure 1 for further details.

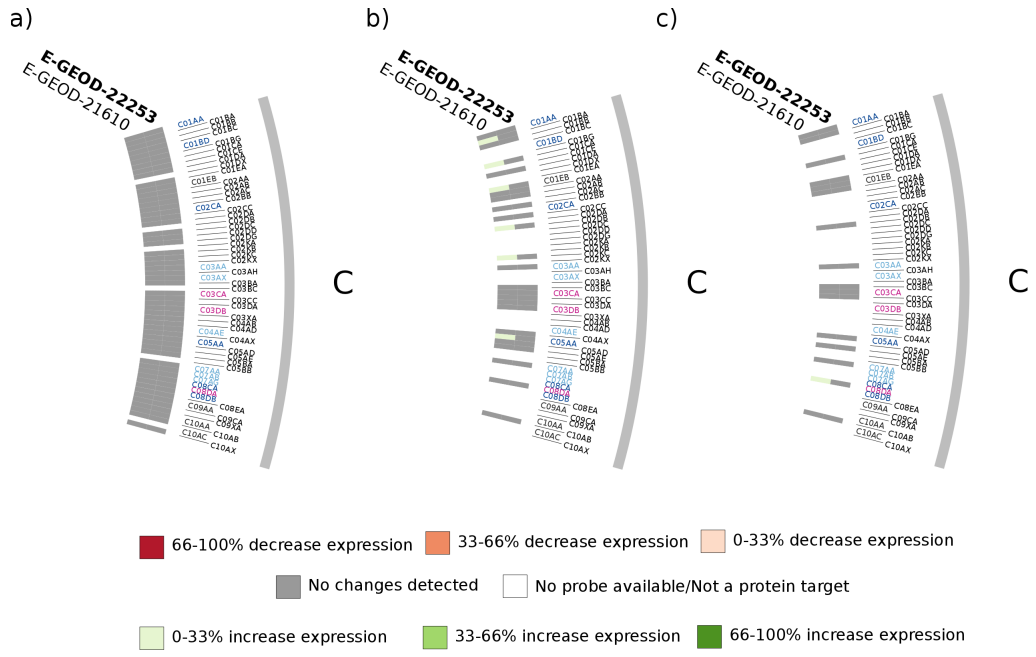


Figure 6: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in human heart per ATC level 4 in the cardiovascular system category (ATC code 'C'). See caption from figure 1 for further details.

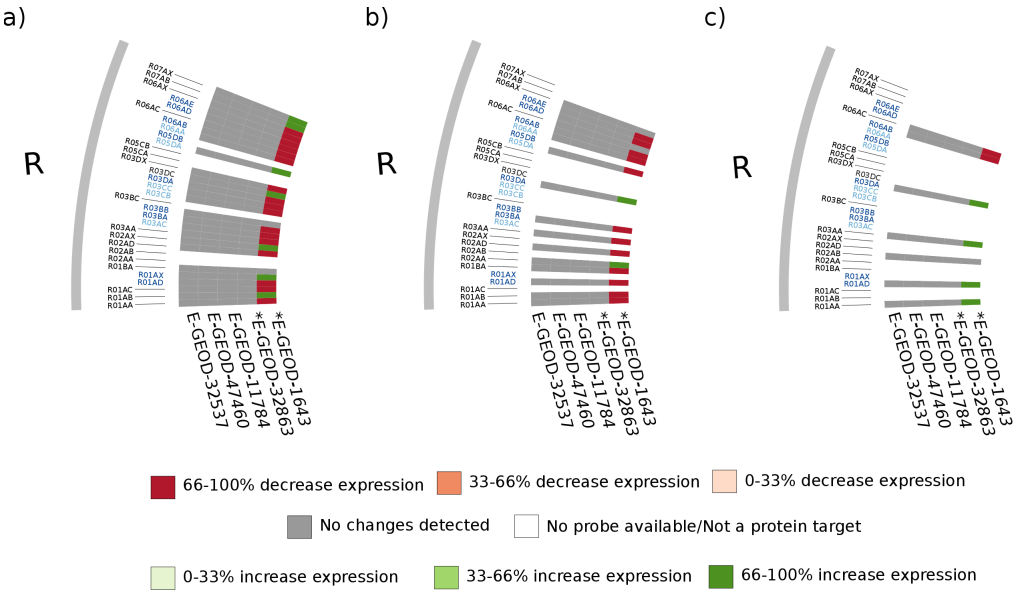


Figure 7: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in human lungs per ATC level 4 in the respiratory system category (ATC code 'R'). See caption from figure 1 for further details.

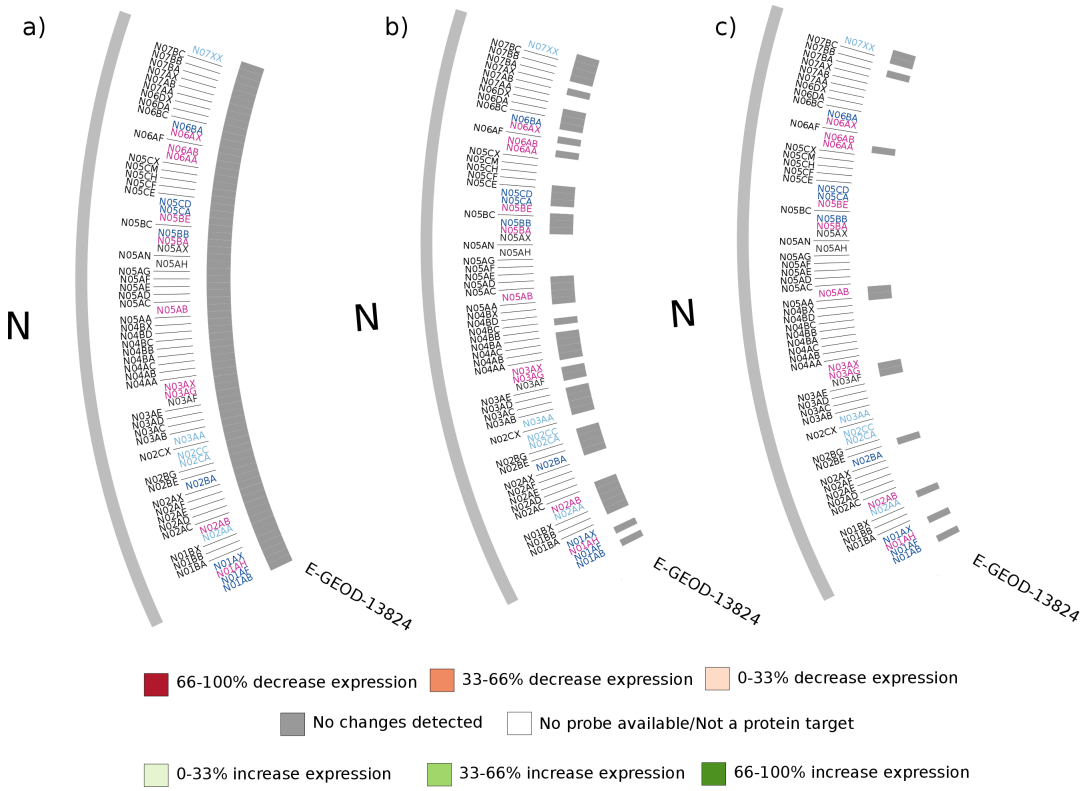


Figure 8: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in macaque hippocampus per ATC level 4 in the nervous system category (ATC code 'N'). See caption from figure 1 for further details.

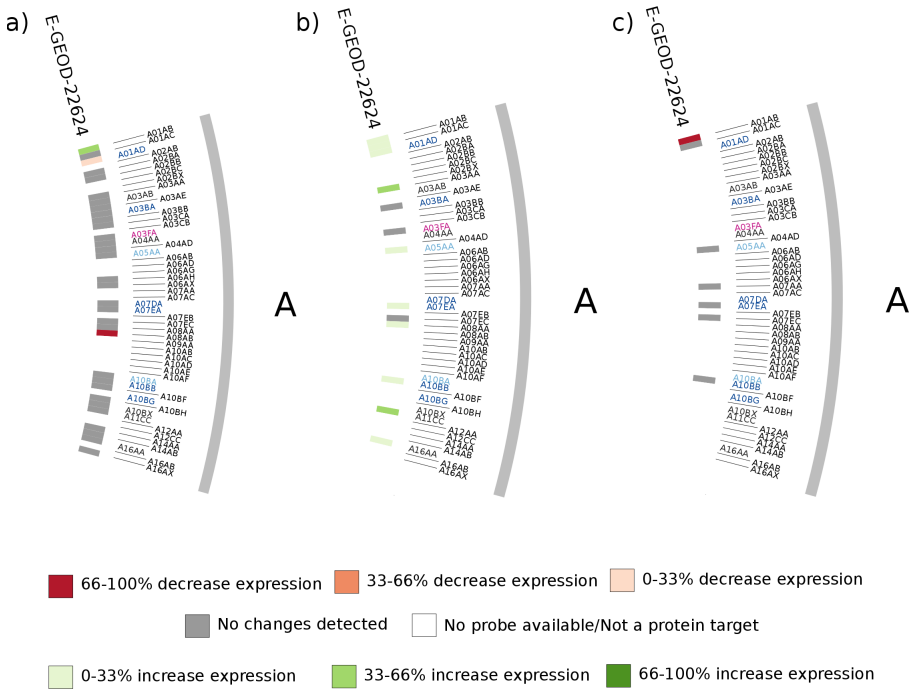


Figure 9: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in macaque liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

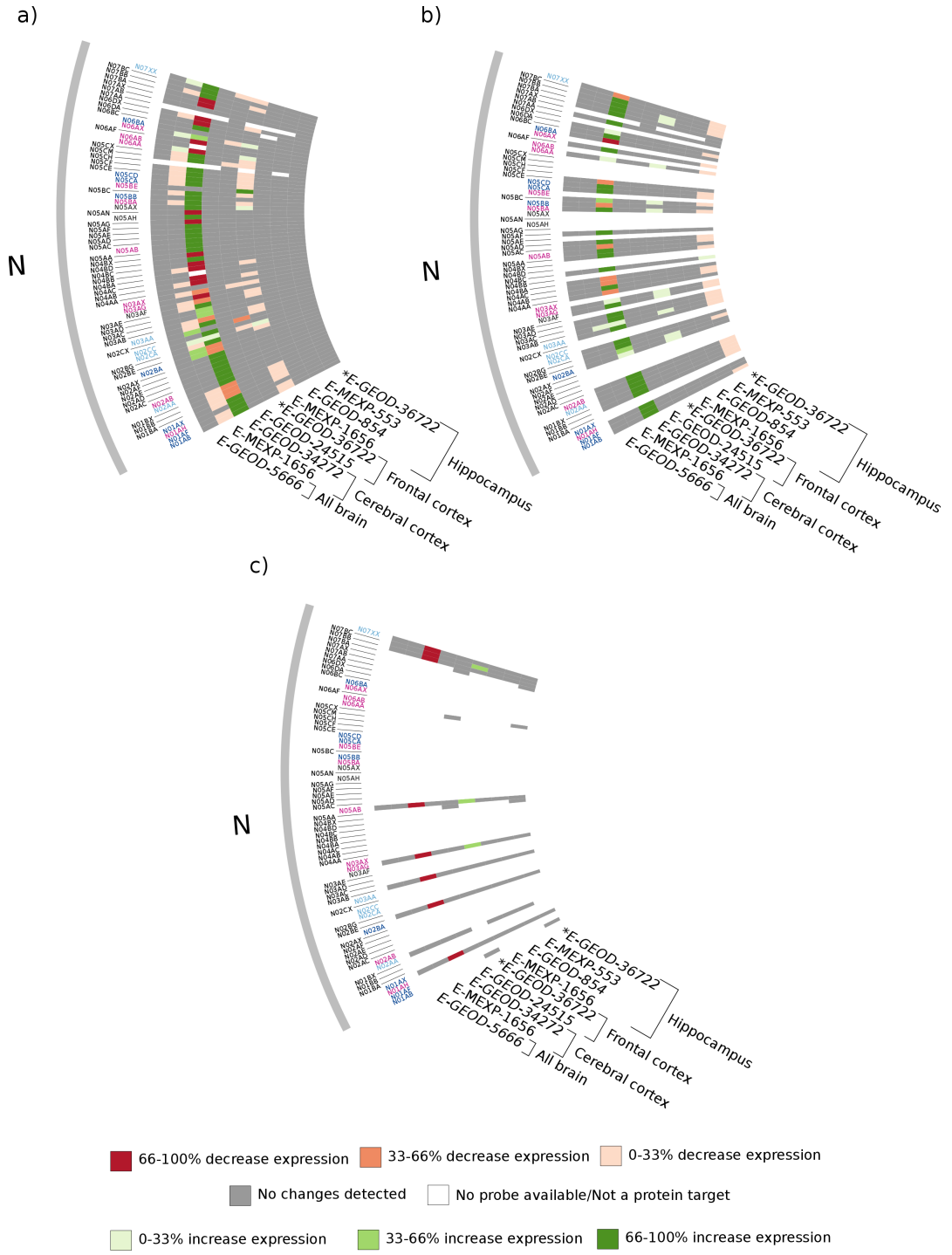


Figure 10: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in rat brain per ATC level 4 in the nervous system category category (ATC code 'A'). See caption from figure 1 for further details.

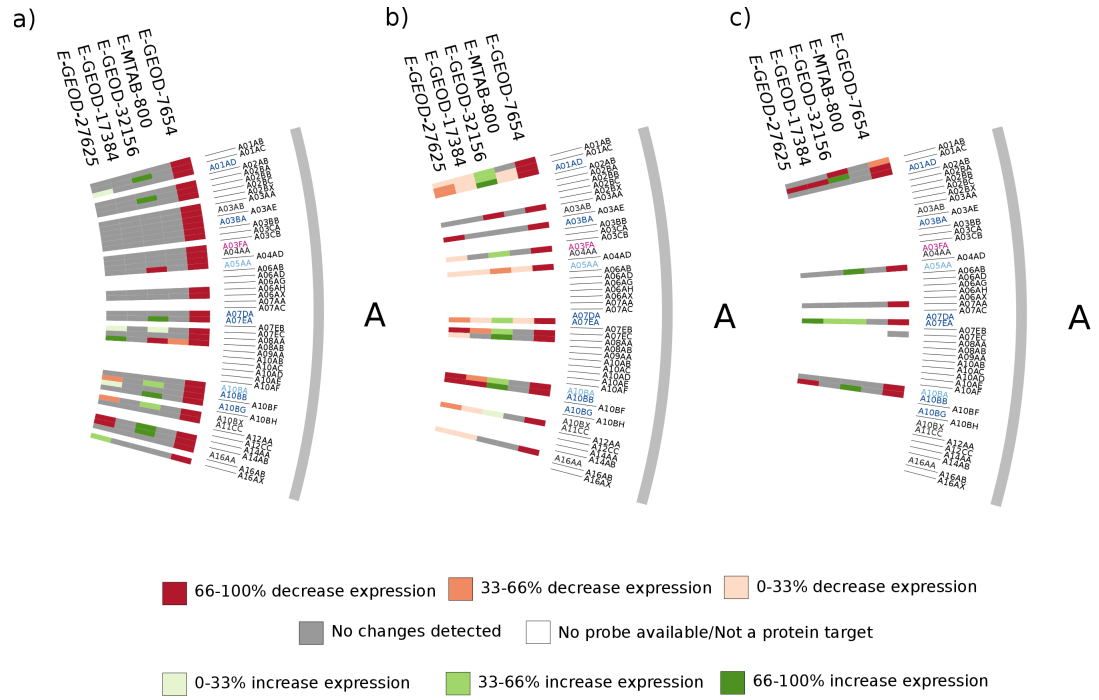


Figure 11: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in rat liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

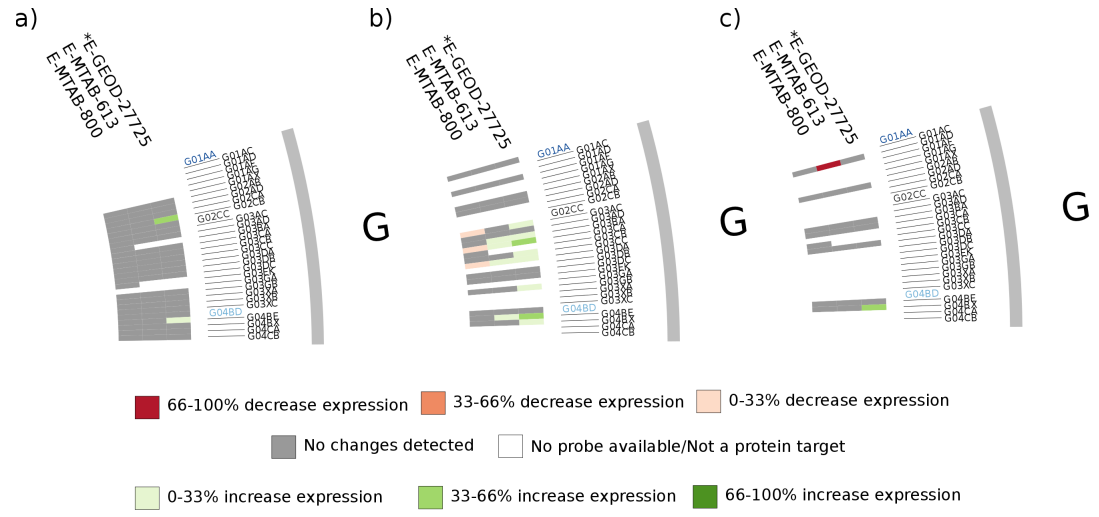


Figure 12: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in rat kidney per ATC level 4 in the genito-urinary system category (ATC code 'G'). See caption from figure 1 for further details.

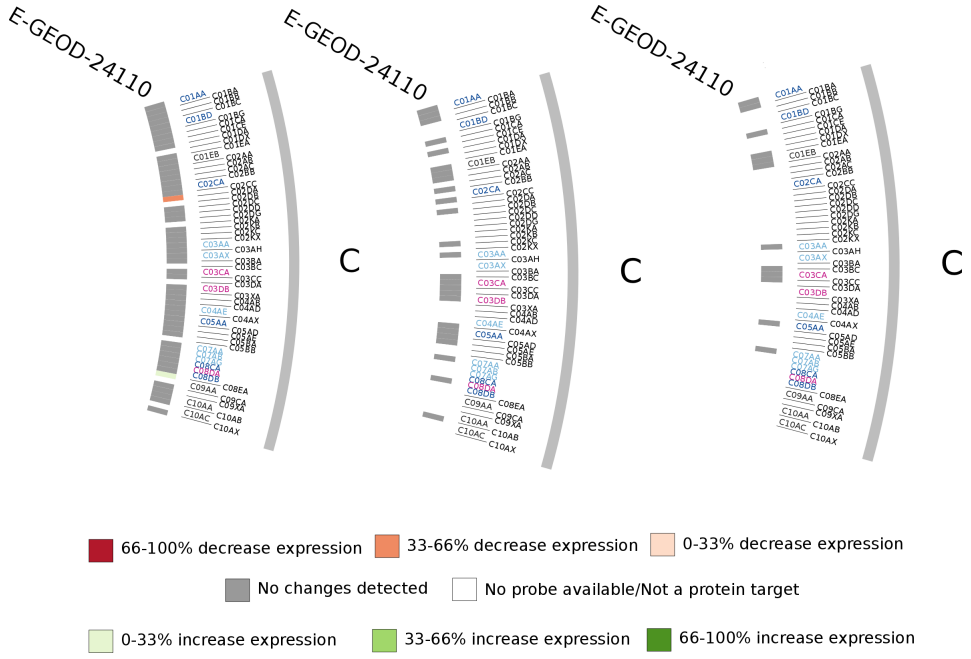


Figure 13: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in rat heart per ATC level 4 in the cardiovascular system category (ATC code 'C'). See caption from figure 1 for further details.

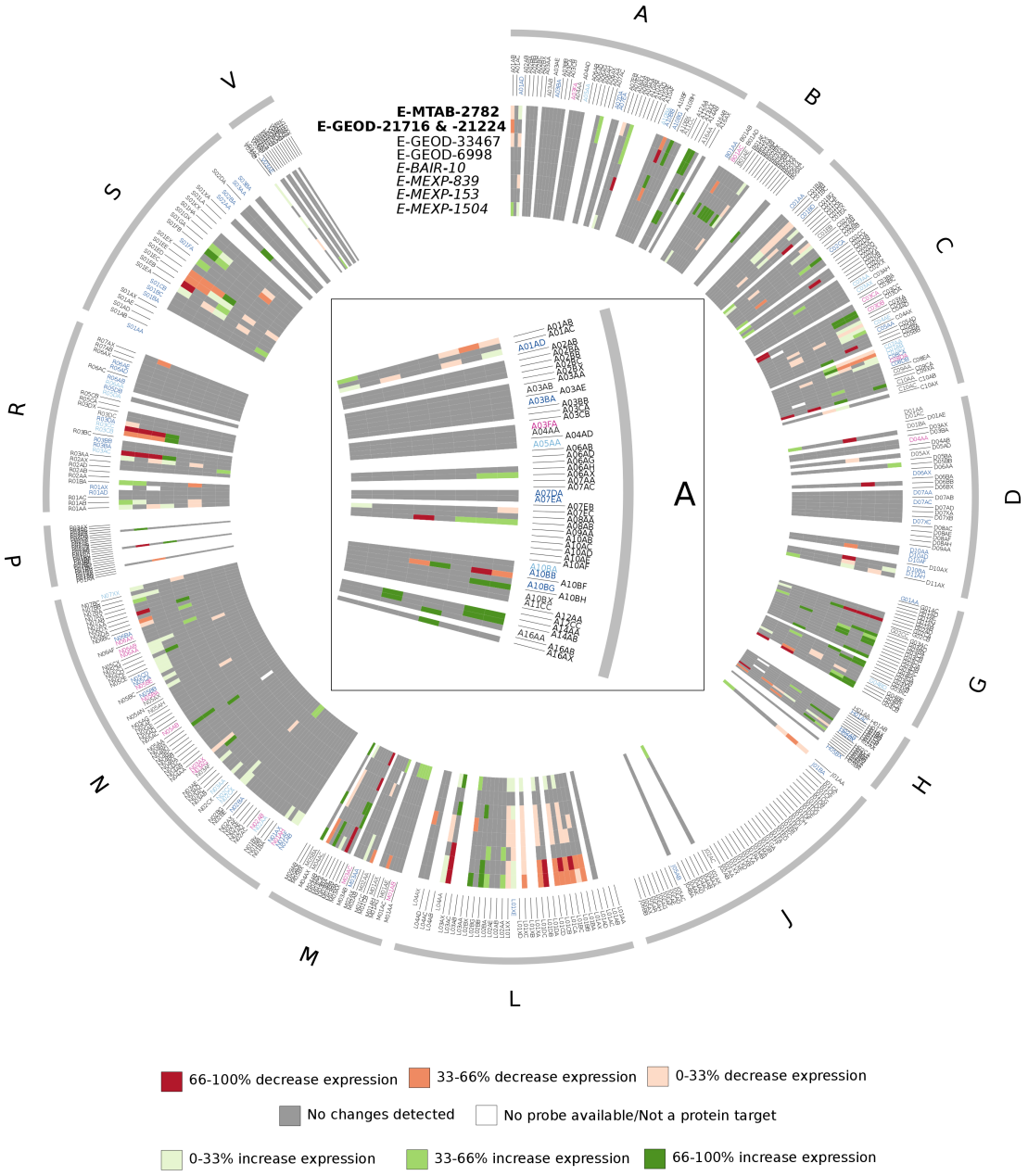


Figure 14: Overview of gene expression changes as a function of age of drug targets in mouse liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

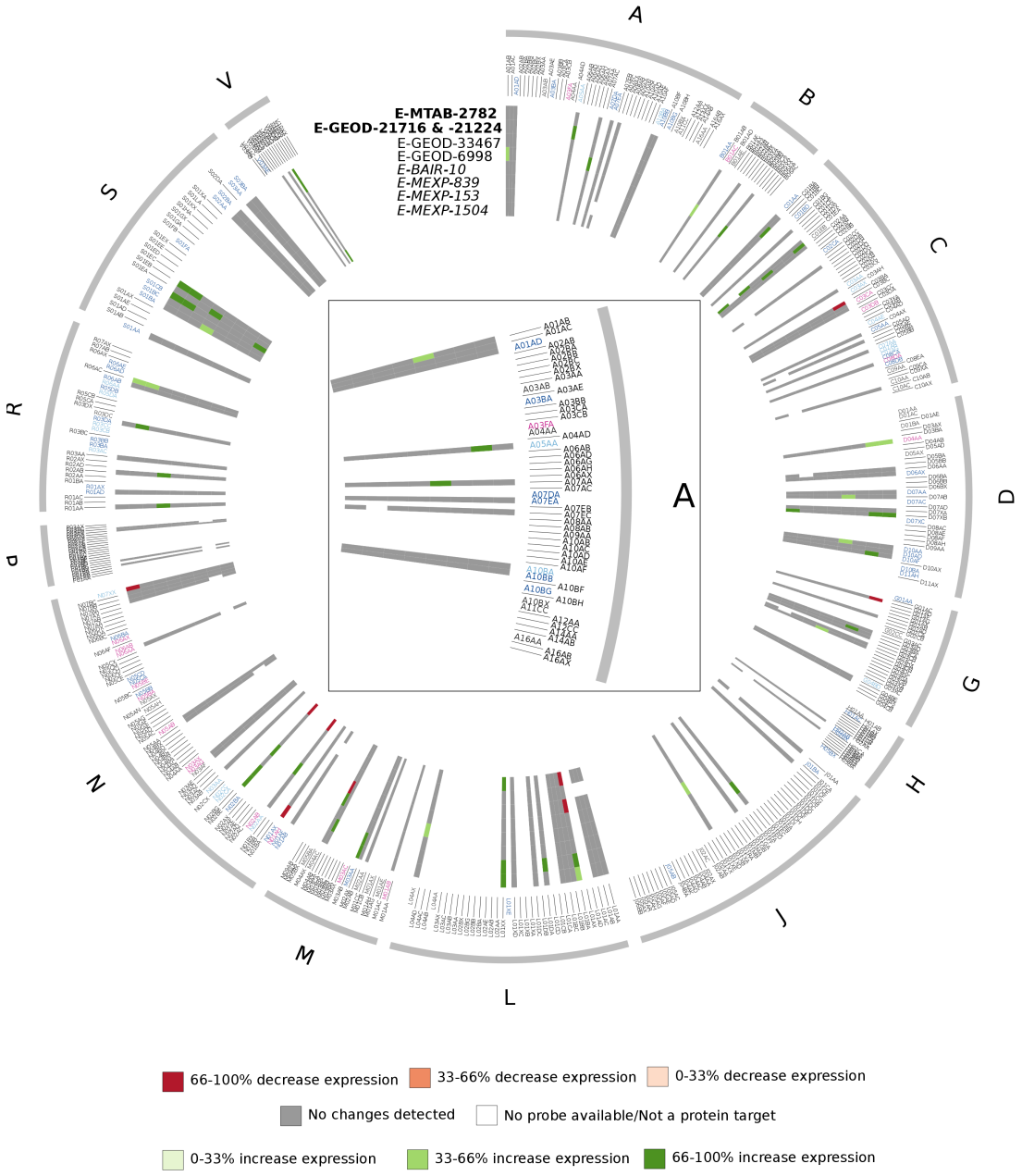
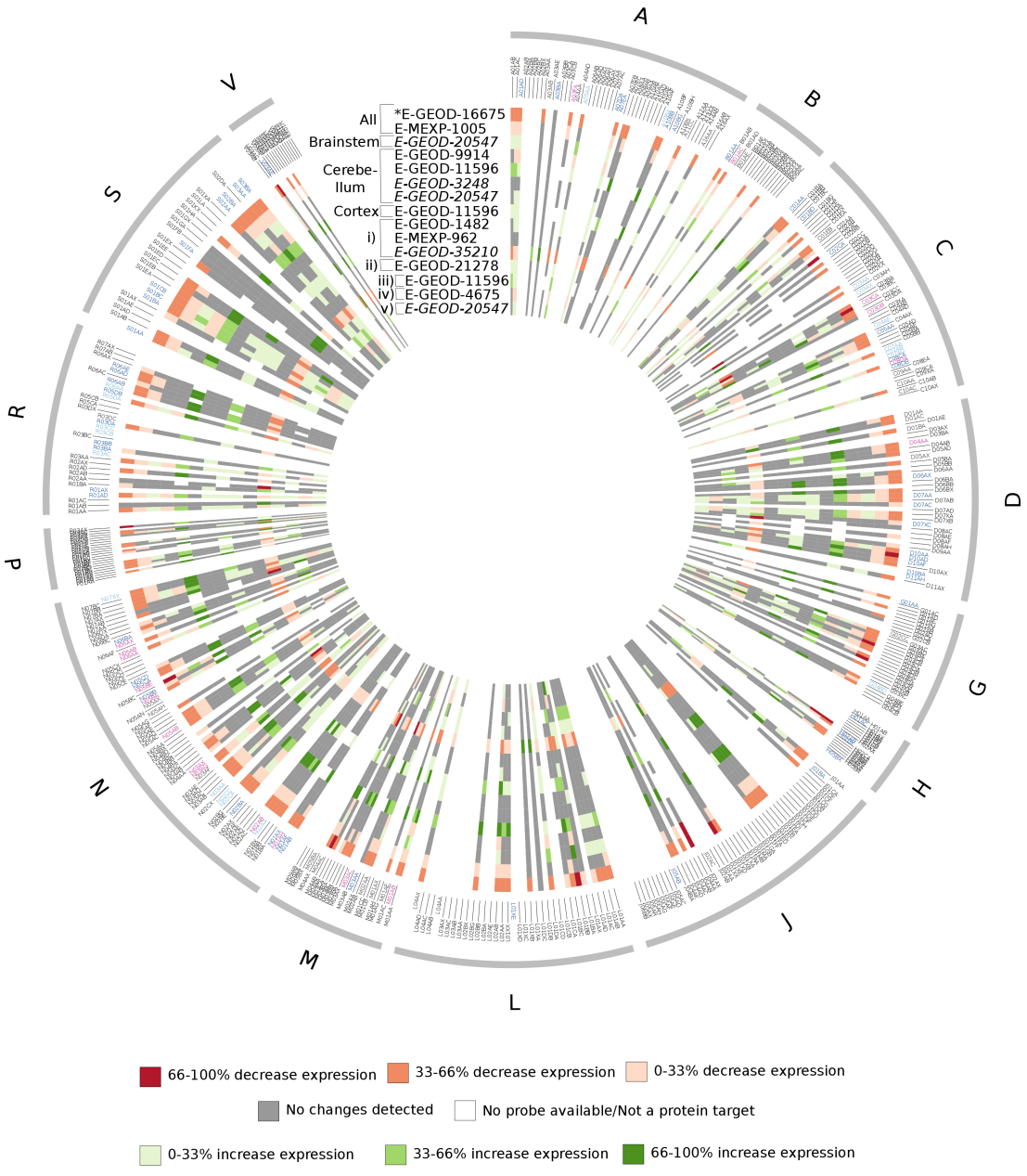


Figure 15: Overview of gene expression changes as a function of age of transporters in mouse liver per ATC level 4 in the alimentary and metabolism category (ATC code 'A'). See caption from figure 1 for further details.

(a)


A
B
C
D
G
H
J
L
M
N
P
R
S
V
All
*E-GEOD-16675
E-MEXP-1005
Brainstem
E-GEOD-20547
E-GEOD-9914
Cerebe-
llum
E-GEOD-11596
E-GEOD-3248
E-GEOD-20547
Cortex
E-GEOD-11596
E-GEOD-1482
i)
E-MEXP-962
E-GEOD-35210
ii)
E-GEOD-21278
iii)
E-GEOD-11596
iv)
E-GEOD-4675
v)
E-GEOD-20547
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

(b)

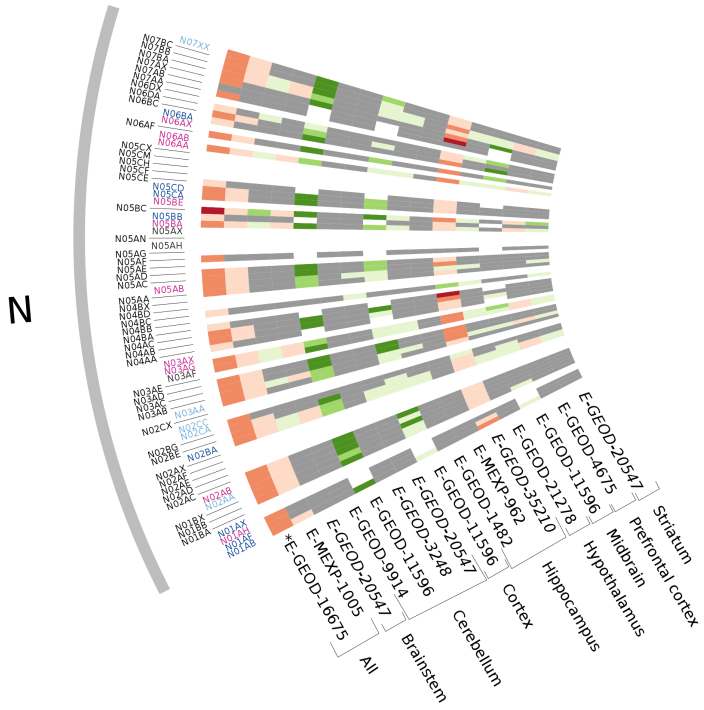
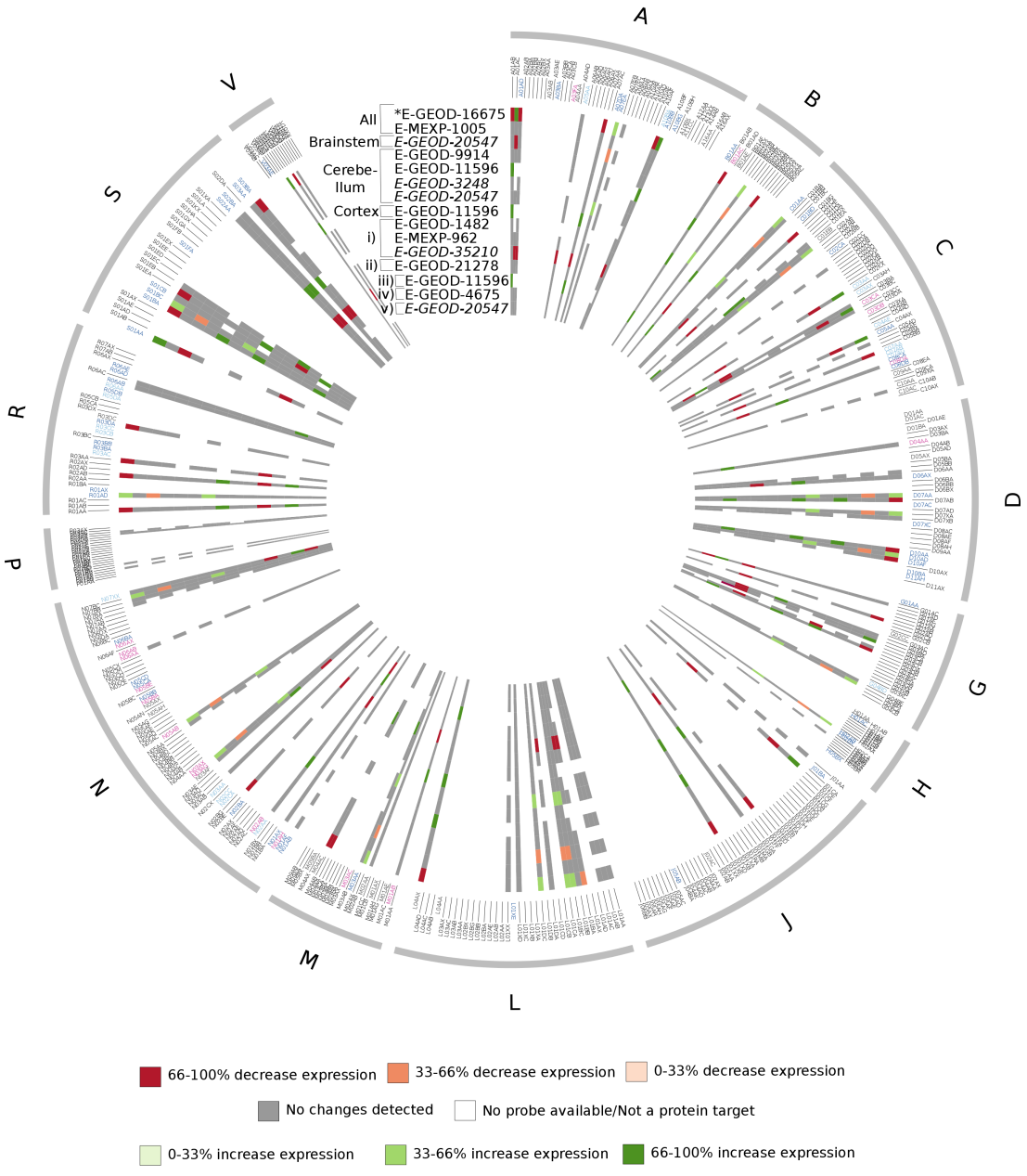


Figure 16: Overview of gene expression changes as a function of age of metabolic enzymes in mouse brain per ATC level 4 (a) in the nervous system category (ATC code 'N') (b). See caption from figure 1 for further details. i) Hippocampus, ii), Hypothalamus, iii) Midbrain, iv) Prefrontal cortex, v) Striatum.

(a)


A
B
C
D
G
H
J
L
M
N
P
R
S
V
All
Brainstem
Cerebellum
Cortex
i)
ii)
iii)
iv)
v)
*E-GEOD-16675
E-MEXP-1005
E-GEOD-20547
E-GEOD-9914
E-GEOD-11596
E-GEOD-3248
E-GEOD-20547
E-GEOD-11596
E-GEOD-1482
E-MEXP-962
E-GEOD-35210
E-GEOD-21278
E-GEOD-11596
E-GEOD-4675
E-GEOD-20547
66-100% decrease expression
33-66% decrease expression
0-33% decrease expression
No changes detected
No probe available/Not a protein target
0-33% increase expression
33-66% increase expression
66-100% increase expression

(b)

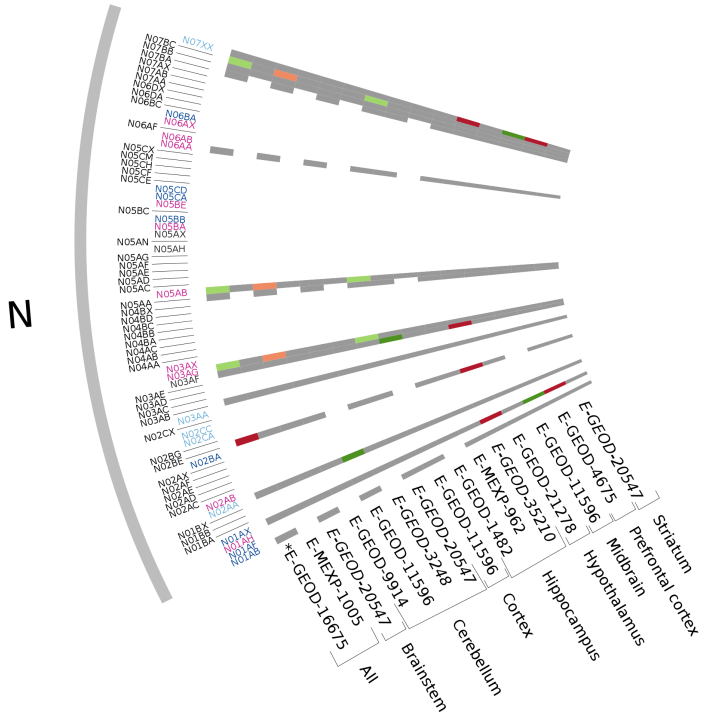


Figure 17: Overview of gene expression changes as a function of age of transporters in mouse brain per ATC level 4 (a) in the nervous system category (ATC code 'N') (b). See caption from figure 1 for further details. i) Hippocampus, ii), Hypothalamus, iii) Midbrain, iv) Prefrontal cortex, v) Striatum.

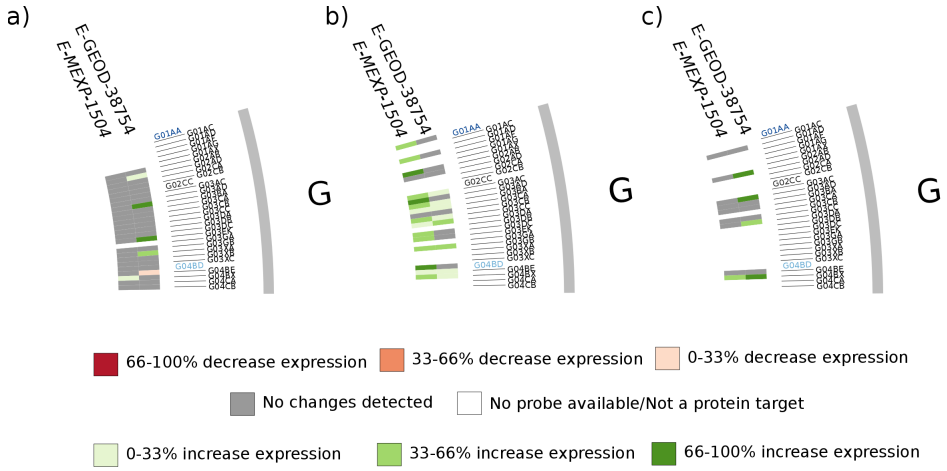


Figure 18: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in mouse kidney per ATC level 4 in the genito-urinary system category (ATC code 'G'). See caption from figure 1 for further details.

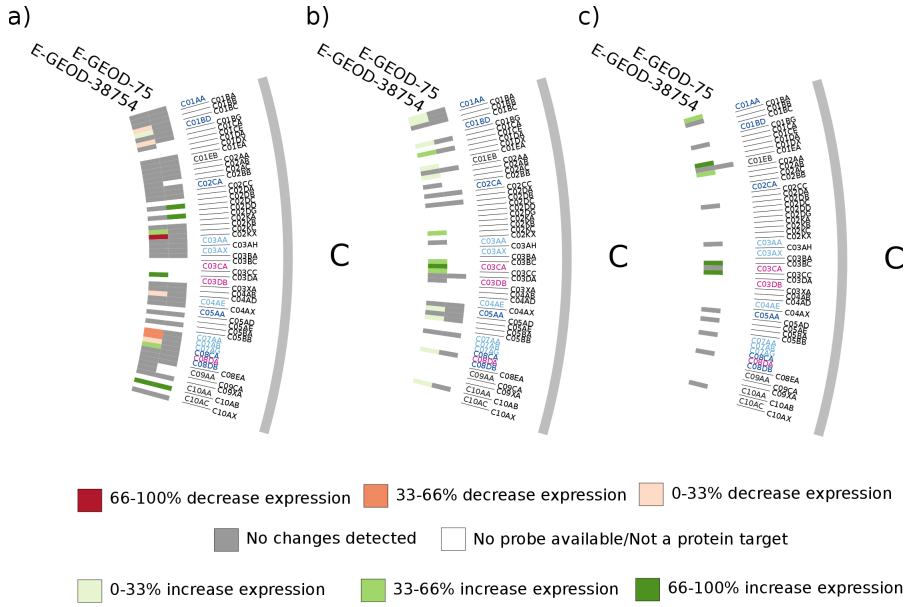


Figure 19: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in mouse heart per ATC level 4 in the cardiovascular system category (ATC code 'C'). See caption from figure 1 for further details.

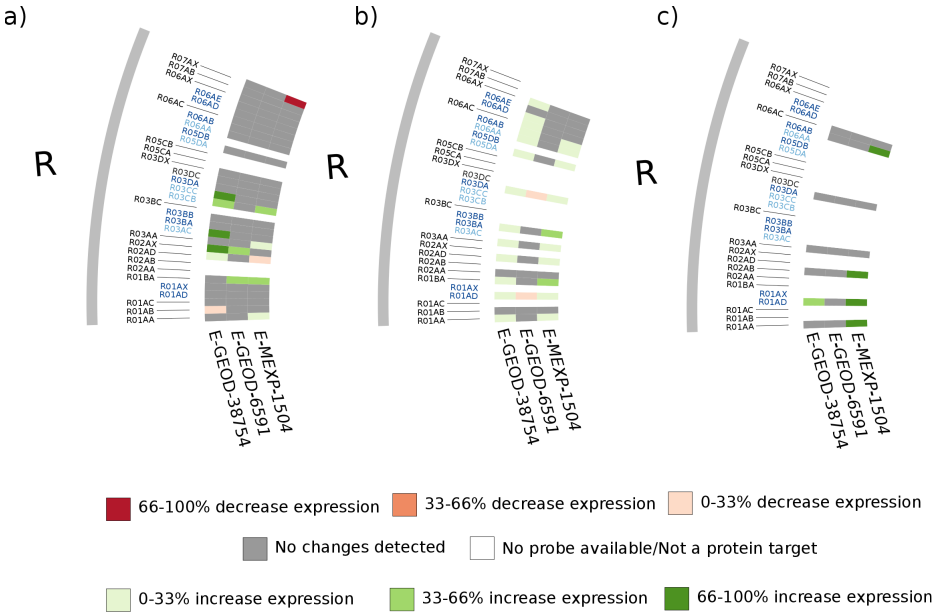


Figure 20: Overview of gene expression changes as a function of age of drug targets (a), metabolic enzymes (b), and transporters (c) in mouse lungs per ATC level 4 in the respiratory system category (ATC code 'R'). See caption from figure 1 for further details.

List of Abbreviations

Gene/Protein nomenclature

ABCA	ATP-binding cassette sub-family A
ABCB	ATP-binding cassette sub-family B
ABCC	ATP-binding cassette sub-family C
ABCG	ATP-binding cassette sub-family G
ACE	angiotensin-converting-enzyme
ADH	Alcohol dehydrogenase
ADRA2A	α-adrenergic receptor 2A
AKR	Aldo-keto reductase
ALDH	Aldehyde dehydrogenase
CACN*	Voltage-gated calcium channel
CHST	Carbohydrate sulfotransferase
CYP	Cytochrome P450
DHRS	Dehydrogenase/reductase SDR family
DRD*	Dopamine type II receptors
FMO	Flavin monoxygenase
GABR*	GABA A receptor
GLR*	Glycine receptor
GPX	Glutathione peroxidase
GRIN*	Glutamate [NMDA] receptor
GST	Glutathione transferase
HERG	Voltage-gated potassium channel subunit Kv11.1
HMGCR	3-hydroxy-3-methylglutaryl-coenzyme A reductase
HTR1A	Serotonin 1a Receptor
KCNK*	Potassium channel subfamily K members
M3	Muscarinic receptor 3

List of Abbreviations

MAOB	Monoamine Oxidase B
MT-ND*	Mitochondrial complex I (NADH dehydrogenase)
NAT	N-acetyltransferase
PBPs	Penicillin-Binding Proteins
SCN*	Sodium channel
SLC	Solute carrier superfamily
SLCO	Solute carrier organic anion transporter family
SULT	Sulfotransferase
TAP	Antigen peptide transporter
UGT	UDP-glucoronosyltransferase

Acronyms

ADHD	Attention Deficit Hyperactivity Disorder
ADME	Absorption, Distribution, Metabolism and Excretion
ADRs	Adverse Drug Reactions
ATC	Anatomical Therapeutic Chemical Classification System code
AUC	Area Under the Curve
BBB	Blood-Brain-Barrier
BNFC	British National Formulary for Children
BoW	Bag-of-Words
BPCA	Best Pharmaceuticals for Children Act
BSA	Body Surface Area
CKD	Chronic Kidney Disease
CNS	Central Nervous System
DALYs	Disability-Adjusted Life Years
DDI	Drug-Drug Interaction
DDR	differential drug responses
DT	Drug target
EC	European Commission
EFO	Experimental Factor Ontology
EMA	European Medicines Agency
EMBL	European Molecular Biology Laboratory
EUGMS	European Union Geriatric Medicine Societies
FDA	Food and Drug Administration

GBD	Global Burden of Disease
GEO	Gene Expression Omnibus
IACUC	Institutional Animal Care and Use Committee
ICH E7	International Conference of Harmonisation
INN	International Nonproprietary Names
logP	Partition coefficient
MCC	Matthews Correlation Coefficient
MDD	Major Depressive Disorder
ME	Metabolic enzyme
MIAME	Minimum information about a microarray experiment
mRNA	messenger RNA
NB	Naïve Bayesian
NSAID	Non-Steroidal Anti-Inflammatory Drugs
PCA	Principal Component Analysis
PK/PD	Pharmacokinetic and Pharmacodynamic models
PSA	Polar Surface Area
RNA-seq	RNA Sequencing or Whole Transcriptome Shotgun Sequencing
ROC	Receiver Operating Characteristic
SAGE	Human Liver cohort
SPC	Summary of Product Characteristics
SSRIs	Selective Serotonin Re-uptake Inhibitors
START	Screening Tool to Alert doctors to Right Treatment
STEP	Safety and Toxicity of Excipients for Pediatrics
STOPP	Screening Tool of Older Persons Prescriptions
TP	Transporter
TTD	Therapeutic targets database
Uniprot	Universal Protein Resource
USAN	United States Adopted Name
WHO	World Health Organisation

References

[1] J. W. von Goethe, "GoodReads - Quote 133314." 1

[2] G. B. Shaw, "GoodReads - Quote 23088." 2

[3] P. Smit-marshall, "Pediatric Trials: A Worldview," *Applied Clinical Trials*, vol. 19, no. 1, p. 32, 2010. 2, 4, 5, 33

[4] A. Cras, M.-A. Conscience, G. Rajzbaum, A. Lillo-Le Louët, N. Lopez, I. Tersen, and Y. Bezie, "Off-label prescribing in a French hospital.," *Pharmacy world & science : PWS*, vol. 29, pp. 97–100, Apr. 2007. 2

[5] A. W. Craft, J. T. Brocklebank, E. N. Hey, and R. H. Jackson, "The 'grey toddler'. Chloramphenicol toxicity.," *Archives of disease in childhood*, vol. 49, pp. 235–7, Mar. 1974. 3

[6] M. Chen, B. Leduc, S. Kerr, D. Howe, and D. A. Williams, "Identification of Human UGT2B7 as the Major Isoform Involved in the O-Glucuronidation of Chloramphenicol," *Drug Metabolism and Disposition*, vol. 38, no. 3, pp. 368–375, 2010. 3, 108

[7] C. D. van der Marel, B. J. Anderson, R. A. van Lingen, N. H. G. Holford, M. A. L. Pluim, F. G. A. Jansman, J. N. van den Anker, and D. Tibboel, "Paracetamol and metabolite pharmacokinetics in infants.," *European journal of clinical pharmacology*, vol. 59, pp. 243–51, July 2003. 3

[8] D. J. Safer and J. M. Zito, "Treatment-emergent adverse events from selective serotonin reuptake inhibitors by age group: children versus adolescents.," *Journal of child and adolescent psychopharmacology*, vol. 16, no. 1-2, pp. 159–69, 2006. 3

[9] D. Bylund and A. Reed, "Childhood and adolescent depression: Why do children and adults respond differently to antidepressant drugs?," *Neurochemistry international*, vol. 51, no. 5, pp. 246–253, 2007. 3

[10] S. M. Hall, "Reye ' s syndrome and aspirin : a review," *Journal of the Royal Society of Medicine*, vol. 79, pp. 596–598, 1986. 3

[11] J. F. Glasgow and B. Middleton, "Reye syndrome–insights on causation and prognosis.," *Archives of disease in childhood*, vol. 85, pp. 351–3, Nov. 2001. 3

[12] S. N. Hilmer, A. J. McLachlan, and D. G. Le Couteur, "Clinical pharmacology in the geriatric patient.," *Fundamental & clinical pharmacology*, vol. 21, pp. 217–30, June 2007. 3, 20, 40, 41

[13] M. W. Bowie and P. W. Slattum, "Pharmacodynamics in older adults: a review.," *The American journal of geriatric pharmacotherapy*, vol. 5, pp. 263–303, Sept. 2007. 3, 40, 41, 45, 109, 111, 146, 151, 152, 162

[14] C. D. van der Marel, B. J. Anderson, M. A. L. Pluim, T. H. R. de Jong, A. Gonzalez, and D. Tibboel, "Acetaminophen in cerebrospinal fluid in children.," *European journal of clinical pharmacology*, vol. 59, pp. 297–302, Aug. 2003. 3, 122

[15] M. M. Laughon, D. K. B. Jr, E. V. Capparelli, G. L. Kearns, K. Berezny, I. M. Paul, K. Wade, J. Barrett, P. B. Smith, and M. Cohen-wolkowiez, "Innovative clinical trial design for pediatric therapeutics," *Expert Review of Clinical Pharmacology*, vol. 4, no. 5, pp. 643–652, 2011. 4

[16] V. Ivanovska, C. M. a. Rademaker, L. van Dijk, and A. K. Mantel-Teeuwisse, "Pediatric Drug Formulations: A Review of Challenges and Progress.," *Pediatrics*, July 2014. 4, 5, 13

[17] S. Conroy and J. McIntyre, "The use of unlicensed and off-label medicines in the neonate.," *Seminars in fetal & neonatal medicine*, vol. 10, pp. 115–22, Apr. 2005. 4

REFERENCES

[18] P. Auby, "Pharmaceutical research in paediatric populations and the new EU Paediatric Legislation: an industry perspective.," *Child and adolescent psychiatry and mental health*, vol. 2, p. 38, Jan. 2008. 4

[19] S. Iyasu and M. D. Murphy, "Pharmacovigelance In Pediatrics & Drugs and the Elderly," in *Pharmacovigelance*, pp. 497–506 & 515–532, Chichester, England: Wiley, 2nd ed., 2007. 5, 10, 11, 13

[20] S. Salunke, G. Giacoia, and C. Tuleu, "The STEP (safety and toxicity of excipients for paediatrics) database. Part 1-A need assessment study.," *International journal of pharmaceutics*, vol. 435, pp. 101–11, Oct. 2012. 5

[21] S. Salunke, B. Brandys, G. Giacoia, and C. Tuleu, "The STEP (Safety and Toxicity of Excipients for Paediatrics) database: part 2 - the pilot version.," *International journal of pharmaceutics*, vol. 457, pp. 310–22, Nov. 2013. 5

[22] P. Munzenberger and P. McKercher, "Pediatric dosing - the pharmacist's dilemma," *Contemporary pharmacy practice*, vol. 3, no. 1, pp. 11–4, 1980. 5

[23] J. M. Daniels and L. M. Smith, "Pediatric Dosage," in *Clinical Calculations: A Unified Approach*, ch. 11, Cengage Learning, 5th ed., 2006. 5, 6

[24] M. Johnston, "Dosage Calculations," in *Pharmacy Calculations: The Pharmacy Technician Series*, pp. Ch.5 1–24, Prentice Hall, 2006. 5

[25] G. P. Elias, C. Antoniali, and R. C. Mariano, "Comparative study of rules employed for calculation of pediatric drug dosage.," *Journal of applied oral science : revista FOB*, vol. 13, pp. 114–9, June 2005. 6

[26] Paediatric Formulary Committee, *BNF for Children 2014-2015 (BNFC)*. London: Pharmaceutical Press, 2014. 6

[27] C. Knibbe, E. Krekels, J. van den Anker, J. DeJongh, G. Santen, M. van Dijk, S. Simons, R. van Lingen, E. Jacqz-Aigrain, M. Danhof, and D. Tibboel, "Morphine Glucuronidation in Preterm Neonates, Infants and Children Younger than 3 Years," *Clinical Pharmacokinetics*, vol. 48, no. 6, pp. 371–385, 2009. 6

[28] E. Krekels, J. DeJongh, R. van Lingen, C. van der Marel, I. Choonara, A. Lynn, M. Danhof, D. Tibboel, and C. Knibbe, "Predictive Performance of a Recently Developed Population Pharmacokinetic Model for Morphine and its Metabolites in New Datasets of (Preterm) Neonates, Infants and Children," *Clinical Pharmacokinetics*, vol. 50, no. 1, pp. 51–63, 2011. 6

[29] B. Agoram, W. S. Woltosz, and M. B. Bolger, "Predicting the impact of physiological and biochemical processes on oral drug bioavailability.," *Advanced drug delivery reviews*, vol. 50 Suppl 1, pp. S41–67, Oct. 2001. 6

[30] V. Sinha, K. Vaarties, S. De Buck, L. Fenu, M. Nijsen, R. Gilissen, W. Sanderson, K. Uytsel, E. Hoeben, A. Peer, C. Mackie, and J. Smit, "Towards a Better Prediction of Peak Concentration, Volume of Distribution and Half-Life after Oral Drug Administration in Man, Using Allometry," *Clinical Pharmacokinetics*, vol. 50, no. 5, pp. 307–318, 2011. 7

[31] I. Mahmood, "Allometric extrapolation of factors VII, VIII and IX clearance in children from adults.," *Journal of thrombosis and haemostasis : JTH*, vol. 10, pp. 1609–13, Aug. 2012. 7

[32] S. Stegemann, F. Ecker, M. Maio, P. Kraahs, R. Wohlfart, J. Breitkreutz, A. Zimmer, D. Bar-Shalom, P. Hettrich, and B. Broegmann, "Geriatric drug therapy: neglecting the inevitable majority.," *Ageing research reviews*, vol. 9, pp. 384–98, Oct. 2010. 7, 33, 150

[33] European Medicines Agency, "Ensuring safe and effective medicines for an ageing population : Workshop Proceedings," (London, UK), 22–23 March 2012. 7, 8, 9

[34] A. Cherubini, S. Del Signore, J. Ouslander, T. Semla, and J.-P. Michel, "Fighting against age discrimination in clinical trials.," *Journal of the American Geriatrics Society*, vol. 58, pp. 1791–6, Sept. 2010. 7, 8, 33

[35] Studies in support of special populations: geriatric E7, "International conference on harmonization of technical requirements for registration of pharmaceuticals for human use.," 1994. 7

[36] Committee for Medicinal Products for Human Use (CHMP), "EMA geriatric medicines strategy," 2011. 8

[37] M. P. Doogue and T. M. Polasek, "Drug dosing in renal disease.," *The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists*, vol. 32, pp. 69–73, May 2011. 8

[38] C. M. Campanelli, "American Geriatrics Society Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults," *Journal of the American Geriatrics Society*, vol. 60, no. 4, pp. 616–631, 2013. 9, 40, 114

[39] H. Hamilton, P. Gallagher, C. Ryan, S. Byrne, and D. O'Mahony, "Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients.," *Archives of internal medicine*, vol. 171, pp. 1013–9, June 2011. 9

[40] P. Gallagher, C. Ryan, S. Byrne, J. Kennedy, and D. O'Mahony, "STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment). Consensus validation.," *International journal of clinical pharmacy*, vol. 42(2), pp. 72–83, 2008. 9, 40, 114

[41] A. Z. Crepeau, B. D. Moseley, and E. C. Wirrell, "Specific safety and tolerability considerations in the use of anticonvulsant medications in children.," *Drug, healthcare and patient safety*, vol. 4, pp. 39–54, Jan. 2012. 10, 11, 12, 14, 16, 17, 20, 108, 150

[42] G. L. Kearns, S. M. Abdel-Rahman, S. W. Alander, D. L. Blowey, J. S. Leeder, and R. E. Kauffman, "Developmental pharmacology–drug disposition, action, and therapy in infants and children.," *The New England journal of medicine*, vol. 349, pp. 1157–67, Sept. 2003. 10, 13, 14, 15, 16, 17, 18, 36

[43] O. Jones and N. Gautam, *The hands-on guide to practical pescribing.* Blackwell Publishing, 2008. 11

[44] H. Gray in *Anatomy of the Human Body* (W. H. Lewis, ed.), ch. XI, Philadelphia: Lea & Febiger, 20 ed., 1918. 11, 20

[45] J. Ali Raza and A. Movahed, "Use of cardiovascular medications in the elderly.," *International journal of cardiology*, vol. 85, pp. 203–15, Oct. 2002. 12, 19, 40

[46] G. Rylance, "Clinical pharmacology: Drugs in children," *British Medical Journal*, vol. 282, pp. 50–51, 1981. 13

[47] L. E. Ramsay and G. T. Tucker, "Clinical pharmacology: Drugs and the elderly," *British Medical Journal*, vol. 282, pp. 125–127, 1981. 13

[48] K. Turnheim, "When drug therapy gets old: pharmacokinetics and pharmacodynamics in the elderly," *Experimental Gerontology*, vol. 38, pp. 843–853, Aug. 2003. 13

[49] M. L. Becker and J. S. Leeder, "Identifying genomic and developmental causes of adverse drug reactions in children.," *Pharmacogenomics*, vol. 11, pp. 1591–602, Nov. 2010. 13

[50] R. Bellmann-Weiler and G. Weiss, "Pitfalls in the diagnosis and therapy of infections in elderly patients–a mini-review.," *Gerontology*, vol. 55, pp. 241–9, Jan. 2009. 14, 20

[51] R. N. Hines, "The ontogeny of drug metabolism enzymes and implications for adverse drug events.," *Pharmacology & therapeutics*, vol. 118, pp. 250–67, May 2008. 15, 16, 17, 18, 20, 98, 108

[52] J. D. Marshall and G. L. Kearns, "Developmental pharmacodynamics of cyclosporine.," *Clinical pharmacology and therapeutics*, vol. 66, pp. 66–75, July 1999. 20

[53] "PubMed/MEDLINE." http://www.pubmed.org. 26

[54] "Europe PubMed Central." http://europepmc.org/. 26

[55] "Embase." http://www.elsevier.com/online-tools/embase. 26

[56] D. Vishnyakova, E. Pasche, and P. Ruch, "Using binary classification to prioritize and curate articles for the Comparative Toxicogenomics Database.," *Database : the journal of biological databases and curation*, vol. 2012, p. bas050, 2012. 27

[57] A. P. Davis, T. C. Wiegers, R. J. Johnson, J. M. Lay, K. Lennon-Hopkins, C. Saraceni-Richards, D. Sciaky, C. G. Murphy, and C. J. Mattingly, "Text mining effectively scores and ranks the literature for improving chemical-gene-disease curation at the comparative toxicogenomics database.," *PloS one*, vol. 8, no. 4, p. e58201, 2013. 27

[58] G. Papadatos, G. J. van Westen, S. Croset, R. Santos, S. Trubian, and J. P. Overington, "A document classifier for medicinal chemistry publications trained on the ChEMBL corpus," *Journal of Chemoinformatics*, vol. 6, p. 40, 2014. 27

[59] P. Domingos and M. Pazzani, "On the Optimality of the Simple Bayesian Classifier under Zero-One Loss," *Machine Learning*, vol. 29, no. 2-3, pp. 103–130, 1997. 28

[60] WHO, "Disease incidence, prevalence and disability," in *The Global Burden of Disease: 2004 Update*, ch. Part 3, pp. 28–37, Geneva: WHO press, 1 ed., 2008. 32, 154

REFERENCES

[61] B. Birmaher, N. D. Ryan, D. E. Williamson, D. a. Brent, J. Kaufman, R. E. Dahl, J. Perel, and B. Nelson, "Childhood and adolescent depression: a review of the past 10 years. Part I.," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 35, pp. 1427–39, Nov. 1996. 32

[62] WHO Collaborating Centre for Drug Statistics Methodology, *ATC classification index with DDDs.* Oslo 2012, 2013. 34, 54, 103, 136

[63] H. Se, M. Sn, J. Mckenzie, P. Sindahl, and M. Proctor, "Selective serotonin reuptake inhibitors (SSRIs) for depressive disorders in children and adolescents (Review)," *Cochrane Database of Systematic Reviews*, no. 3, p. CD004851, 2007. 36, 43, 44, 153

[64] R. Steinbrook, "Testing medications in children," *The New England journal of medicine*, vol. 347, no. 18, pp. 1462–1470, 2002. 36, 149

[65] H. Goldberg-Stern, L. Feldman, T. Eidlitz-Markus, U. Kramer, S. Perez, L. Pollak, and A. Phatal-Valevski, "Levetiracetam in children, adolescents and young adults with intractable epilepsy: efficacy, tolerability and effect on electroencephalogram–a pilot study.," *European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society*, vol. 17, pp. 248–53, May 2013. 36, 45

[66] G. D. Anderson, "Children versus adults: pharmacokinetic and adverse-effect differences.," *Epilepsia*, vol. 43 Suppl 3, pp. 53–9, Jan. 2002. 36

[67] J. Dunne, W. J. Rodriguez, M. D. Murphy, B. N. Beasley, G. J. Burckart, J. D. Filie, L. L. Lewis, H. C. Sachs, P. H. Sheridan, P. Starke, and L. P. Yao, "Extrapolation of adult data and other data in pediatric drug-development programs.," *Pediatrics*, vol. 128, pp. e1242–9, Nov. 2011. 36

[68] L. B. Bacharier and T. W. Guilbert, "Diagnosis and management of early asthma in preschool-aged children.," *The Journal of allergy and clinical immunology*, vol. 130, pp. 287–96, Aug. 2012. 36

[69] E. Kirino, "Escitalopram for the management of major depressive disorder: a review of its efficacy, safety, and patient acceptability.," *Patient preference and adherence*, vol. 6, pp. 853–61, Jan. 2012. 36

[70] A. Henry, M. D. Kisicki, and C. Varley, "Efficacy and safety of antidepressant drug treatment in children and adolescents.," *Molecular psychiatry*, vol. 17, pp. 1186–93, Dec. 2012. 37, 113

[71] J. Fashner, K. Ericson, and S. Werner, "Treatment of the common cold in children and adults.," *American family physician*, vol. 86, pp. 153–9, July 2012. 37

[72] S. Smith, K. Schroeder, and T. Fahey, "Over-the-counter (OTC) medications for acute cough in children and adults in ambulatory settings," *Cochrane Database of Systematic Reviews*, no. 8, p. CD001831, 2012. 37

[73] A. K. Ghose, T. Herbertz, R. L. Hudkins, B. D. Dorsey, and J. P. Mallamo, "Knowledge-Based, Central Nervous System (CNS) Lead Selection and Lead Optimization for CNS Drug Discovery.," *ACS chemical neuroscience*, vol. 3, pp. 50–68, Jan. 2012. 38

[74] C. F. Gravina, M. Bertolami, and G. H. Rodrigues, "Dyslipidemia: evidence of efficacy of the pharmacological and non-pharmacological treatment in the elderly.," *Journal of geriatric cardiology : JGC*, vol. 9, pp. 83–90, June 2012. 41

[75] E. Hancock, "Treatment of infantile spasms," *Cochrane Database of Systematic Reviews*, no. 6, p. CD001770, 2008. 42

[76] D. Elbe and Z. Lalani, "Review of the pharmacotherapy of irritability of autism.," *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Académie canadienne de psychiatrie de l'enfant et de l'adolescent*, vol. 21, pp. 130–46, May 2012. 43

[77] L. P. M. M. Wijlaars, I. Nazareth, and I. Petersen, "Trends in depression and antidepressant prescribing in children and adolescents: a cohort study in The Health Improvement Network (THIN).," *PloS one*, vol. 7, p. e33181, Jan. 2012. 44

[78] V. N. Yewale and D. Dharmapalan, "Promoting appropriate use of drugs in children.," *International journal of pediatrics*, vol. 2012, p. 906570, Jan. 2012. 45

[79] Accelrys Software Inc., *Pipeline Pilot.* 2011. 46

[80] R Core Team, *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2012. 48, 134, 135

[81] J. Drews, "Genomic sciences and the medicine of tomorrow," *Nature biotechnology*, vol. 14, pp. 1516–1518, 1996. 50

[82] J. Drews and S. Ryser, "The role of innovation in drug development," *Nature biotechnology*, vol. 15, pp. 1318–1319, 1997. 50

[83] X. Chen, Z. L. Ji, and Y. Z. Chen, "TTD: Therapeutic Target Database.," *Nucleic acids research*, vol. 30, pp. 412–5, Jan. 2002. 50

[84] F. Zhu, B. Han, P. Kumar, X. Liu, X. Ma, X. Wei, L. Huang, Y. Guo, L. Han, C. Zheng, and Y. Chen, "Update of TTD: Therapeutic Target Database.," *Nucleic acids research*, vol. 38, pp. D787–91, Jan. 2010. 50

[85] A. L. Hopkins and C. R. Groom, "The druggable genome.," *Nature reviews. Drug discovery*, vol. 1, pp. 727–30, Sept. 2002. 50

[86] J. B. Golden, "Prioritizing the human genome: knowledge management for drug discovery.," *Curr Opin Drug Discov Devel*, vol. 6, pp. 310–316, May 2003. 50

[87] J. P. Overington, B. Al-Lazikani, and A. L. Hopkins, "How many drug targets are there?," *Nature reviews. Drug discovery*, vol. 5, pp. 993–6, Dec. 2006. 50

[88] P. Imming, C. Sinning, and A. Meyer, "Drugs, their targets and the nature and number of drug targets," *Nature reviews Drug discovery*, vol. 5, pp. 821–835, 2006. 50

[89] D. S. Wishart, C. Knox, A. C. Guo, S. Shrivastava, M. Hassanali, P. Stothard, Z. Chang, and J. Woolsey, "DrugBank: a comprehensive resource for in silico drug discovery and exploration.," *Nucleic acids research*, vol. 34, pp. D668–72, Jan. 2006. 50

[90] C. Knox, V. Law, T. Jewison, P. Liu, S. Ly, A. Frolkis, A. Pon, K. Banco, C. Mak, V. Neveu, Y. Djoumbou, R. Eisner, A. C. Guo, and D. S. Wishart, "DrugBank 3.0: a comprehensive resource for 'omics' research on drugs.," *Nucleic acids research*, vol. 39, pp. D1035–41, Jan. 2011. 50

[91] S. Günther, M. Kuhn, M. Dunkel, M. Campillos, C. Senger, E. Petsalaki, J. Ahmed, E. G. Urdiales, A. Gewiess, L. J. Jensen, R. Schneider, R. Skoblo, R. B. Russell, P. E. Bourne, P. Bork, and R. Preissner, "SuperTarget and Matador: resources for exploring drug-target relationships.," *Nucleic acids research*, vol. 36, pp. D919–22, Jan. 2008. 50

[92] M. Rask-Andersen, M. S. Almén, and H. B. Schiöth, "Trends in the exploitation of novel drug targets," *Nature Reviews Drug Discovery*, vol. 10, pp. 579–590, Aug. 2011. 50

[93] B. Munos, "A Forensic Analysis of Drug Targets from 2000 through 2012," *Clinical Pharmacology & Therapeutics*, vol. 94, pp. 407–411, June 2013. 50

[94] P. Agarwal, P. Sanseau, and L. R. Cardon, "Novelty in the target landscape of the pharmaceutical industry.," *Nature reviews. Drug discovery*, vol. 12, pp. 575–6, Aug. 2013. 50

[95] The Uniprot Consortium, "Update on activities at the Universal Protein Resource (UniProt) in 2013.," *Nucleic acids research*, vol. 41, pp. D43–7, Jan. 2013. 51

[96] "PharmaADME Consortium." http://pharmaadme.org, 2009. 52

[97] "Pharmacology weekly." http://www.pharmacologyweekly.com, 2008. 52

[98] Indiana University, "P450 drug interaction table." http://medicine.iupui.edu/clinpharm/ddis/main-table/, 2011. 52

[99] S. Preissner, K. Kroll, M. Dunkel, C. Senger, G. Goldsobel, D. Kuzman, S. Guenther, R. Winnenburg, M. Schroeder, and R. Preissner, "SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions.," *Nucleic acids research*, vol. 38, pp. D237–43, Jan. 2010. 52

[100] A. P. Bento, A. Gaulton, A. Hersey, L. J. Bellis, J. Chambers, M. Davies, F. A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, and J. P. Overington, "The ChEMBL bioactivity database: an update.," *Nucleic acids research*, vol. 42, pp. D1083–90, Nov. 2014. 54

[101] H. P. Rang, M. M. Dale, J. M. Ritter, R. J. Flower, and G. Henderson, *Pharmacology*. Student consult, Elsevier Health Sciences UK, 7th ed., 2012. 57, 105

[102] H. J. Kempen, M. Vermeer, E. de Wit, and L. M. Havekes, "Vastatins inhibit cholesterol ester accumulation in human monocyte- derived macrophages," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 11, pp. 146–153, Jan. 1991. 58

[103] A. Parkinson and B. W. Ogilvie, "Biotransformation of Xenobiotics," in *Casarett and Doull's toxicology : the basic science of poisons*, ch. 6, McGraw-Hill, 7 ed., 2008. 61, 126

REFERENCES

[104] O. A. Barski, S. M. Tipparaju, and A. Bhatnagar, "The Aldo-Keto Reductase Superfamily and its Role in Drug Metabolism and Detoxification," *Drug Metabolism Review*, vol. 40, no. 4, pp. 553–624, 2008. 61

[105] S. A. Marchitti, C. Brocker, D. Stagos, and V. Vasiliou, "Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily.," *Expert opinion on drug metabolism & toxicology*, vol. 4, pp. 697–720, June 2008. 61

[106] M. Strolin Benedetti, R. Whomsley, and E. Baltes, "Involvement of enzymes other than CYPs in the oxidative metabolism of xenobiotics.," *Expert opinion on drug metabolism & toxicology*, vol. 2, pp. 895–921, Dec. 2006. 61

[107] C. Beedham, "The role of non-P450 enzymes in drug oxidation.," *Pharmacy world & science : PWS*, vol. 19, pp. 255–63, Dec. 1997. 61

[108] K. M. Giacomini, S.-M. Huang, D. J. Tweedie, L. Z. Benet, K. L. R. Brouwer, X. Chu, A. Dahlin, R. Evers, V. Fischer, K. M. Hillgren, K. a. Hoffmaster, T. Ishikawa, D. Keppler, R. B. Kim, C. a. Lee, M. Niemi, J. W. Polli, Y. Sugiyama, P. W. Swaan, J. a. Ware, S. H. Wright, S. W. Yee, M. J. Zamek-Gliszczynski, and L. Zhang, "Membrane transporters in drug development.," *Nature reviews Drug discovery*, vol. 9, pp. 215–36, Mar. 2010. 61

[109] D. Croft, G. O'Kelly, G. Wu, R. Haw, M. Gillespie, L. Matthews, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, S. Jupe, I. Kalatskaya, S. Mahajan, B. May, N. Ndegwa, E. Schmidt, V. Shamovsky, C. Yung, E. Birney, H. Hermjakob, P. D'Eustachio, and L. Stein, "Reactome: a database of reactions, pathways and biological processes.," *Nucleic acids research*, vol. 39, pp. D691–7, Jan. 2011. 66

[110] M. Milacic, R. Haw, K. Rothfels, G. Wu, D. Croft, H. Hermjakob, P. D'Eustachio, and L. Stein, "Annotating cancer variants and anti-cancer therapeutics in reactome.," *Cancers*, vol. 4, pp. 1180–211, Jan. 2012. 66

[111] J. Pevsner in *Bioninformatics and Functional Genomics*, ch. 8, New Jersey: Wiley, 2nd ed., 2009. 88

[112] M. Bantscheff, S. Lemeer, M. M. Savitski, and B. Kuster, "Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present.," *Analytical and bioanalytical chemistry*, vol. 404, pp. 939–65, Sept. 2012. 88

[113] S. Orchard, "Data standardization and sharing-the work of the HUPO-PSI.," *Biochimica et biophysica acta*, vol. 1844, pp. 82–7, Jan. 2014. 88

[114] Y. Guo, P. Xiao, S. Lei, F. Deng, G. G. Xiao, Y. Liu, X. Chen, L. Li, S. Wu, Y. Chen, H. Jiang, L. Tan, J. Xie, X. Zhu, S. Liang, and H. Deng, "How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes," *Acta Biochimica et Biophysica Sinica*, vol. 40, pp. 426–436, May 2008. 88, 151

[115] M. Gry, R. Rimini, S. Strömberg, A. Asplund, F. Pontén, M. Uhlén, and P. Nilsson, "Correlations between RNA and protein expression profiles in 23 human cell lines.," *BMC genomics*, vol. 10, p. 365, Jan. 2009. 88, 151

[116] H. Parkinson, U. Sarkans, N. Kolesnikov, N. Abeygunawardena, T. Burdett, M. Dylag, I. Emam, A. Farne, E. Hastings, E. Holloway, N. Kurbatova, M. Lukk, J. Malone, R. Mani, E. Pilicheva, G. Rustici, A. Sharma, E. Williams, T. Adamusiak, M. Brandizi, N. Sklyar, and A. Brazma, "ArrayExpress update–an archive of microarray and high-throughput sequencing-based functional genomics experiments.," *Nucleic acids research*, vol. 39, pp. 2010–2012, Nov. 2010. 88, 90, 132, 134

[117] T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. a. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, and A. Soboleva, "NCBI GEO: archive for functional genomics data sets–update.," *Nucleic acids research*, vol. 41, pp. D991–5, Jan. 2013. 88, 90

[118] J. Rung and A. Brazma, "Reuse of public genome-wide gene expression data," *Nature Reviews Genetics*, vol. 14, pp. 89–99, Feb. 2013. 89, 90, 93

[119] T. F. Rayner, F. I. Rezwan, M. Lukk, X. Z. Bradley, A. Farne, E. Holloway, J. Malone, E. Williams, and H. Parkinson, "MAGETabulator, a suite of tools to support the microarray data format MAGE-TAB," *Bioinformatics (Oxford, England)*, vol. 25, pp. 279–280, Jan. 2009. 95

[120] J. Malone, E. Holloway, T. Adamusiak, M. Kapushesky, J. Zheng, N. Kolesnikov, A. Zhukova, A. Brazma, and H. Parkinson, "Modeling sample variables with an Experimental Factor Ontology," *Bioinformatics (Oxford, England)*, vol. 26, pp. 1112–1118, Apr. 2010. 95

[121] J. G. Fox, S. Barthold, M. Davisson, C. E. Newcomer, F. W. Quimby, and A. Smith, "Mouse Models in Aging Research," in *The mouse in biomedical research vol.2*, American College of Laboratory Animal Medicine, pp. 637–673, Elsevier Science, 2nd ed., 2007. 97

[122] R. Asghari, M. S. Y. Lung, P. M. Pilowsky, and M. Connor, "Sex differences in the expression of serotonin-synthesizing enzymes in mouse trigeminal ganglia," *Neuroscience*, vol. 199, no. 13, pp. 429–437, 2011. 99

[123] K. S. Yip, A. Suvorov, J. Connerney, N. J. Lodato, and D. J. Waxman, "Changes in mouse uterine transcriptome in estrus and proestrus," *Biology of reproduction*, vol. 89, no. 13, pp. 1–12, 2013. 99

[124] M. Ringnér, "What is principal component analysis?," *Nature biotechnology*, vol. 26, pp. 303–4, Mar. 2008. 101

[125] M. E. Futschik and B. Carlisle, "Noise-robust soft clustering of gene expression time-course data," *Journal of bioinformatics and computational biology*, vol. 3, pp. 965–988, Aug. 2005. 102, 135, 136

[126] M. J. Kumar and J. K. Andersen, "Perspectives on MAO-B in Aging and Neurological Disease," *Molecular Neurobiology*, vol. 30, no. 1, pp. 77–89, 2004. 105

[127] E. E. Schadt, C. Molony, E. Chudin, K. Hao, X. Yang, P. Y. Lum, A. Kasarskis, B. Zhang, S. Wang, C. Suver, J. Zhu, J. Millstein, S. Sieberts, J. Lamb, D. GuhaThakurta, J. Derry, J. D. Storey, I. Avila-Campillo, M. J. Kruger, J. M. Johnson, C. a. Rohl, A. van Nas, M. Mehrabian, T. a. Drake, A. J. Lusis, R. C. Smith, F. P. Guengerich, S. C. Strom, E. Schuetz, T. H. Rushmore, and R. Ulrich, "Mapping the genetic architecture of gene expression in human liver.," *PLoS biology*, vol. 6, p. e107, May 2008. 108

[128] X. Yang, B. Zhang, C. Molony, E. Chudin, K. Hao, J. Zhu, A. Gaedigk, C. Suver, H. Zhong, J. S. Leeder, F. P. Guengerich, S. C. Strom, E. Schuetz, T. H. Rushmore, R. G. Ulrich, J. G. Slatter, E. E. Schadt, A. Kasarskis, and P. Y. Lum, "Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver.," *Genome research*, vol. 20, pp. 1020–36, Aug. 2010. 108

[129] M. Kadmiel and J. A. Cidlowski, "Glucocorticoid receptor signaling in health and disease.," *Trends in pharmacological sciences*, vol. 34, pp. 518–30, Sept. 2013. 114

[130] D. Roberts and S. Dalziel, "Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth (Review)," *Cochrane Database of Systematic Reviews*, no. 3, p. CD004454, 2013. 114

[131] S. B. Koukouritaki, P. Simpson, C. K. Yeung, A. E. Rettie, and R. N. Hines, "Human Hepatic Flavin-Containing Developmental Expression," *Pediatric research*, vol. 51, no. 2, pp. 236–243, 2002. 126

[132] R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang, "Bioconductor: open software development for computational biology and bioinformatics," *Genome biology*, vol. 5, no. 10, p. R80, 2004. 134, 135

[133] A. Kauffmann, R. Gentleman, and W. Huber, "arrayQualityMetrics - A bioconductor package for quality assessment of microarray data," *Bioinformatics*, vol. 25, no. 3, pp. 415–416, 2009. 134

[134] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, "Exploration, normalization, and summaries of high density oligonucleotide array probe level data.," *Biostatistics (Oxford, England)*, vol. 4, pp. 249–64, Apr. 2003. 134

[135] L. Gautier, L. Cope, B. M. Bolstad, and R. A. Irizarry, "affy-analysis of Affymetrix GeneChip data at the probe level," *Bioinformatics*, vol. 20, no. 3, pp. 307–315, 2004. 134

[136] M. J. Dunning, M. L. Smith, M. E. Ritchie, and S. Tavaré, "beadarray: R classes and methods for Illumina bead-based data," *Bioinformatics*, vol. 23, no. 16, pp. 2183–2184, 2007. 134

[137] M. E. Ritchie, J. Silver, A. Oshlack, M. Holmes, D. Diyagama, A. Holloway, and G. K. Smyth, "A comparison of background correction methods for two-colour microarrays," *Bioinformatics (Oxford, England)*, vol. 23, pp. 2700–2707, Oct. 2007. 134

[138] J. D. Silver, M. E. Ritchie, and G. K. Smyth, "Microarray background correction: maximum likelihood estimation for the normal-exponential convolution," *Biostatistics (Oxford, England)*, vol. 10, pp. 352–363, Apr. 2009. 134

[139] B. M. Bolstad, R. A. Irizarry, M. Astrand, and T. P. Speed, "A comparison of normalization methods for high density oligonucleotide array data based on variance and bias," *Bioinformatics (Oxford, England)*, vol. 19, pp. 185–193, Jan. 2003. 134

[140] G. K. Smyth, "Limma: linear models for microarray data," in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, and W. Huber, eds.), pp. 397–420, New York: Springer, 2005. 134, 135

[141] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995. 135, 136

REFERENCES

[142] M. Futschik, "Mfuzz: Soft clustering of time series gene expression data." http://itb.biologie.hu-berlin.de/~futschik/software/R/Mfuzz/, 2012. 135, 136

[143] S. Anders and W. Huber, "Differential expression analysis for sequence count data," *Genome Biology*, vol. 11, p. R106, 2010. 135, 136

[144] R. J. Kinsella, A. Kähäri, S. Haider, J. Zamora, G. Proctor, G. Spudich, J. Almeida-King, D. Staines, P. Derwent, A. Kerhornou, P. Kersey, and P. Flicek, "Ensembl BioMarts: a hub for data retrieval across taxonomic space.," *Database : the journal of biological databases and curation*, vol. 2011, p. bar030, Jan. 2011. 136, 198, 229, 246

[145] P. Flicek, M. R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, S. Fitzgerald, L. Gil, L. Gordon, M. Hendrix, T. Hourlier, N. Johnson, A. K. Kähäri, D. Keefe, S. Keenan, R. Kinsella, M. Komorowska, G. Koscielny, E. Kulesha, P. Larsson, I. Longden, W. McLaren, M. Muffato, B. Overduin, M. Pignatelli, B. Pritchard, H. S. Riat, G. R. S. Ritchie, M. Ruffier, M. Schuster, D. Sobral, Y. A. Tang, K. Taylor, S. Trevanion, J. Vandrovcova, S. White, M. Wilson, S. P. Wilder, B. L. Aken, E. Birney, F. Cunningham, I. Dunham, R. Durbin, X. M. Fernández-Suarez, J. Harrow, J. Herrero, T. J. P. Hubbard, A. Parker, G. Proctor, G. Spudich, J. Vogel, A. Yates, A. Zadissa, and S. M. J. Searle, "Ensembl 2012.," *Nucleic acids research*, vol. 40, pp. D84–90, Jan. 2012. 136, 198, 229, 246

[146] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones, and M. a. Marra, "Circos: an information aesthetic for comparative genomics.," *Genome research*, vol. 19, pp. 1639–45, Sept. 2009. 136

[147] The Inkscape Team, *Inkscape project*. Inkscape project, 2004. 136

[148] S. Mathew, "Ketamine for treatment-resistant unipolar depression: current evidence," *CNS Drugs*, vol. 26, no. 3, pp. 189–204, 2013. 143

[149] I. Coupry, V. Lachaud, R.-A. Podevin, E. Koenig, and A. Parini, "Different Affinities of α2-Agonists for Imidazoline and α2-Adrenergic Receptors," *American Journal of Hypertension*, vol. 2, no. 6 Pt 1, pp. 468–470, 1989. 143

[150] D. J. Posey and C. J. McDougle, "Guanfacine and guanfacine extended release: treatment for ADHD and related disorders.," *CNS drug reviews*, vol. 13, pp. 465–74, Jan. 2007. 143

[151] M. K. Mulligan, X. Wang, A. L. Adler, K. Mozhui, L. Lu, and R. W. Williams, "Complex control of GABA(A) receptor subunit mRNA expression: variation, covariation, and genetic regulation.," *PloS one*, vol. 7, p. e34586, Jan. 2012. 143, 165

[152] World health Organisation, *Promoting Safety of Medicines for Children*. Geneva: WHO press, 2007. 149

[153] H. S. Jensen, K. Nichol, D. Lee, and B. Ebert, "Clobazam and its active metabolite N-desmethylclobazam display significantly greater affinities for α2- versus α1-GABA(A)-receptor complexes.," *PloS one*, vol. 9, p. e88456, Jan. 2014. 152, 164

[154] V. V. Ragavan, *Drug Development For Women*. Chichester, England: Wiley, 1998. 154

[155] E. B. Weller, A. Kloos, J. Kang, and R. A. Weller, "Depression in children and adolescents: does gender make a difference?," *Current psychiatry reports*, vol. 8, pp. 108–14, Apr. 2006. 154

[156] A. J. Ferrari, F. J. Charlson, R. E. Norman, S. B. Patten, G. Freedman, C. J. L. Murray, T. Vos, and H. a. Whiteford, "Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010.," *PLoS medicine*, vol. 10, p. e1001547, Nov. 2013. 154

[157] H. Whitley and W. Lindsey, "Sex-based differences in drug activity.," *American family physician*, vol. 80, pp. 1254–8, Dec. 2009. 154

[158] H. J. Kang, Y. I. Kawasawa, F. Cheng, Y. Zhu, X. Xu, M. Li, A. M. M. Sousa, M. Pletikos, K. a. Meyer, G. Sedmak, T. Guennel, Y. Shin, M. B. Johnson, v. Krsnik, S. Mayer, S. Fertuzinhos, S. Umlauf, S. N. Lisgo, A. Vortmeyer, D. R. Weinberger, S. Mane, T. M. Hyde, A. Huttner, M. Reimers, J. E. Kleinman, and N. Šestan, "Spatio-temporal transcriptome of the human brain," *Nature*, vol. 478, pp. 483–489, Oct. 2011. 155

[159] N. J. Guzman, "Epidemiology and management of hypertension in the Hispanic population: a review of the available literature.," *American journal of cardiovascular drugs : drugs, devices, and other interventions*, vol. 12, pp. 165–78, June 2012. 156

[160] M. C. S. Wong, W. W. S. Tam, C. S. K. Cheung, E. L. H. Tong, A. C. H. Sek, G. John, N. T. Cheung, B. P. Y. Yan, C. M. Yu, S. Leeder, and S. Griffiths, "Initial antihypertensive prescription and switching: a 5 year cohort study from 250,851 patients.," *PloS one*, vol. 8, p. e53625, Jan. 2013. 156

[161] N. Saijo, "The role of pharmacoethnicity in the development of cytotoxic and molecular targeted drugs in oncology.," *Yonsei medical journal*, vol. 54, pp. 1–14, Jan. 2013. 156

[162] K. Haug, R. M. Salek, P. Conesa, J. Hastings, P. de Matos, M. Rijnbeek, T. Mahendraker, M. Williams, S. Neumann, P. Rocca-Serra, E. Maguire, A. González-Beltrán, S.-A. Sansone, J. L. Griffin, and C. Steinbeck, "MetaboLights–an open-access general-purpose repository for metabolomics studies and associated meta-data.," *Nucleic acids research*, vol. 41, pp. D781–6, Jan. 2013. 157

[163] R. Goodacre, D. Broadhurst, A. K. Smilde, B. S. Kristal, J. D. Baker, R. Beger, C. Bessant, S. Connor, G. Capuani, A. Craig, T. Ebbels, D. B. Kell, C. Manetti, J. Newton, G. Paternostro, R. Somorjai, M. Sjöström, J. Trygg, and F. Wulfert, "Proposed minimum reporting standards for data analysis in metabolomics," *Metabolomics*, vol. 3, pp. 231–241, Aug. 2007. 157

[164] D. Lacroix, M. Sonnier, A. Moncion, G. Cheron, and T. Cresteil, "Expression of CYP3A in the human liver–evidence that the shift between CYP3A7 and CYP3A4 occurs immediately after birth.," *European journal of biochemistry / FEBS*, vol. 247, pp. 625–34, July 1997. 162

[165] A. C. Fitzgerald, B. T. Wright, and S. A. Heldt, "The behavioral pharmacology of zolpidem: evidence for the functional significance of α1-containing GABA(A) receptors.," *Psychopharmacology*, vol. 231, pp. 1865–96, May 2014. 164

[166] "eHospital." http://www.my-ehospital.org. 164

[167] D. T. Odom, R. D. Dowell, E. S. Jacobsen, W. Gordon, T. W. Danford, K. D. MacIsaac, P. A. Rolfe, C. M. Conboy, D. K. Gifford, and E. Fraenkel, "Tissue-specific transcriptional regulation has diverged significantly between human and mouse.," *Nature genetics*, vol. 39, pp. 730–2, June 2007. 164

[168] K. Stefflova, D. Thybert, M. D. Wilson, I. Streeter, J. Aleksic, P. Karagianni, A. Brazma, D. J. Adams, I. Talianidis, J. C. Marioni, P. Flicek, and D. T. Odom, "Cooperativity and rapid evolution of cobound transcription factors in closely related mammals.," *Cell*, vol. 154, pp. 530–40, Aug. 2013. 164