



Genetic diversity and distribution of pneumococcal surface lipoproteins and implications on potential protein-based vaccines

Ebrima Bojang¹, Senghore M¹, Chaguza C², Harris S², Weinert L³, Worwui A¹, Antonio M¹, Breiman R⁴, McGee L⁵, Bentley SD², Kwambana-Adams B¹

1. MRC Unit The Gambia at The London School of Hygiene and Tropical Medicine, Atlantic Boulevard, Fajara, P.O Box 273, 2. The Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK, 3. University of Cambridge, department of Veterinary Medicine, Cambridge CB3 0ES, UK, 4. Hubert Department of Global Health, Emory Global Health Institute, Atlanta, GA, USA, 5. Centers for Disease Control and Prevention, Respiratory Diseases Branch, Atlanta, GA, USA

Introduction

- The licensed pneumococcal conjugate vaccines are of limited serotype valency
- Protein candidates that reached clinical trials have not been successful so far (i.e. Pneumolysin and pneumococcal histidine triad protein D)
- Therefore, conserved pneumococcal proteins such as lipoproteins (Figure 1) represent alternative vaccine candidates to combat pneumococcal diseases.

This study was designed to identify and evaluate all conserved surface pneumococcal lipoproteins with positive signatures for lipoproteins for their potential as protein vaccine candidates.

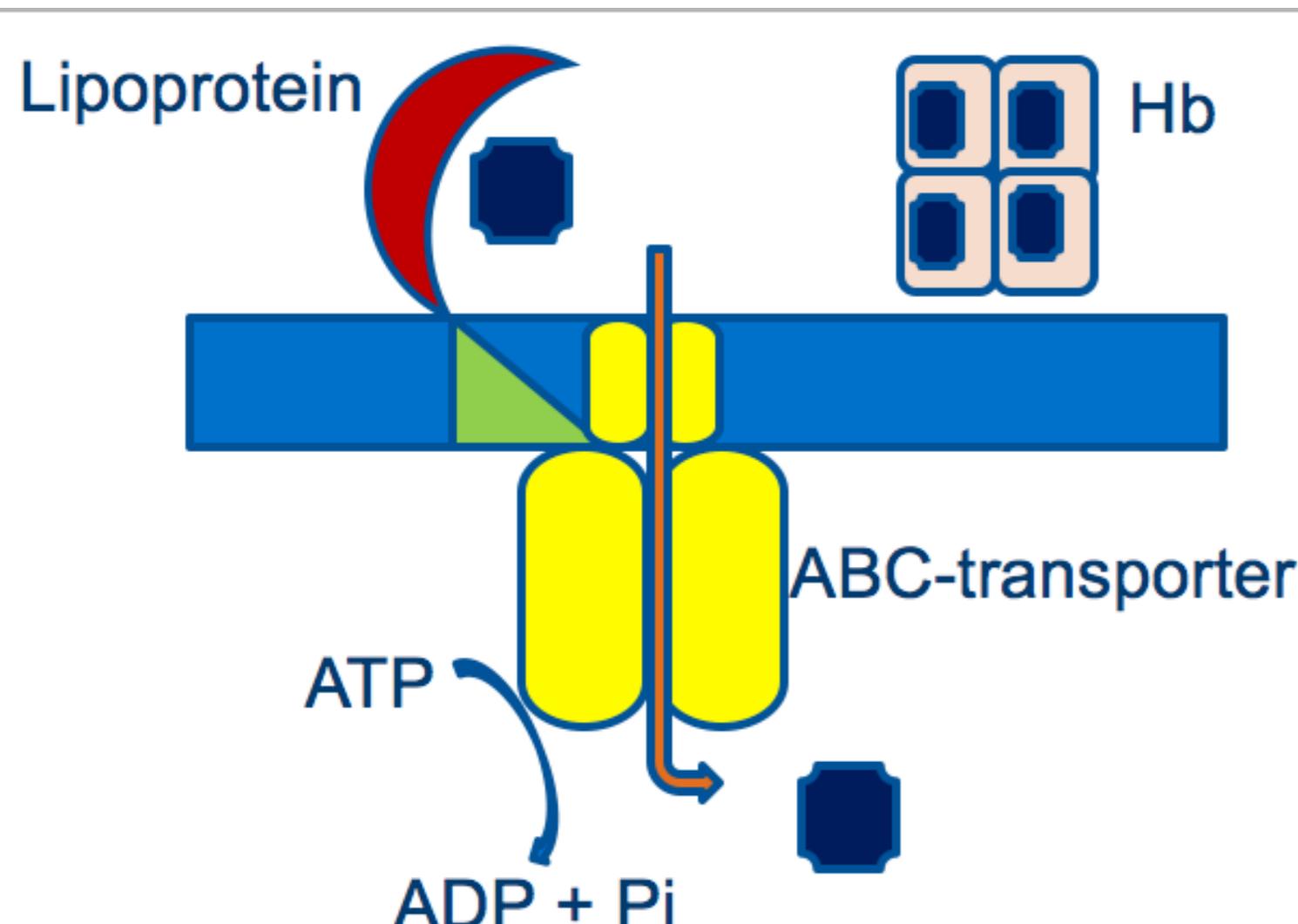


Figure 1. An illustration of the components of a typical gram-positive substrate binding lipoprotein.

Methods

- Study included 1769 annotated pneumococcal genomes from The Gambia
- Each gene was screened using a custom pipeline (Figure 2)
- Genes were scored based on size, prevalence, allele count, number of chains and proportion predicted as epitope

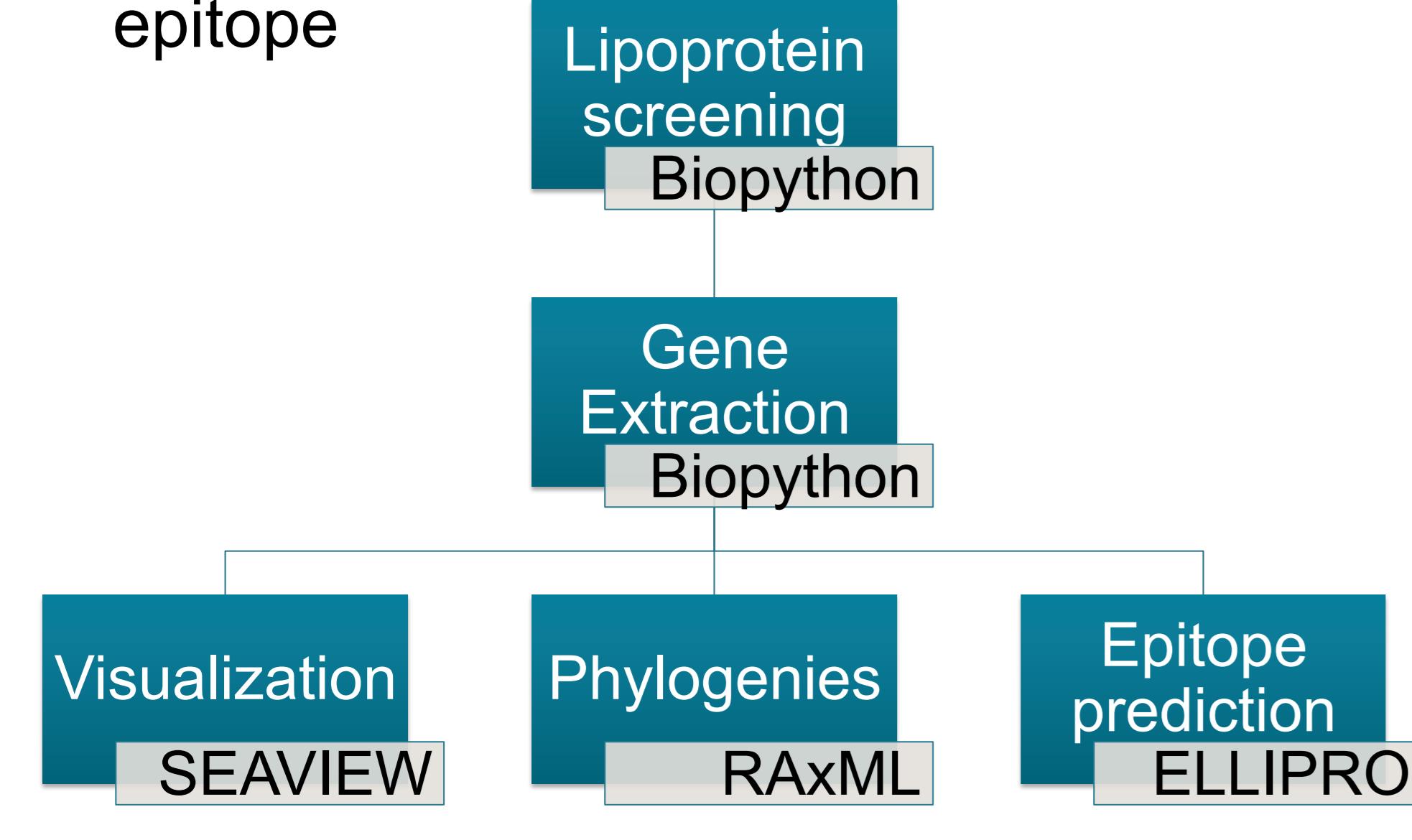


Figure 2 A flowchart showing the general steps employed in our gene screening.

Results

- 169 lipoprotein hits were initially obtained using established lipoprotein pattern searches
- 30 candidate proteins were more than 90% prevalent and were predicted to be surface expressed

Table 1. Top 5 ranked predicted proteins.

Lipoprotein	AA length	Allele count	% proportion as epitope	Chains	Prevalence	Total points	Rank
Group_2005	503	26	53%	2	100%	98	1
TauA	335	15	55.1%	1	100%	96.1	2
Group_2074	188	11	58%	1	100%	96	3
AmiA	660	37	54%	1	99.8%	95.8	4
PsaA	309	17	51.6%	2	100%	94.6	5

Note: Predicted proteins with name starting "Group_" are hypothetical proteins without a generic gene name. The table only shows the different categories used for scoring and their total score.

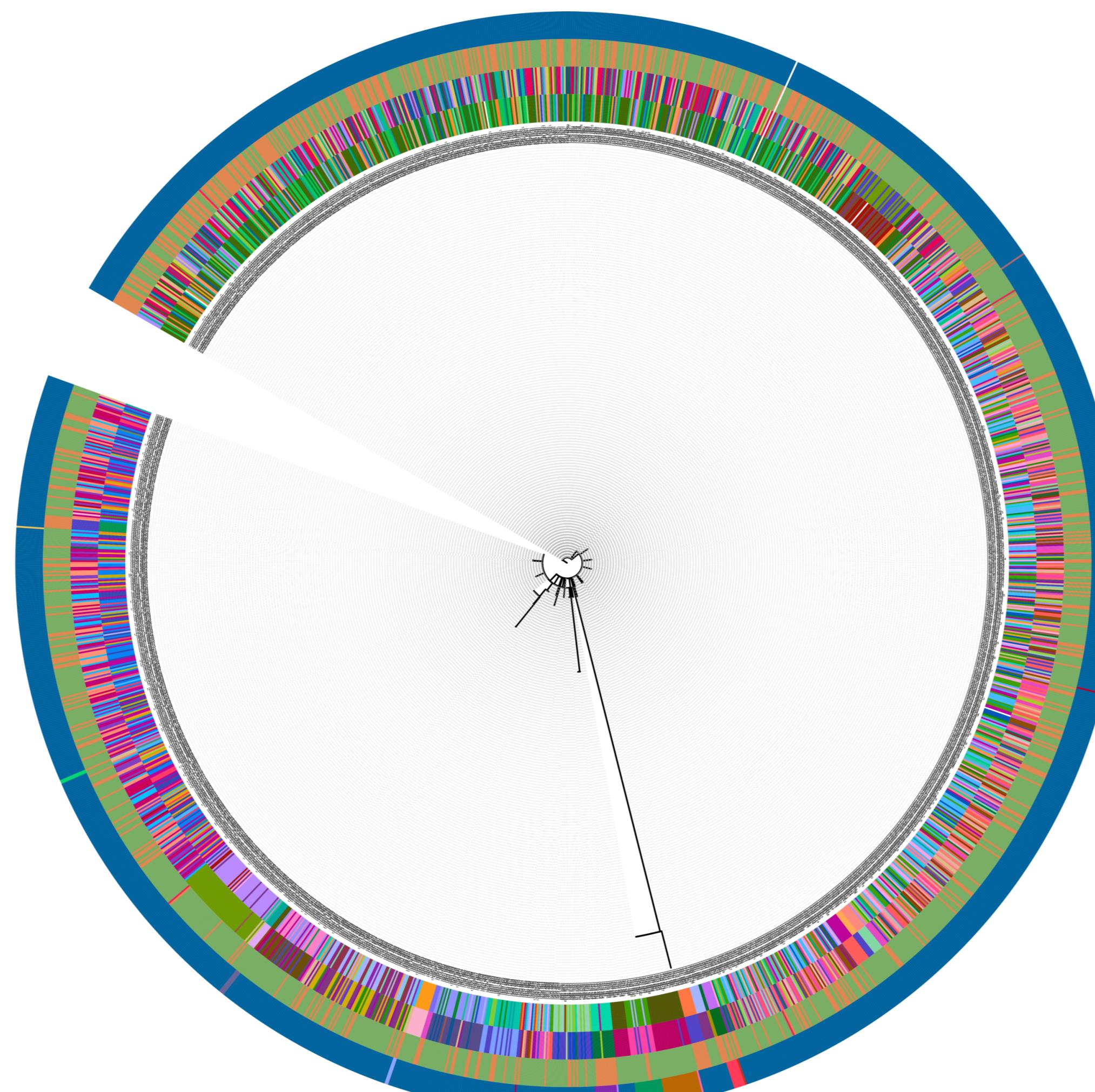


Figure 3. Phylogenetic gene tree of TauA. The lanes from inside out represent the lane IDs, serotypes, BAPS cluster (lineages), disease status (carriage or disease) and alleles. The dominant allele covers both disease and carriage strains and is not serotype or lineage specific

Key findings for TauA

- Predicted as a periplasmic binding protein-like II family
- 100% prevalent in the genomes
- >50% of protein predicted as immunogenic
- <6 AA divergence within the proteins
- Long branch represents NTs driven by 5 AA divergence

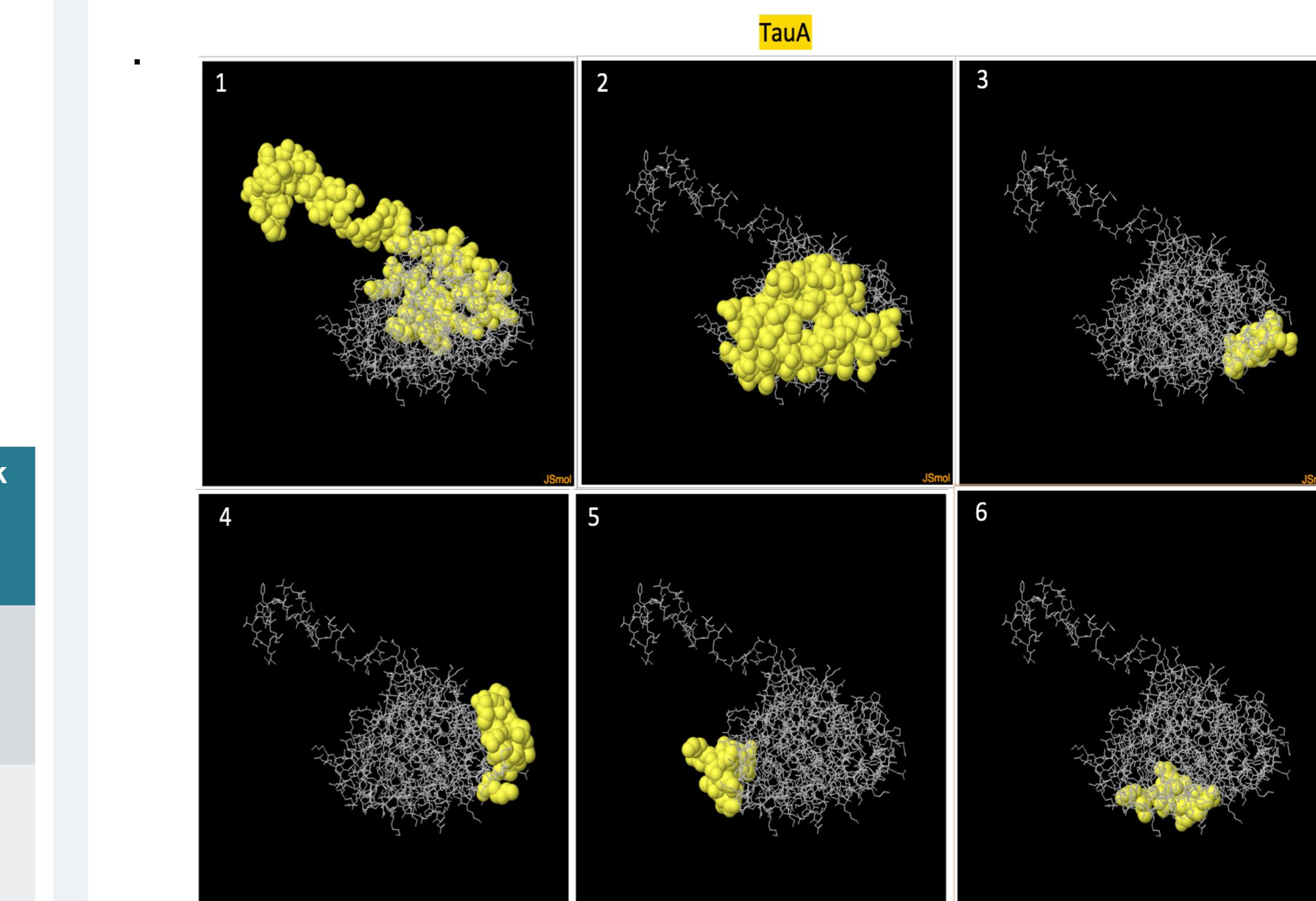


Figure 4. ElliPro predicted epitopes of TauA. Yellow represents residues predicted as epitopes. Diagram 1 is the prediction with the highest confidence and 6 has the least confidence with ElliPro.

Key findings

- 15 lipoproteins were at least 97% conserved, ≥ 99% prevalent and immunogenic
- TauA is an example of a good vaccine candidate identified (Figure 3)
- The predominant allele covered 96% of isolates irrespective of serotypes, disease status or lineage
- Most proteins have good immunogenicity predictions with >50% of the protein predicted as epitope (Table 1 and Figure 4)

The way forward

- Experimental (animal) models needed to validate the novel candidates
- Multi-protein vaccine may be the way forward since some candidates have redundant functions.
- In-silico screening of candidate proteins is a sensitive and quicker way of identifying potential vaccine candidates

Acknowledgements

I would like to acknowledge the Wellcome Sanger Institute for funding my MPhil, the Global Pneumococcal Sequencing project, Dr Assan Jaye for his mentorship and the study participants.