



Renal Gene Ontology Annotation

## Renal GO Annotation Initiative

*providing full GO annotation to genes associated with renal development and disease processes*

Newsletter May 2009

On 1<sup>st</sup> April 2009 Dr. Yasmin Alam-Faruque started as the new curator in the GOA team at the European Bioinformatics Institute (EBI, Cambridge UK). She will be project managing the Renal Gene Ontology Annotation Initiative; a new three year project funded by Kidney Research UK, led by Dr. Rolf Apweiler, Principal Investigator of the Protein and Nucleotides Database group at the EBI, and Professor Peter Scambler, a renal and cardiovascular expert at the Institute of Child Health (London UK). This project will provide a unique public resource for renal research by generating a detailed functional dataset for mammalian gene products implicated in renal development and disease.

### Developments so far:

#### New Renal Interest Group mailing list

A Renal Interest Group mailing list on the Gene Ontology Consortium web pages (<http://www.geneontology.org/GO.interests.shtml?all#renal>) has now been set up, so please feel free to subscribe and use it to provide suggestions/ advice/ discussions on renal gene/ protein-related issues. Information about the renal project can also be viewed at <http://www.geneontology.org/GO.renal.shtml>.

#### Renal target list

A list of over 680 renal-related proteins has been created which will act as an initial set of curation targets and this can be viewed at: <http://www.ebi.ac.uk/GOA/kidney/>.

The preliminary target list of proteins has been comprehensively analyzed and any duplicate accessions, gene names or protein names were removed. Also any proteins with a UniProtKB/TrEMBL accession number were replaced with the accession of the associated parent UniProtKB/SwissProt. The resulting list has been tagged with the term "KRUK priority" in the GOA group's Protein2GO annotation tool and also in the publicly available QuickGO browser (<http://www.ebi.ac.uk/QuickGO>). In terms of annotation, this will allow easy visibility of this dataset and any overlapping proteins on parallel Cardiovascular and Reference Genome GO annotation projects, as well as the creation of gene association files relating to this project. This will also enable members of the medical research community to search and view the dataset easily in the GOA group's QuickGO browser (<http://www.ebi.ac.uk/QuickGO/GAnnotation?protein=KRUK>).

This list has also been checked against the Reactome database at the EBI, from which it was discovered that 165 renal-related proteins, from a range of reactions and pathways, are present in



Reactome. This has led to a potential collaboration with the Reactome team to include more proteins involved in renal-specific processes into the Reactome database thereby providing an even more unique and comprehensively detailed functional dataset for mammalian gene products implicated in renal development and disease.

It is planned that the list of priority renal proteins will be expanded using the PRIDE database at the EBI, which holds experimental mass spectrometry proteomic data, from experiments associated with the HEK-293 cell-line, kidney tumour cell lines, renal glomerulus, and polycystic kidney autosomal recessive samples.

#### Outreach/Meetings

I have attended a 1 day symposium “A Half-Century of Renal Tubular Disease” on 27<sup>th</sup> March 2009 at the Sir William Wells Atrium, Royal Free Hampstead NHS Trust in London, held in celebration of the 50th anniversary of Professor Oliver Wrong’s seminal paper of April 1959 on renal tubular acidosis.

I also attended the 3<sup>rd</sup> International Meeting for Biocuration held in Berlin 16-19<sup>th</sup> April 2009. This led to the planning of further meetings with the Edinburgh group of GUDMAP (The GenitoUrinary Development Molecular Anatomy Project), an outcome of a visit to the GUDMAP poster.

#### Gene Annotation

The initial approach to this project has been taken to begin annotating a list of proteins involved in a particular process/pathway involved in a particular kidney function. This idea was the result of attending the 1 day symposium on renal tubular acidosis. Therefore, the initial annotation targets are highly investigated proteins involved in the proton pump process and ammonium transport; both processes being important for the maintenance of acid-base homeostasis in the kidney. The gene list includes *SLC4A1*; *SLC26A5*; *SLC26A4*; *AQP2*; *OCLR1*; *CLNC5*; *TSHZ3*; *RHBG*; *RHCG*; *RHAG*; *CA11*; *SLC12A3*; *VPH1*; *STV1*; *TFP1*; *VMA1*; *PFKM*; *ATP6V0A4*; *ATP6V0D1*; *ATP6V0D2*; *ATP6V1D*; *ATP6V1F*.

To date, this initiative has associated over 286 GO terms to almost 98 proteins, of which 22 are human proteins. Additionally, 2 prioritized genes: *RHBG* and *RHCG* have been comprehensively annotated using GO terms.

#### Gene Ontology Development

So far 2 new terms have been added to GO to describe the function of a particular set of proteins involved in ammonium transport: *i.e.*: GO:0070633 transepithelial transport and GO:0070634 transepithelial ammonium transport.

#### **Call for contributions from the Renal research community**

If you have a specific gene/ protein of interest that is not on the KRUK priority list, or you would like a particular one to be prioritized for GO annotation then please do let me know. Similarly, let me know if you are aware of any large dataset or are interested in a particular set of proteins that are involved in a particular kidney function/ development/ disease pathway requiring annotation.

If you know of anyone in the clinical/scientific research community working on any aspect involving the genitourinary tract who you feel would be interested in this effort, please could you forward this newsletter onto them - your assistance with this would be greatly appreciated.

---

Editor:  **Dr. Yasmin Alam-Faruque**

