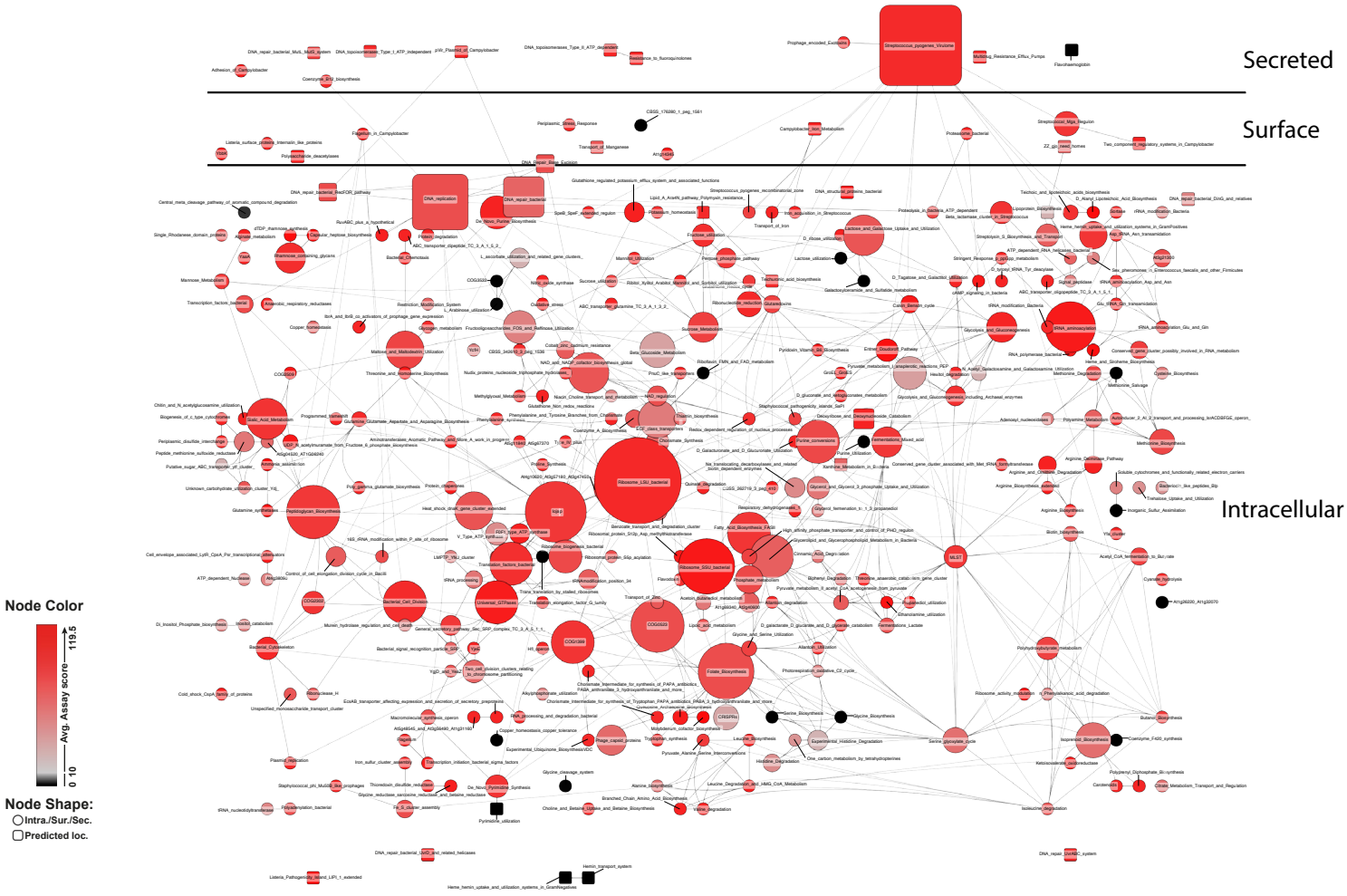
Supplementary Information
Proteome-wide selected reaction monitoring assays for the human pathogen *Streptococcus pyogenes*

Christofer Karlsson[1, #], Lars Malmström[2, #], Ruedi Aebersold[2, 3], Johan Malmström[1,4 *]

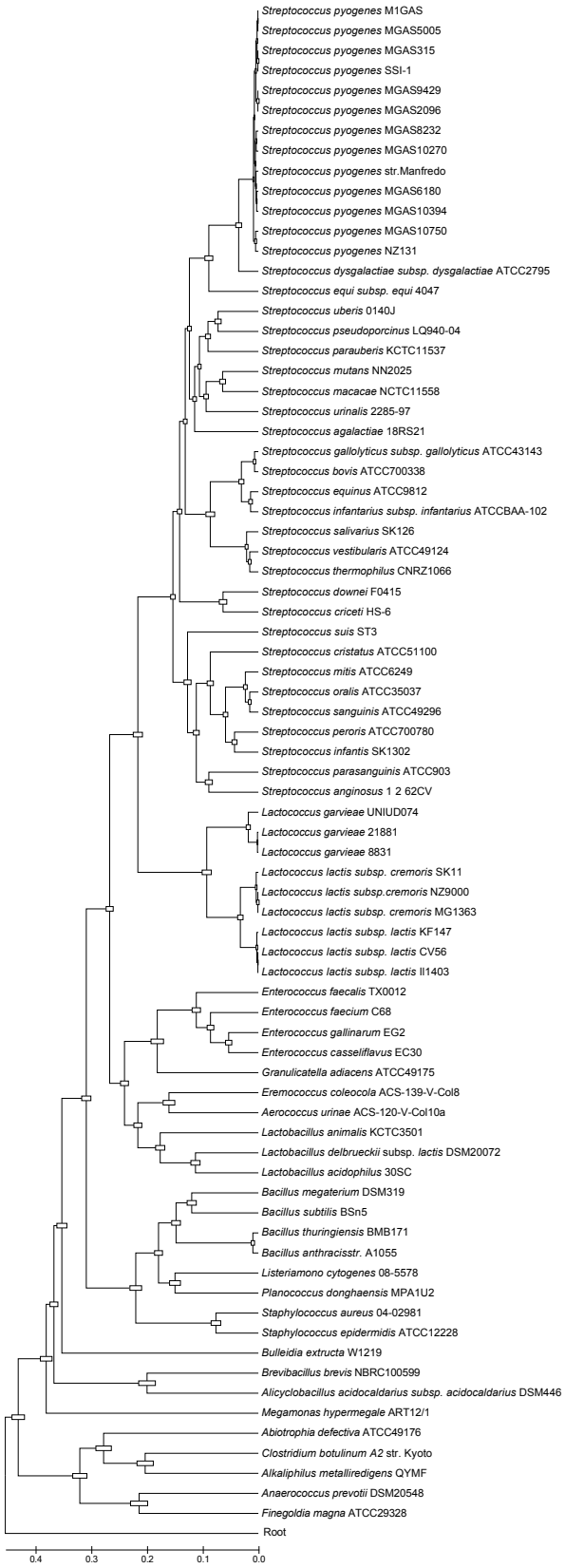
[1)] Department of Immunotechnology, Lund, Sweden
[2)] Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland
[3)] Faculty of Science, University of Zurich, Zurich, Switzerland
[4)] Biognosys AG, Zurich, Switzerland
[#)] Equal author contribution

Supplementary Figure S1 | Outline of the GAS proteome network topology and assay score.



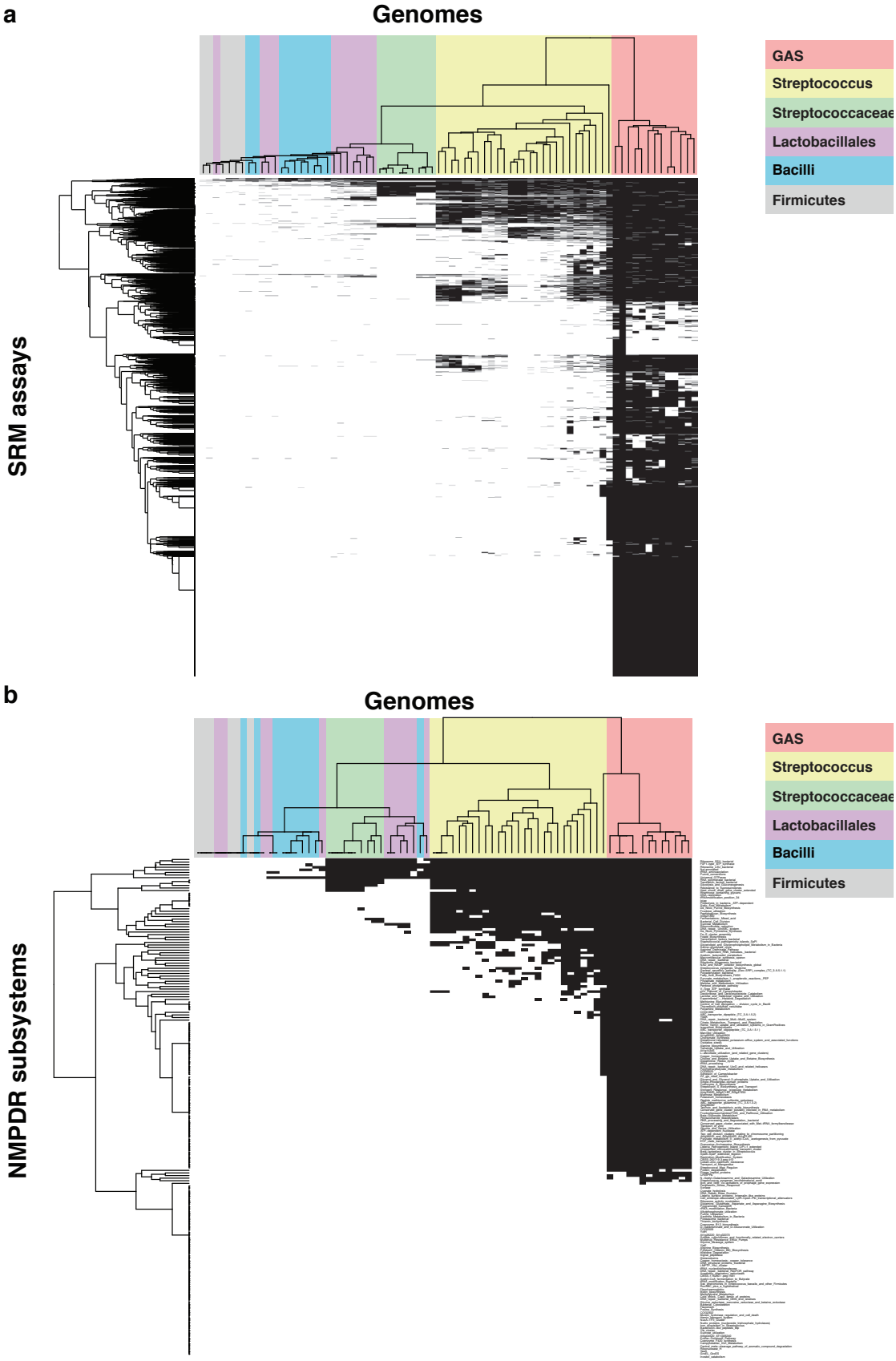
Circles represent NMPDR subsystems where all proteins predominantly have the same subcellular location, secreted, surface associated or intracellular, according to the subcellular protein profiles in Figure 4a-d. Rectangles represent NMPDR subsystems where an equal number of members have opposing subcellular location profiles. The localization of the rectangles in the network is influenced by the edges, which represent protein members that belong to more than one NMPDR subsystem. Increasing node size represents increasing number of member proteins. The color represents average SRM assay score, where red indicates NMPDR subsystems with high-average SRM assay score and black indicating NMPDR subsystems with low average SRM assays score.

Supplementary Figure S2 | Phylogenetic relationships among selected species in the Firmicutes phylum



A maximum likelihood phylogenetic tree for selected bacterial species based on the Tamura-Nei model[49] was built upon an alignment of nucleotide sequences of *rpoB*[50] extracted from respective genome (downloaded from PATRIC Bacterial Bioinformatics Resource Center[28]). The tree with the highest log likelihood (-76672.8395) is shown. Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was < 100 or less than one fourth of the total number of sites, the maximum parsimony method was used; otherwise BIONJ method with MCL distance matrix was used. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 76 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 2323 positions in the final dataset. Evolutionary analyses were conducted in MEGA5[51].

Supplementary Figure S3 | Transportability of GAS SF370 SRM assays within the Firmicutes phylum.



To determine which of the generated SRM assays targets homologous proteins in other GAS strains or other members of the Firmicutes phylum, all SRM assays (Supplementary Dataset 2) were mapped onto 75 selected genomes (see Supplementary Figure S2). (**a**) Black color represents SRM assay conservation in respective taxa. (**b**) The identity of functional proteins categories with transferable SRM assays were addressed by determining which NMPDR subsystems contained at least 10 conserved SRM assays in respective taxa, indicated by black color. Rows and columns of the heatmaps have been re-orded according to hierarchical clustering represented by the dendrograms. Respective taxonomic rank is indicated with colors. The heatmaps were generated with R programming language (http://www.r-project.org/).

Supplementary Table S1 | Genome sequenced GAS strains used in this work.

PATRIC Genome Info Id[28]	Genome Name	Serotype	CDS†	Publication‡
79812	*Streptococcus pyogenes* M1 GAS (SF370)	M1	1919	11296296[52]
7871	*Streptococcus pyogenes* MGAS5005	M1	1931	16088826[31]
120482	*Streptococcus pyogenes* MGAS10270	M2	2024	16636287[30]
110589	*Streptococcus pyogenes* SSI-1	M3	2009	12799345[53]
96585	*Streptococcus pyogenes* MGAS315	M3	2010	12122206[54]
25933	*Streptococcus pyogenes* MGAS10750	M4	2030	16636287[30]
41547	*Streptococcus pyogenes* str. Manfredo	M5	1935	17012393[55]
30273	*Streptococcus pyogenes* MGAS10394	M6	1960	15272401[56]
37061	*Streptococcus pyogenes* MGAS9429	M12	1893	16636287[30]
133020	*Streptococcus pyogenes* MGAS2096	M12	1953	16636287[30]
4396	*Streptococcus pyogenes* MGAS8232	M18	2007	11917108[57]
35950	*Streptococcus pyogenes* MGAS6180	M28	1956	16088825[58]
129711	*Streptococcus pyogenes* NZ131	M49	1876	18820018[32]

† Coding sequences from PATRIC annotation source

‡ PUBMED identifier

Supplementary References

49. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526 (1993).
50. Case, R. J. *et al.* Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol* **73**, 278–288 (2007).
51. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
52. Ferretti, J. J. *et al.* Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc Natl Acad Sci USA* **98**, 4658–4663 (2001).
53. Nakagawa, I. *et al.* Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res* **13**, 1042–1055 (2003).
54. Beres, S. B. *et al.* Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proc Natl Acad Sci USA* **99**, 10078–10083 (2002).
55. Holden, M. T. G. *et al.* Complete genome of acute rheumatic fever-associated serotype M5 *Streptococcus pyogenes* strain manfredo. *J Bacteriol* **189**, 1473–1477 (2007).
56. Banks, D. J. *et al.* Progress toward characterization of the group A Streptococcus metagenome: complete genome sequence of a macrolide-resistant serotype M6 strain. *J Infect Dis* **190**, 727–738 (2004).
57. Smoot, J. C. *et al.* Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks. *Proc Natl Acad Sci USA* **99**, 4668–4673 (2002).
58. Green, N. M. *et al.* Genome sequence of a serotype M28 strain of group a streptococcus: potential new insights into puerperal sepsis and bacterial disease specificity. *J Infect Dis* **192**, 760–770 (2005).