**Comprehensive Gene Ontology annotation of ciliary genes in the laboratory mouse**

Karen R. Christie and Judith A. Blake; The Jackson Laboratory

**ABSTRACT**

Interest in primary cilia has increased dramatically over the last ten years as it has become clear that ciliopathies are an underlying cause of numerous human diseases including some types of retinitis pigmentosa and polycystic kidney disease. Once thought to be restricted to a few cell types, it is now clear that primary cilia are found on almost all vertebrate cells and are critical to Sonic hedgehog (Shh) signaling. Mouse models play a key role in developing our understanding of the role of primary cilia in control of Shh signaling in development throughout the embryo and in ongoing maintenance of structures such as photoreceptors.

To maximize the utility of the wealth of experimental data generated by these mouse ciliopathy models, we have initiated a project to comprehensively annotate ciliary genes of the laboratory mouse using Gene Ontology (GO) terms to describe their molecular functions, biological roles, and cellular locations. We are guided by the SysCilia gold standard of known human ciliary components as a starting point, but will also include additional genes experimentally shown to be involved in ciliary function in the mouse. If needed, we will also update the Gene Ontology to add new terms representing recent advances in our understanding of ciliary biology. Comprehensive GO annotation of ciliary genes in the mouse will be a great resource to those doing high throughput studies or comparative genomics analysis across species, and may help us better understand the similarities and differences between mouse and human in the role of Sonic hedgehog signaling in development.

This work is funded by HG 002273 to the Gene Ontology Consortium.

**Improving the representation of cilia in the Gene Ontology**

Prior to the rearrangement, the term "intraflagellar transport particle" was not directly connected to the term "cilium". Now, the new term "cilium" is related to the term "intraflagellar transport particle" by a direct link that indicates that a normal cilium must contain an intraflagellar transport particle. In addition, changes were made to unify terms relevant to eukaryotic flagella and cilia, including changing term names to use forms of the word "cilium" to refer to all eukaryotic flagella and cilia, making it much easier to identify the correct terms for eukaryotic flagella and cilia, as distinct from terms for bacterial flagella.

In this ontology development work, we contributed to a project led by Paola Roncaglia (EBI; Gene Ontology Consortium) working with John van Dam (Radboud UMC; SysCilia Consortium) to improve the representation of cilia within the Cellular Component section of the Gene Ontology.

**Identification of the mouse equivalents of the human genes in SysCilia Gold Standard version 1**

297 mouse genes with 1:1 homology with a human gene
- 263 – both Panther and HomoloGene
- 29 – HomoloGene only
- 4 – Panther only

10 mouse genes with 2 or more homologies to a human gene
- 6 – 2 mouse to 1 human
- 4 – 4 mouse to 1 human

No clear mouse homolog to 1 human gene (SLC47A2)
- 2 human genes in the same Panther family (P7H11206)
  - only HsSLC47A2 on SysCilia list
  - 2 mouse genes more closely related to HsSLC47A2 than to HsSLC47A2

**Prioritizing curation of mouse genes by paucity of annotations and availability of literature**

**Availability of literature for 187 candidate ciliary genes in mouse lacking any annotation to ciliary GO terms**

- 187 genes
- 187 genes with other literature
- 76 genes
- 76 genes with no papers selected for GO curation
- 81 genes
- 81 genes with ciliary literature
- 30 genes
- 30 genes with no papers selected for GO curation

**Plan:**

To maximize the our annotation effort, we are focusing on mouse genes that are not currently annotated or only partially annotated with ciliary terms in the Cellular Component (CC) or Biological Process (BP) sections of the ontology, based on the availability of literature for each gene. However, we will make additional annotations for previously annotated genes when we come across supporting experimental data.

**Priorities:**

- 30 genes – no ciliary annotations; ciliary literature available
- 76 genes – no ciliary annotations; other literature available
- 18 genes – some ciliary annotations; ciliary literature available
- 36 genes – some ciliary annotations; other literature available
- 119 genes – no literature available
- 28 genes – already annotated with ciliary terms

**Summary:**

Our goal is to achieve comprehensive annotation of ciliary genes in the laboratory mouse based on experimental data by mid 2015. Based on our early results, we were able to annotate some genes that were not associated in the Mouse Genome Database (MGD) with literature tagged to be curated for Gene Ontology. In some cases, this was due to publication of new papers after acquisition of the initial data set. However, more frequently, we encountered papers containing a single experiment with data for multiple genes, but where many of these genes were not indexed to the paper in MGD as they were not a major focus of the paper. When we are finished, it will be interesting to look for conservation, or changes, in function between mouse and human ciliary genes, utilizing annotations by the SysCilia Consortium for human genes.

**Progress (Oct ‘13 – Jan ‘14):**

- Completed: 7 genes
- New exp. Annotations: 6 genes

**Genes with ciliary literature**

- 18 genes
- 13 genes
- 6 genes

**Genes with ciliary literature**

- 36 genes
- 6 genes
- 5 genes
- 1 gene