The American Journal of Human Genetics, Volume *91*

Supplemental Data

TMEM165 Deficiency Causes a Congenital Disorder of Glycosylation

François Foulquier, Mustapha Amyere, Jaak Jaeken, Renate Zeevaert, Els Schollen, Valérie Race, Riet Bammens, Willy Morelle, Claire Rosnoblet, Dominique Legrand, Didier Demaegd, Neil Buist, David Cheillan, Nathalie Guffon, Pierre Morsomme, Willem Annaert, Hudson H. Freeze, Emile Van Schaftingen, Miikka Vikkula, and Gert Matthijs

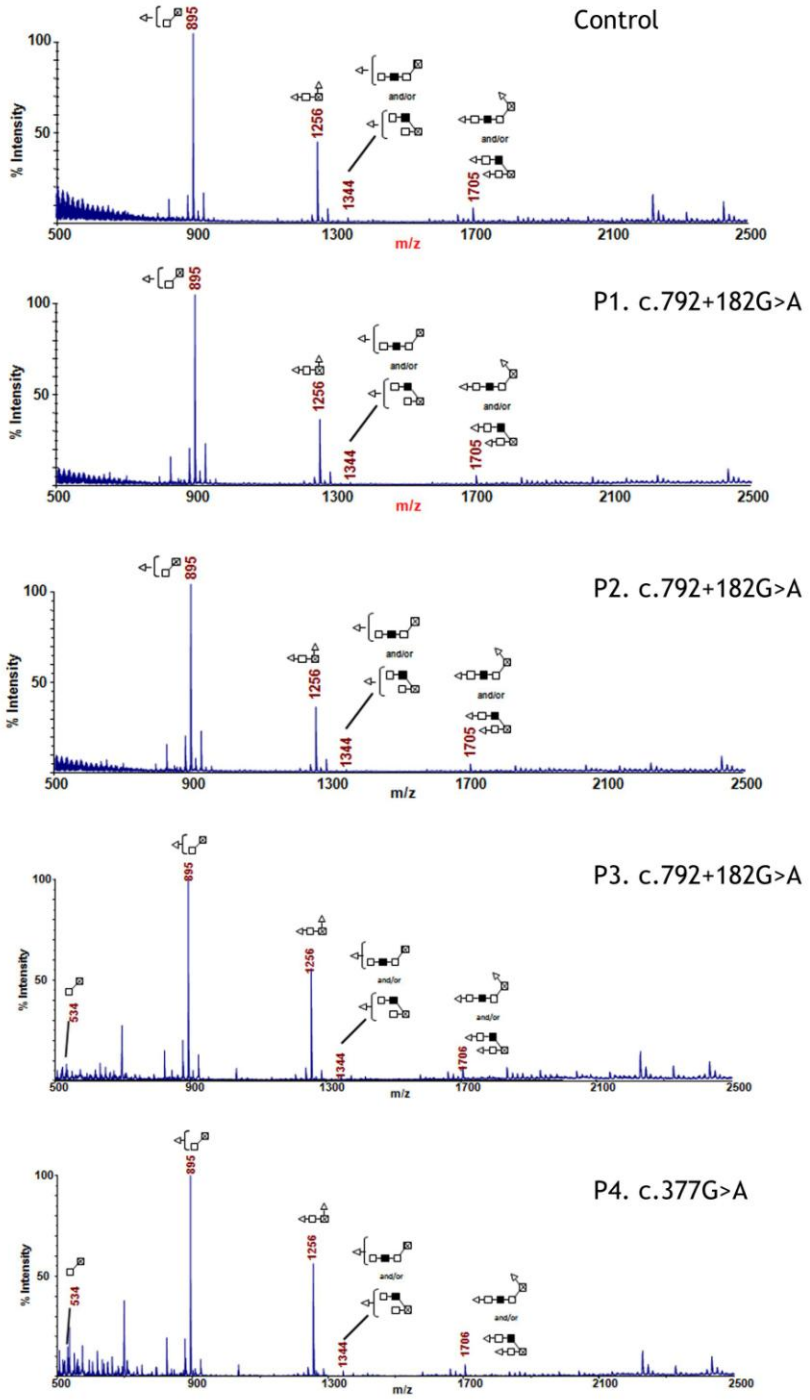

Figure S1. MALDI-TOF-MS Spectra of the Permethylated O-Glycans from Sera of Control and TMEM165-Deficient Individuals

The permethylated derivatives were analyzed in positive-ion reflective mode, as $[M+Na]^+$. Symbols: open square, galactose; closed square, N-acetylglucosamine; crossed square, reduced N-acetylgalactosamine; open triangle, N-acetyl neuraminic acid.

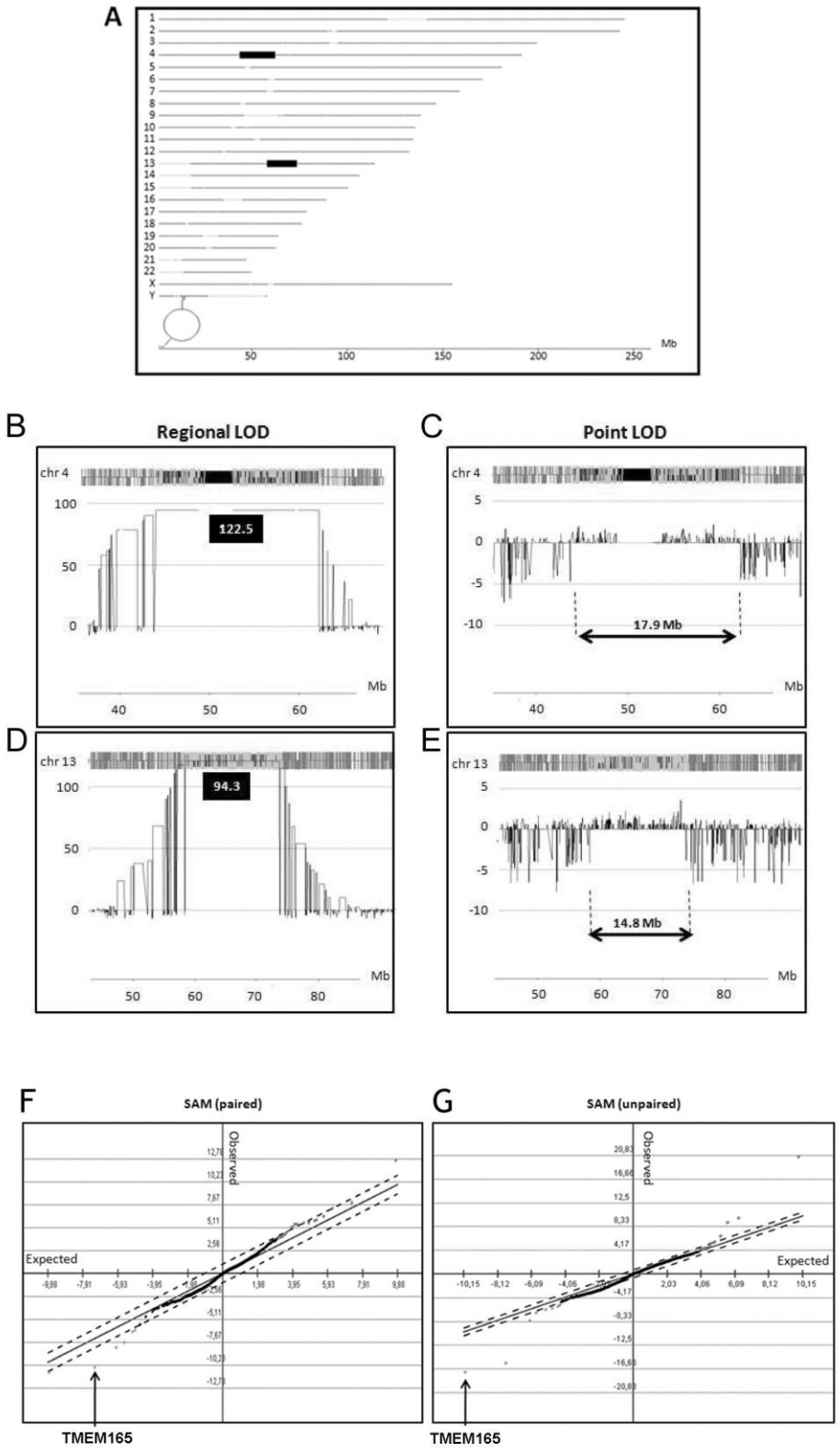

Figure S2. Mapping and Gene identification

(A) Whole-genome view showing regional LOD scores above 80, based on a control population. Two haplotype blocks presented with black boxes in chr4 and in chr13.

(B and D**)** The highest regional logarithm of odds (LOD) score is 94.3 in chr4 and 122.5 in chr13.

(C and E) Point LOD score plots indicate an autozygous region in chr4 delimited by single-nucleotide polymorphism (SNPs) rs6828553 and rs1505663 between 44149314 bp and 62131840 bp, and an autozygous region in chr13 delimited by rs10507629 and rs9318228 from position 59714684 bp to

74500960 bp. Vertical bars, SNPs. Autozygosity mapping calculated by Genespring GT v2 package (Agilent).

(F and G) Significance Analysis of Microarrays (SAM): Scatter plots showing the observed relative difference of expression between two affected individuals and three controls, compared to that determined by permutation analysis. Significantly underexpressed genes locate in lower left of each plot and significantly overexpressed genes locate in upper right. Paired analysis, direct comparison between each affected member and control (F); unpaired analysis with significantly different mean expression levels of genes between affected and control samples in (G). Significantly overexpressed genes are located in upper right and significantly underexpressed genes locate in lower left. *TMEM165* gene is highlighted in black box. 31 genes in paired and 20 genes unpaired analysis were significantly changed. Parameters: (fold change = 1.5x ; FDRpaired & unpaired = 5; deltapaired = 1.03 ; deltaunpaired = 0.66. SAM analysis was performed with MultiExperiment Viewer package v4.5.

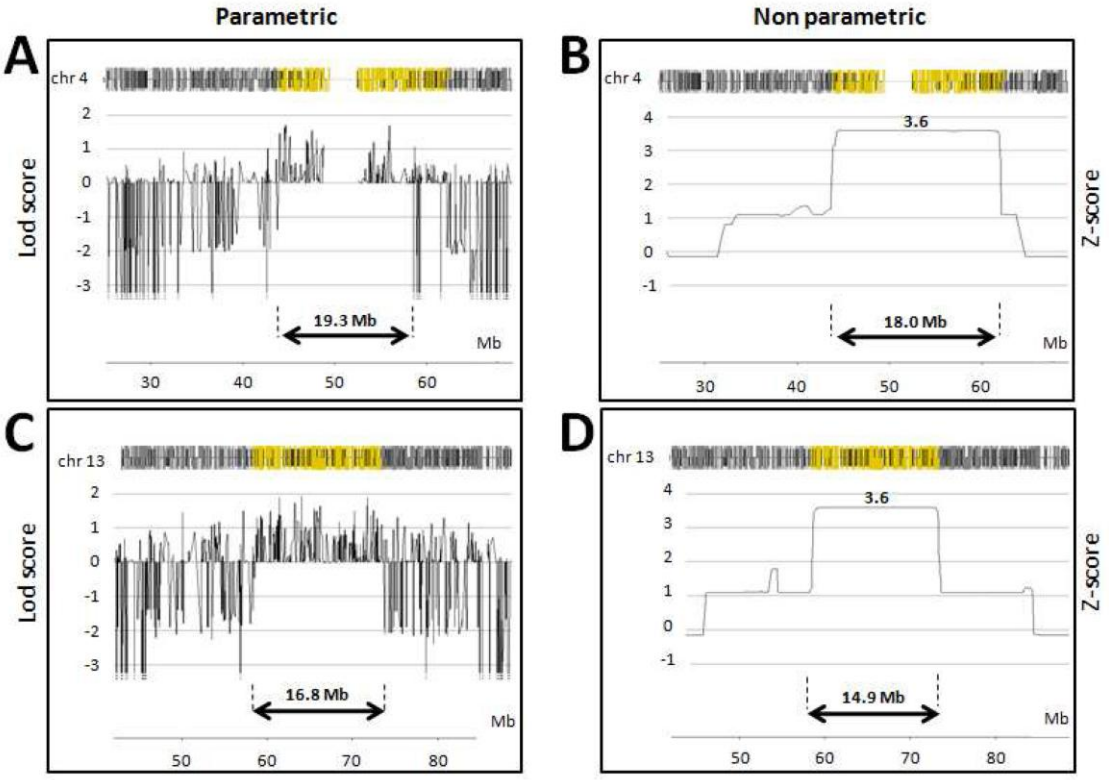

Figure S3. Multipoint Linkage Studies

(A and C) Parametric linkage analysis plots. Positive LOD score region between rs720961 and rs335337 at position 43054407 bp and 62332299 bp in chr4 and between rs9316935 and rs2184853 at position 58019973 bp and 74841547 bp in chr13.

(B and D) Non parametric linkage with max Z-score 3.6, between rs6828553 and rs2342717 in chr4 and between rs9317052 and rs1536025 in chr13, respectively.

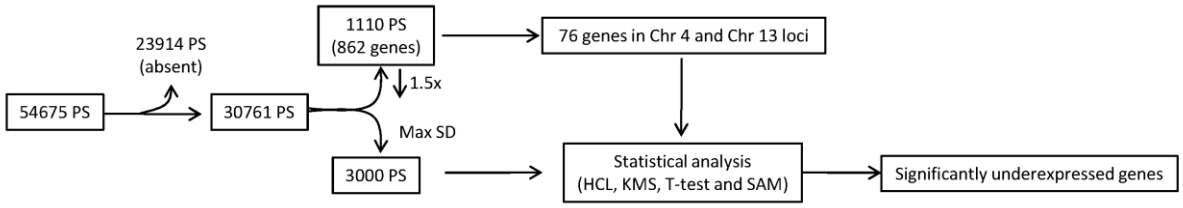

Figure S4. Statistical Approach Used to Analyze Expression Data

Detection call (Present/Absent) generated by Affymetrix microarray suite version 5 software (MAS5) was used to remove data not reliably detected in all samples before further analysis: 23914 probe sets (PS) discarded form initial 54675 PS. Two subsequent filtering methods were employed:

(A) 1.5x fold change between affected members and controls: 862 significant genes were selected of which only 76 were located in the autozygous and linked regions in chr4 and chr13.

(B) 3000 PS selected by maximum standard deviation in all samples. Our initial hypothesis was that in the affected individuals, the causative mutation, as is common, in metabolic diseases, causes loss of function of the implicated protein. This is often due to a premature stop codon (PSC) mutation, which leads to mRNA instability via PSC mediated mRNA decay. Therefore, statistical analyses were performed on the prefiltered data to select significantly underexpressed genes in affected samples as compared to controls.

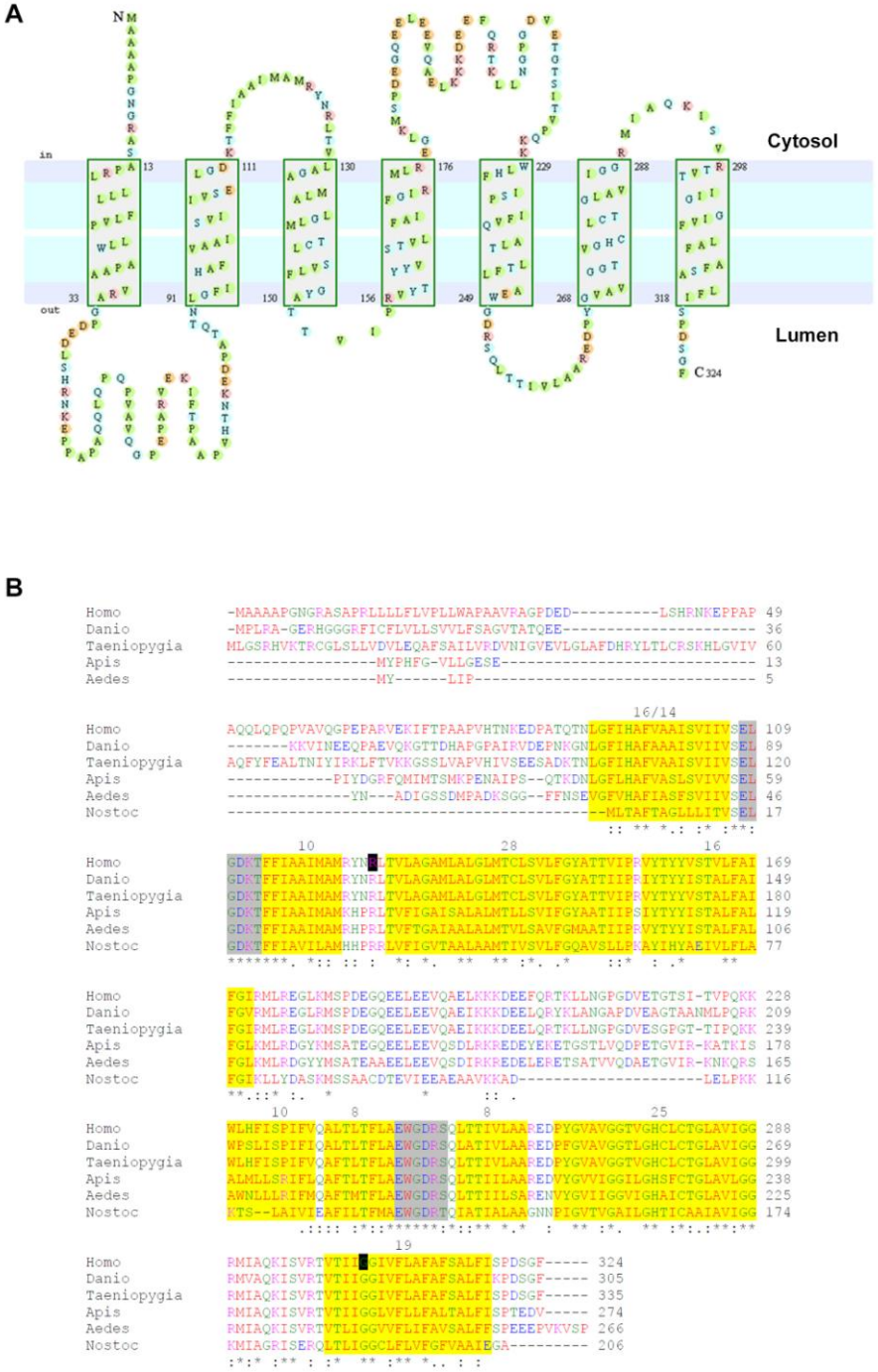

Figure S5. TMEM165 Protein

(A) Schematic representation of the predicted membrane spanning topology. The putative signal peptide is represented in blue. The changed amino acid residues and premature stop codon as well as the missense mutation are indicated (in red). Membrane spanning helices were predicted using the TMHMM software17.

(B) Alignment of Human TMEM165 with other proteins of the same family. The following sequences are shown : *Homo sapiens :* NP_060945.2; *Taeniopygia guttata* : XP_002192775.1; *Danio rerio* : NP_997848.1; *Apis mellifera* : XP_623837.1; *Aedes aegyptii* : XP_001658730.1; *Nostoc punctiforme* pcc 73102 : YP_001866240.1. The strictly conserved EXGDK/RT/S motif is highlighted in grey and the hydrophobic regions are highlighted in yellow. The residues that are strictly conserved are indicated with an asterisk below the alignment, whereas the semi-conserved and conserved residues are indicated with 1 or 2 dots, respectively. The mutated residue in case 4 is indicated with a black box.

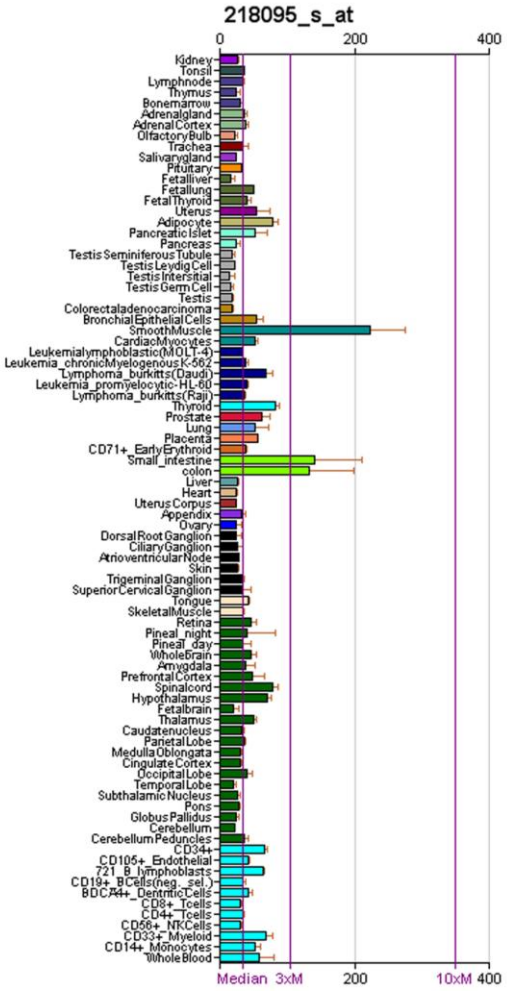

Figure S6. The Expression Profile of Human *TMEM165* Transcript in Various Tissues

The data is taken from http://symatlas.gnf.org/Symatlas/. The human GNF1H gene chip and MAS5 algorithm were used for the analysis of *TMEM165* transcript.