

Metrics for the Human Proteome Project 2016: Progress on Identifying and Characterizing the Human Proteome, Including Post-Translational Modifications

Gilbert S. Omenn,^{×} Lydie Lane,[∞] Emma K. Lundberg,^μ Ronald C. Beavis,^α*

Christopher M. Overall^φ, and Eric W. Deutsch^π

[×] Department of Computational Medicine and Bioinformatics

[∞] CALIPHO Group, SIB Swiss Institute of Bioinformatics and Department of Human Protein
Science

^μ SciLifeLab Stockholm and School of Biotechnology, KTH

^α Biochemistry & Medical Genetics, University of Manitoba

^φ Overall Lab, University of British Columbia

^π Institute for Systems Biology

SUPPORTING INFORMATION

TABLE OF CONTENTS

Supplementary Table 1. The 485 proteins that qualify for PE1 in neXtProt under guidelines version 1.0 which have been excluded in neXtProt 2016-02 under HPP guidelines version 2.0.

Supplementary Table 2. The 2949 PE2,3,4 missing proteins in neXtProt 2016-02.

Supplementary Table 3. Chromosome-by-chromosome list of numbers of neXtProt entries by PE level.

Supplementary Table 4. Human Protein Atlas 2016-05 list of the 879 testis-specific or testis-enriched proteins.

Supplementary Figure 1. Circular chart showing proportions of protein-coding genes on each chromosome by neXtProt PE level.

Supplementary Tables and Figure Legends

Supplementary Table 1: neXtProt 2016-02 list of the 485 Proteins that would have qualified for PE 1 under the previous guidelines, but are excluded from PE 1 now under the new Guidelines v2.0; these 485 proteins now are classified under candidate missing proteins (PE2,3,4).

Supplementary Table 2: Complete list of the 2949 neXtProt “missing proteins” PE2,3,4 as of 2016-02. This list includes the 485 in Supplementary Table 1. It matches the red wedge in Figure 2.

Obtaining a list of missing proteins: An up-to-date listing of the missing proteins may be extracted at neXtProt as follows:

- 1) Go to <http://www.nextprot.org/>
- 2) Click on Advanced Search (<https://search.nextprot.org/>)
- 3) Click on the Advanced Search radio button
- 4) In the search box type “PE=2” to see and select saved query NXQ_00204
- 5) Click [Download] to download the whole list

An up-to-date listing of the missing proteins may be extracted at PeptideAtlas as follows:

- 1) Go <http://www.peptideatlas.org/>
- 2) Click on the large C-HPP icon
- 3) In the second table, click on the number of neXtProt entries in the PE=2 row
- 4) Click on the [Form] tab
- 5) Hold down the CONTROL key (or COMMAND key on Apple) and click on the Protein Evidence Level constraint PE=3 and PE=4 (thereby selecting all three PE=2, PE=3, PE=4) and click [QUERY]
- 6) At the bottom of the resulting page, click on [Excel] or a similar button to download the whole list

Two authors each confirmed that these instructions work as of 25 July.

Supplementary Table 3: Chromosome-by-Chromosome list of numbers of neXtProt protein entries by PE level [download from neXtProt.org].

Supplementary Table 4: Human Protein Atlas 2016-05 list of the 879 testis-specific or testis-enriched proteins. Altogether 854 have UniProt IDs, including three with multiple IDs, while 354 are canonical in PeptideAtlas 2016-01. HPA data kindly provided by Asa Sievertsson of HPA, based on v15 (updated from v14 in Djureinovic et al, 2015).

Supplementary Figure

Supplementary Figure 1. Circular chart showing proportions of protein-coding genes on each chromosome (including Y and Mitochondria) that are PE1, PE2, PE3, PE4, or PE5 in neXtProt. The corresponding table with current numbers (April 2016) for each level for each chromosome is Supplementary Table 2.

