**Paul Medvedev**
**Assistant Professor**
Department of Computer Science and Engineering
Department of Biochemistry and Molecular Biology
The Pennsylvania State University

506B Wartik Lab
University Park, PA 16802-5807
pzm11@psu.edu
814-865-0193

## LEGENDS FOR SUPPLEMENTARY DATA FILES

**Supplementary Dataset 1**. Splice coordinates (chrY) of the 668 alignments from the 168 known database transcripts used as a database of known sites. Each row corresponds to a distinct alignment, and transcripts with multiple alignments have multiple rows. An empty splice coordinate column means that the transcript had no splice junction.

**Supplementary Dataset 2**. Sequences of IsoCon transcripts, their alignments, and the gene family graphs.