**Description of Supplemental Data**

**Redefinition of the human mast cell transcriptome by deep-CAGE sequencing**

Efthymios Motakis,1,\* Sven Guhl,2,\* Yuri Ishizu,1 Masayoshi Itoh,1,3 Hideya Kawaji,1,3 Michiel de Hoon,1 Timo Lassmann,1 Piero Carninci,1 Yoshihide Hayashizaki,3 Torsten Zuberbier,2 Alistair R R Forrest,1¶ Magda Babina2¶ and the FANTOM consortium1

\*These authors contributed equally

1RIKEN Centre for Life Science Technologies, Division of Genomic Technologies, Yokohama, Kanagawa, 230-0045 Japan

2Department of Dermatology and Allergy, Charité Campus Mitte, Berlin, Germany

3RIKEN Preventive Medicine and Diagnosis Innovation Program

¶To whom correspondence should be addressed

The dataset supporting material consists of the following kind of information:

* Worksheet “DPI Counts – 49 samples”: It includes the raw count data for all 184,827 DPI regions (Column A contains their location and column B their annotation) for 49 blood related samples extracted from the FANTOM5 data collection. Columns C - K contain the raw counts of the Mast Cell samples. Columns L - N contain the raw counts of the Basophils samples. The rest of the columns contain the raw counts of other blood related samples.
* Worksheet “Mean RLE Counts (RefSeq genes)”: It summarizes the Relative Log Expression (RLE) based normalized counts by the means of replicates of Ex-vivo, Expanded, Expanded & Stimulated (Mast Cells), Basophils, FANTOM5 without Mast Cells, FANTOM5 without Basophils and all FANTOM5. The average values are estimated for all genes that are expressed in at least one of the above conditions. Column A contains the RefSeq gene annotation. Columns B - C contain the associated promoter location(s) and annotation(s). The rest of the columns contain the average RLE normalized expression levels.
* Worksheet “Mean RLE Counts (RefSeq promot)”: It summarizes the Relative Log Expression (RLE) based normalized counts by the means of replicates of Ex-vivo, Expanded, Expanded & Stimulated (Mast Cells), Basophils, FANTOM5 without Mast Cells, FANTOM5 without Basophils and all FANTOM5. The average values are estimated for all promoters of the expressed genes. Column A contains the RefSeq promoter annotation. Column B contains the promoter location. The rest of the columns contain the average RLE normalized expression levels.
* Worksheet “Data Labels”: It includes the FANTOM5 data labels of all samples used in this analysis.