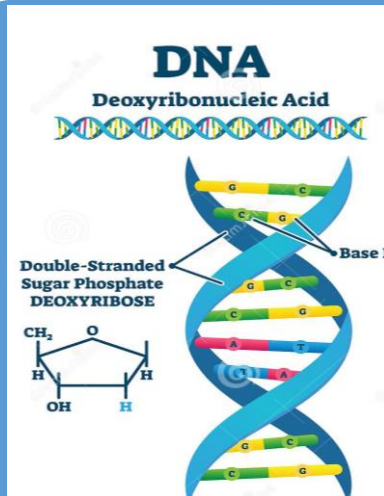


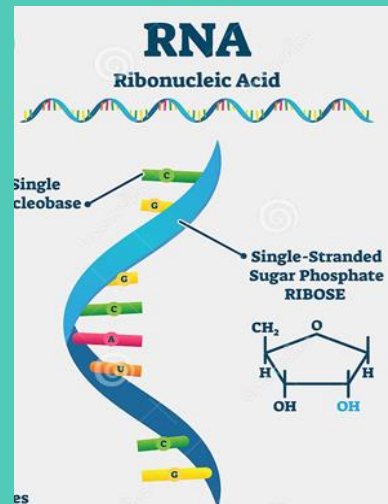
Proteomics and its applications

James Nyagwange, BHKi seminar 20/07/2022

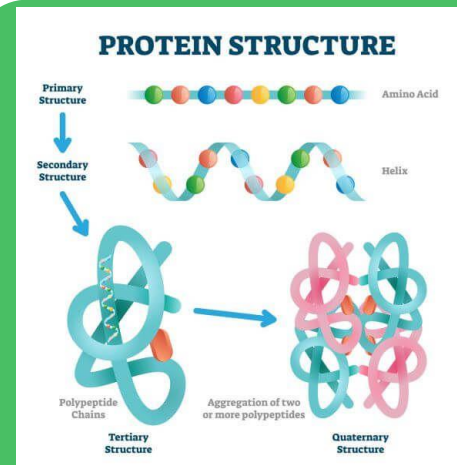
Possibly, probably, actually and more



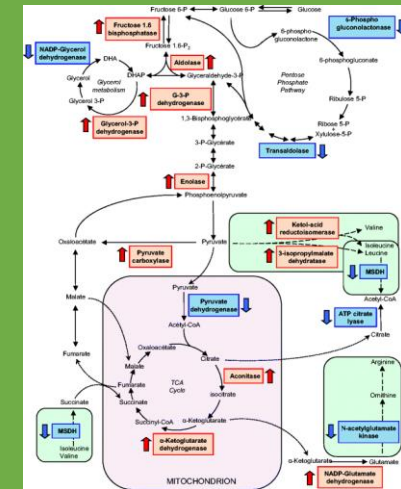
Genomics



Transcriptomics



Proteomics

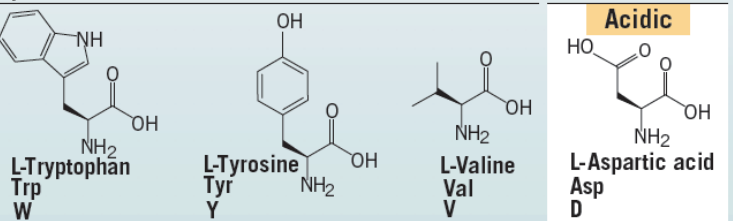
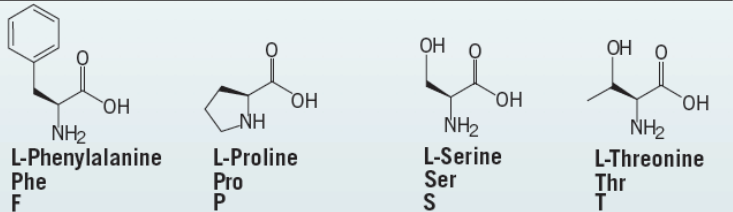
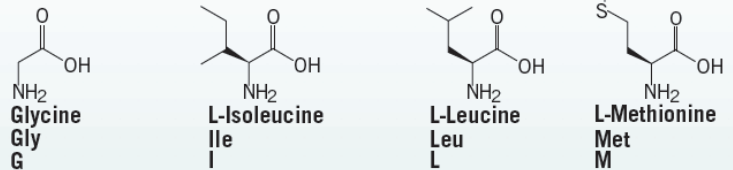
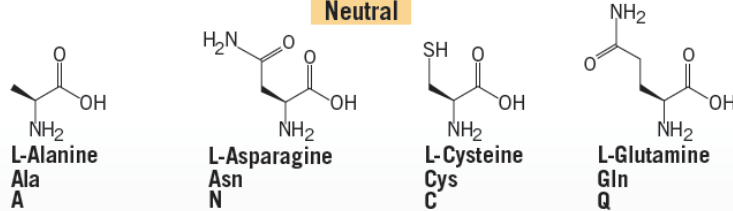


Metabolomics

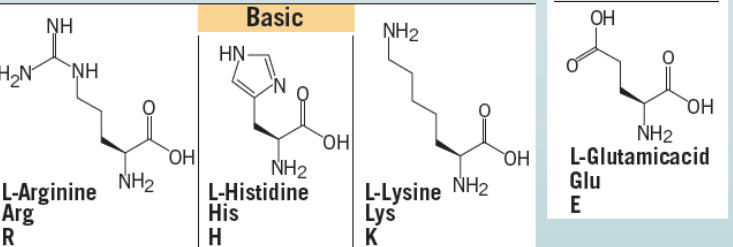
Proteomics study large sets of proteins

AMINO ACID STRUCTURES AND ABBREVIATIONS

Neutral

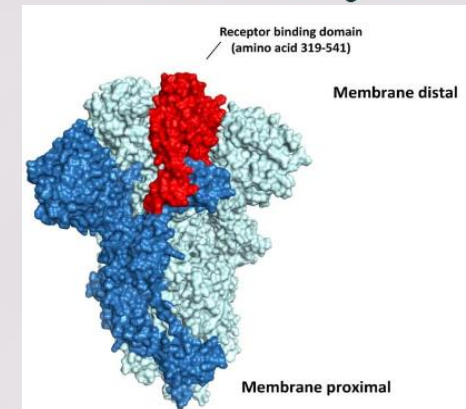
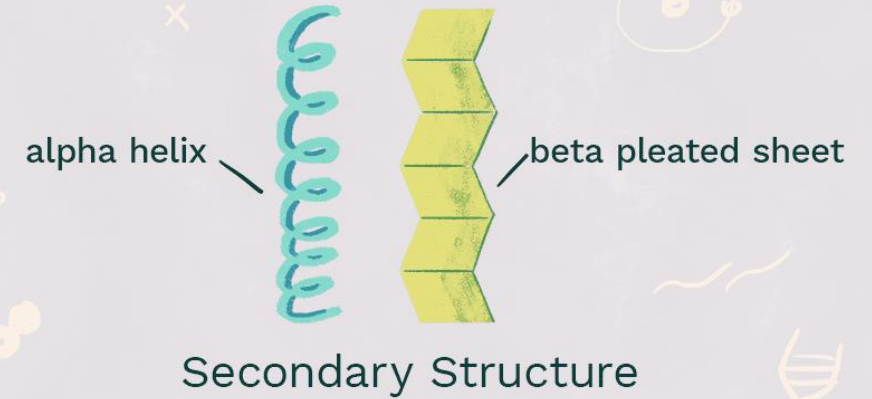
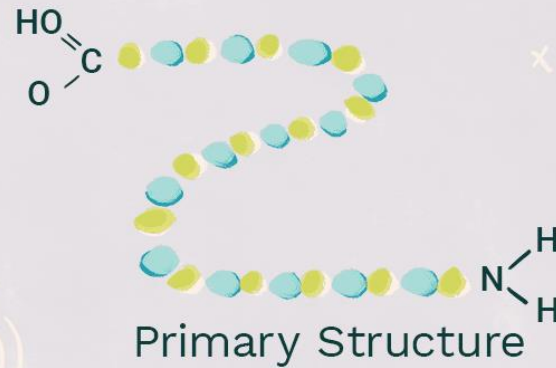


Basic



Lubrizol Life Science

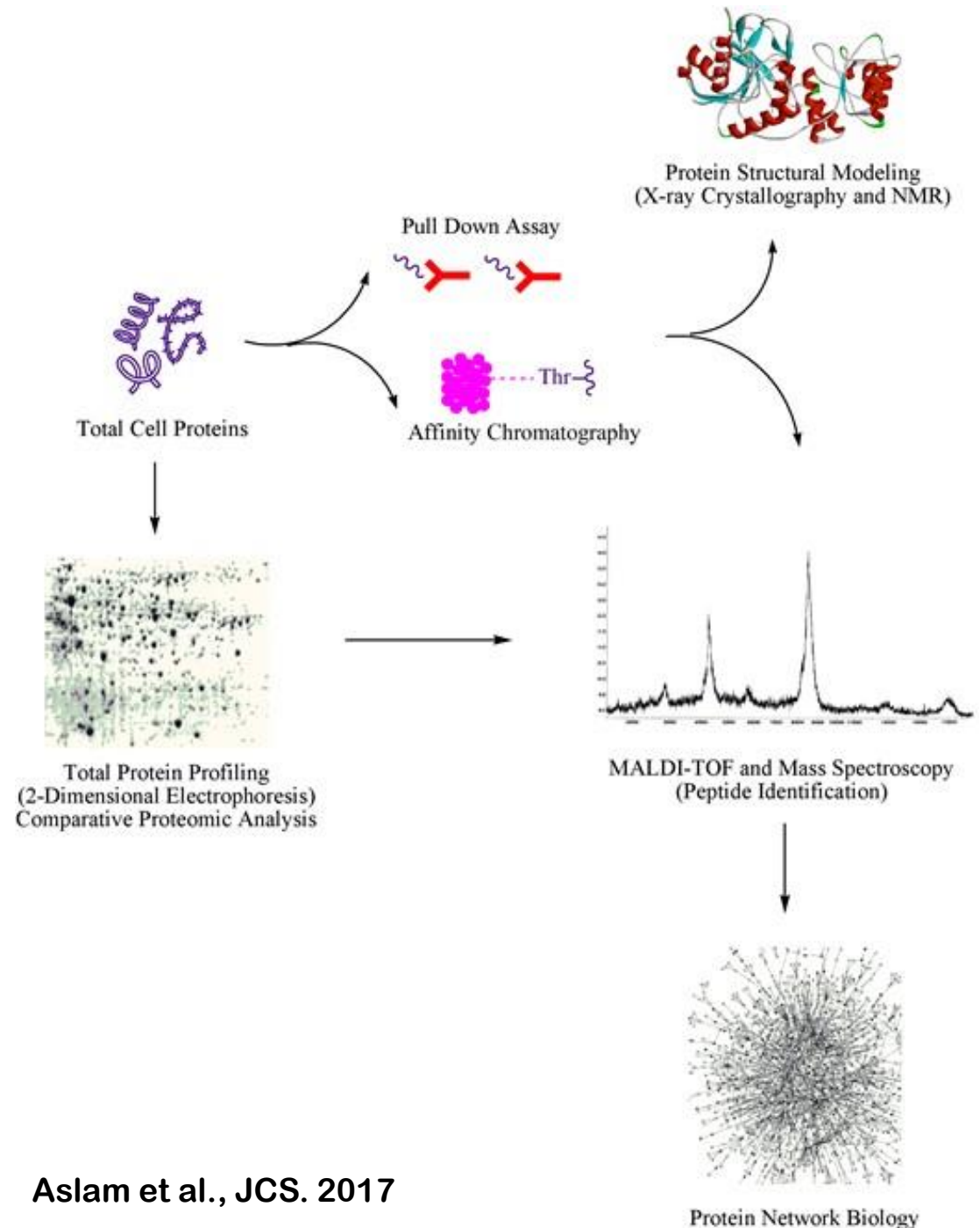
Types of Protein Structures



ThoughtCo.

Protein analysis workflow

- Purification –chromatography based
- Analysis –ELISA, WB, protein micro-array
- Characterization-gel based, mass spec
- Seq analysis-Edman
- Quantification-ICAT, SILAC, iTRAQ
- Structural analysis-X-ray crystallography, NMR spectro
- Bioinformatic analysis



Resources

Databases

- GenBank
- RefSeq
- nr
- UniProt
- UniRef
- UniParc
- TrEMBL
- SwissProt

Protein structure databases

- PDB
- SwissModel
- SCOP
- ModBase
- CATH
- MMDD

Alignments (BLAST)

- BLASTP BLAST
- TBLASTN
- PSI-BLAST Position Specific Iterated BLAST
- PHI-BLAST Pattern Hit Initiated BLAST
- DELTA-BLAST Domain Enhanced Lookup Time Accelerated BLAST

Protein structure prediction sever

- PredictProtein
- O-GlycoBase
- PhosphoBase
- SwissModel
- WhatIf
- ESyPred3D
- EBI

Protein structure analysis server

- ConSurf
- CASTp
- ProtSkin
- LigandProtein

Protein motif search tools

- PATTINPROT
- SIRW

Motif based alignment server

- Match Box
- MEME
- Dialign
- BlockMakei

Protein domain servers

- InterProScan
- CD server
- ProWleScan
- ScanProsite

Proteomic applications

- Delineation of altered protein expression in disease
- Development of novel biomarkers for diagnosis and early detection of disease
- Potential for accelerating drug development through more effective strategies to evaluate therapeutic effect and toxicity
- Identification of new targets for therapeutics and vaccines



ELSEVIER

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara



Characterization of the *Theileria parva* sporozoite proteome

James Nyagwange^{a,b}, Edwin Tijhaar^b, Nicola Ternette^c, Fredrick Mobegi^f, Kyle Tretina^d,
Joana C. Silva^{d,e}, Roger Pelle^a, Vishvanath Nene^{a,*}

^a International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya

^b Cell Biology and Immunology Group, Wageningen University, The Netherlands

^c The Jenner Institute, Nuffield Department of Medicine, University of Oxford, UK

^d Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA

^e Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, USA

^f Department of Infection and Immunity, South Australian Health and Medical Research Institute, North Terrace, Adelaide 5000, South Australia, Australia



The problem

- East Coast fever (ECF) cattle diseases causing losses >\$300M/year in Africa
- *Theileria parva* (*Tp*) causes ECF
- The sporozoite stage of *Tp*, harbored and released from the salivary glands of the tick *Rhipicephalus appendiculatus* during feeding
- Sporozoites invade and establish infection in bovine lymphocytes



Possible solution

- Block sporozoites from invading bovine lymphocytes by vaccine
- Vaccine requires identification of surface proteins
- Proteins used to induce cattle immune response
- Proteomics can help

Workflow

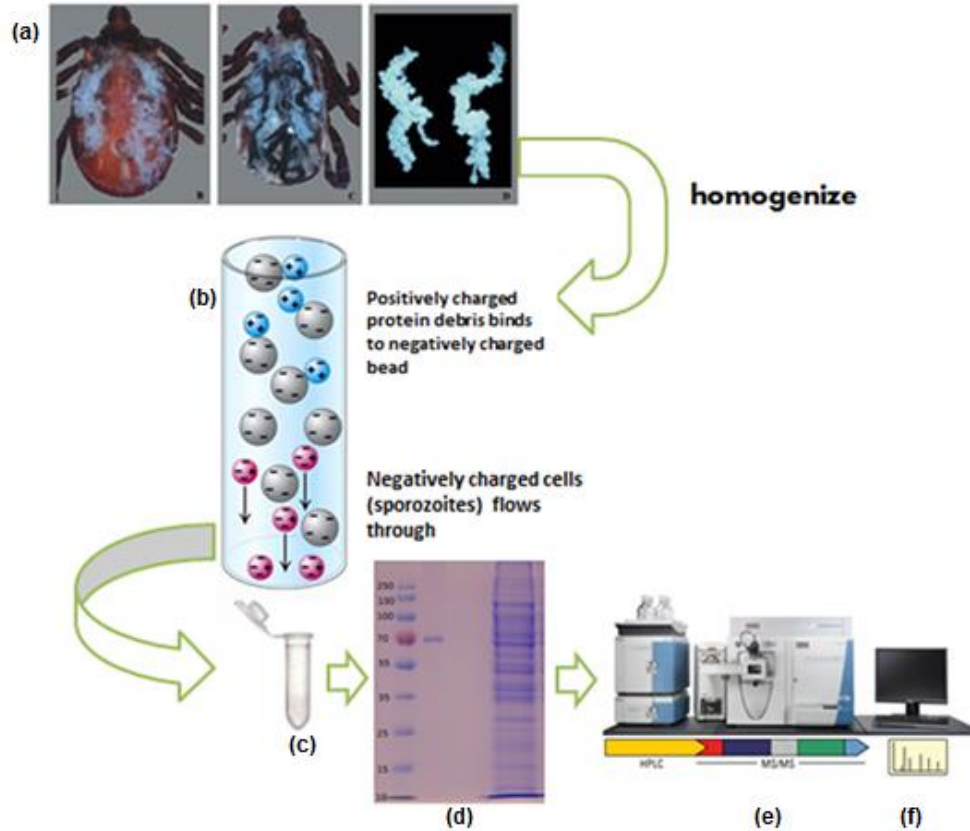


Figure 1. Overview of experimental procedures. Infected ticks were dissected and salivary glands removed (a), homogenized, and passed through a DE-52 column (b). The resulting sporozoite-enriched suspension was lysed (c), and size fractionated by SDS-PAGE (d) followed by LC-MS/MS (e), and subsequent bioinformatical analysis (f).

Trypsinization

- Gel divided in 4 fractions, de-stained, dehydrated, reduced, alkylated, washed
- Proteolytic digestion by trypsin

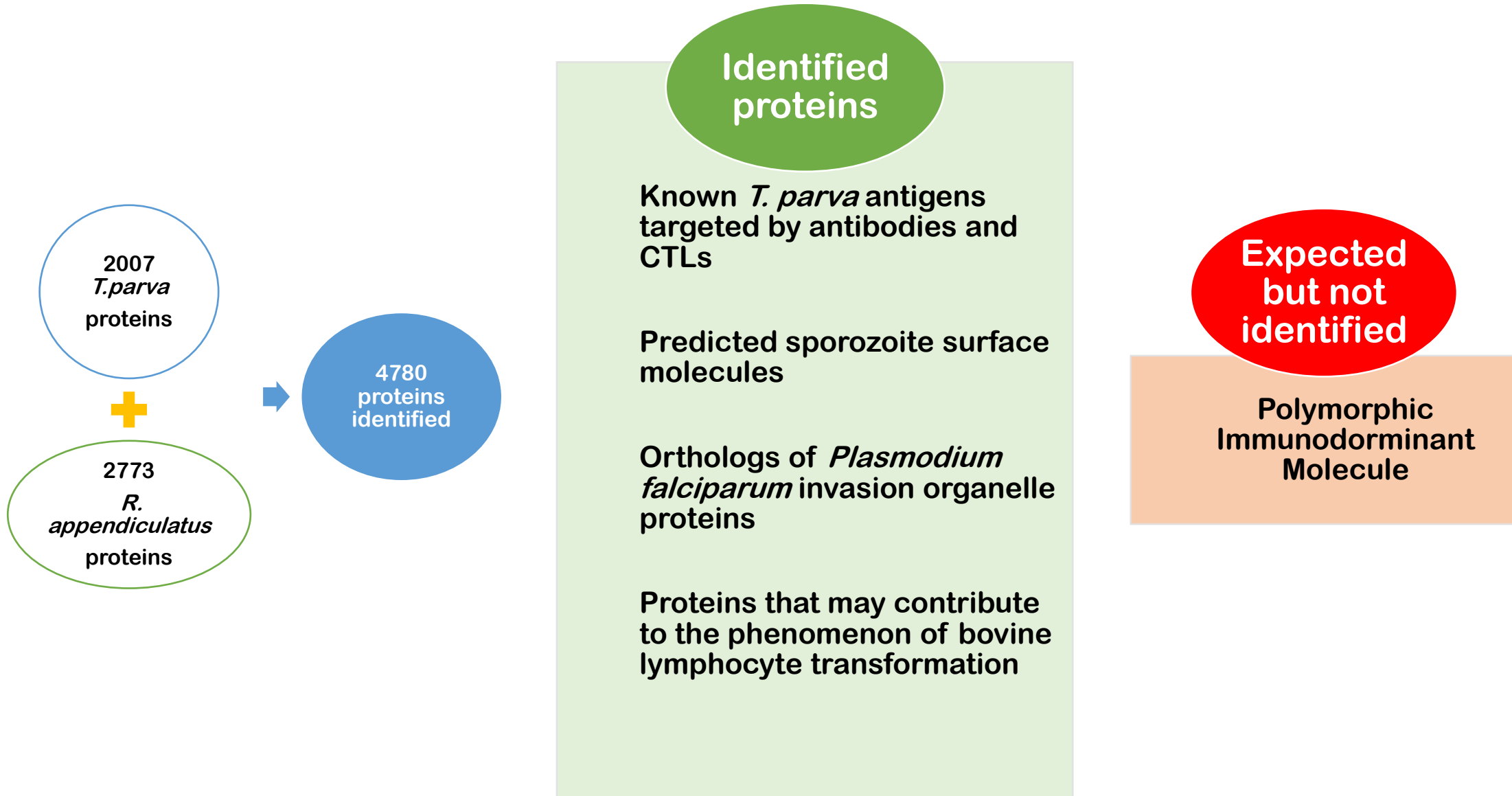
LC-MS/MS

- Gel fractions separated by column (PepMap C18)
- Then introduced into mass spec (QExactive)

Data analysis

- Raw MS/MS files analysed with Peaks software using db with Tp (parasite) and R.app (tick)
- FDR of 1.2% applied and only proteins with unique peptide id included in proteome list
- Exponentially Modified Protein Abundance Index (emPAI) by Mascot
- Classification by subcellular localization (TargetP, TMHMM, SignalP, GPI-SOM)
- Classification by orthology with Pf3D7 (OrthoMCL)

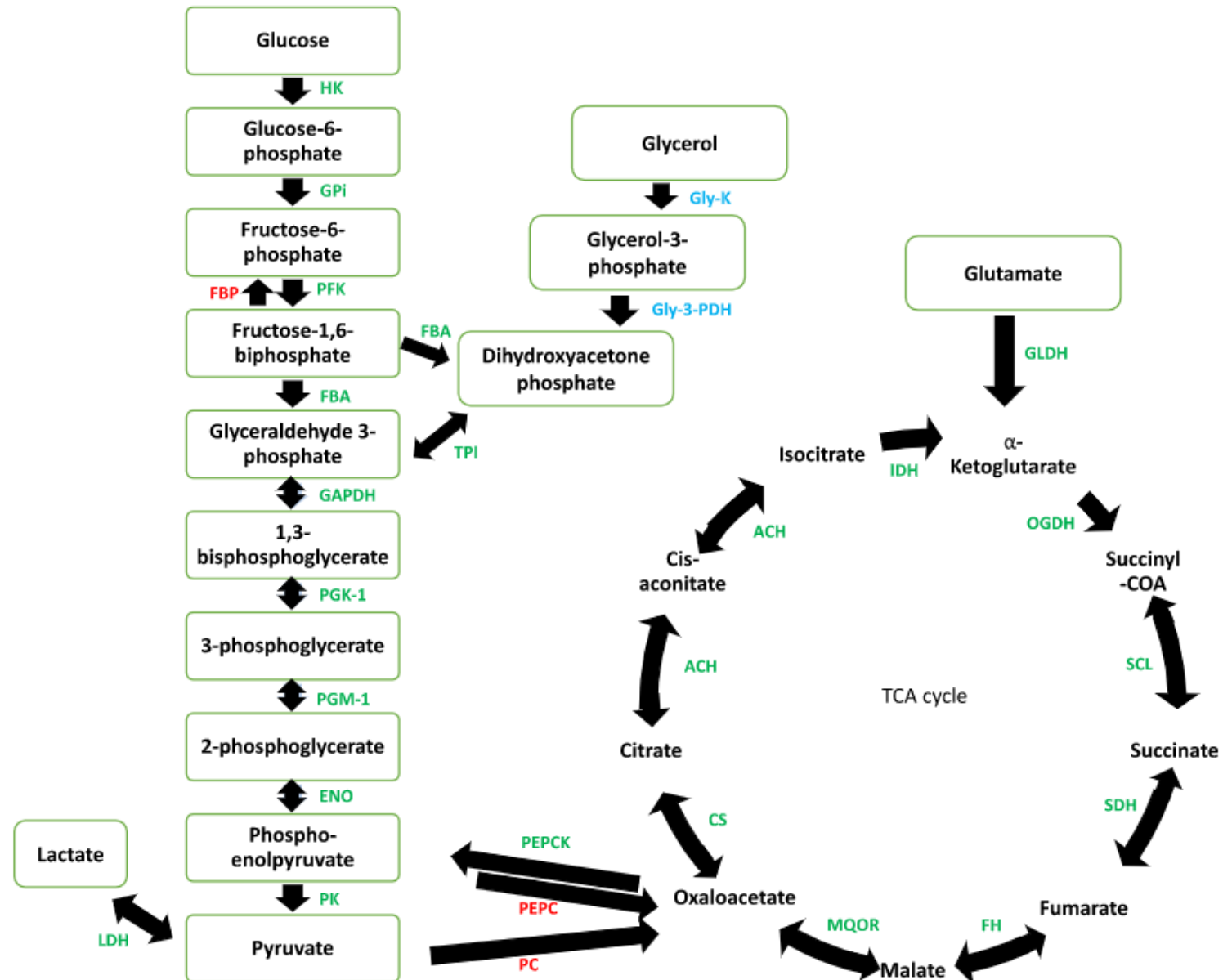
Results



Abundant *T. parva* sporozoite proteins

ORF locus tag	Description ^a	emPAI
TpMuguga_04g00404	Histone H2B.1	7228.34
TpMuguga_04g00071	Histone H2B	900.39
TpMuguga_04g00675	Histone H4	457.35
TpMuguga_04g00690	phosphoglycerate mutase 1 family	105.17
TpMuguga_04g00050	Ribosomal protein S19	99.38
TpMuguga_03g00655	hypothetical protein	96.5
TpMuguga_04g00036	AhpC/TSA family	84.68
TpMuguga_02g00487	Ribosomal protein S6e	67.36
TpMuguga_02g00903	Actin	60.62
TpMuguga_04g00322	Histone H2A	60.26
TpMuguga_03g00067	hypothetical protein	54.06
TpMuguga_01g00726	eIF-Tu GTP binding domain	52.64
TpMuguga_01g00541	hypothetical protein	52.36
TpMuguga_04g00383	GAPDH NAD binding domain	51.37
TpMuguga_01g00701	RAP-1	47.73
TpMuguga_02g00148	Heat shock 70 kDa protein	42.22
TpMuguga_04g00179	RanBP1 domain	41.34
TpMuguga_03g00700	hypothetical protein	40
TpMuguga_03g00287	Sporozoite P67 surface antigen	39.16
TpMuguga_01g00067	Acyl CoA binding protein	39.11

Proteins involved in metabolism identified



Identification of sporozoite antigens that mediate the infection process

Approach

In silico; GPI* and SP* predictions, conserved fragments of proteins

Protein subunit with adjuvant

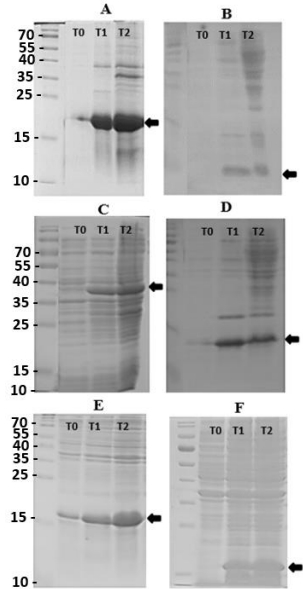
Murine studies
- Good antibody titres
In vitro assays
-ELISA,
seroneutralization

#	Gene locus	Protein product	Signal P	Full length protein (kDa)	Expressed fragment	Identified by LC-MS
1	TpMuguga_01g00095	hypothetical protein	YES	29	29%	NO
2	TpMuguga_01g00876	hypothetical protein	YES	13	50%	YES
3	TpMuguga_04g00437	104 kDa antigen	YES	104	12%	YES
4	TpMuguga_01g00575	hypothetical protein	YES	198	10%	NO
5	TpMuguga_01g00972	hypothetical protein	YES	38	32%	YES
6	TpMuguga_01g00939	hypothetical protein (p34)	YES	34	40%	YES
7	TpMuguga_03g00136	hypothetical protein	YES	21	NO	NO
8	TpMuguga_04g02375	hypothetical protein	NO	40	NO	NO
9	TpMuguga_03g00844	hypothetical protein	YES	14	NO	YES
10	TpMuguga_02g00792	hypothetical protein	YES	15	NO	YES

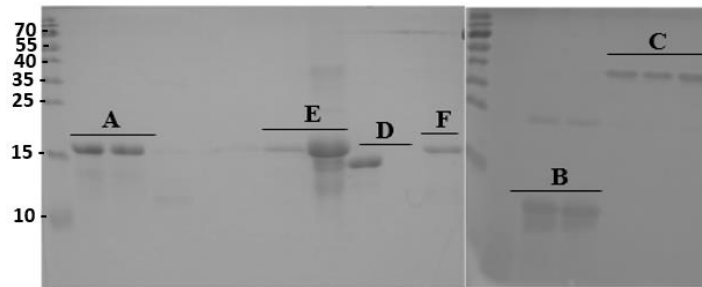
* Glycosylphosphatidylinositol (GPI), signal peptide (SP)

Expression of antigens and immunogenicity studies

Recombinant protein expressed in *E. coli*

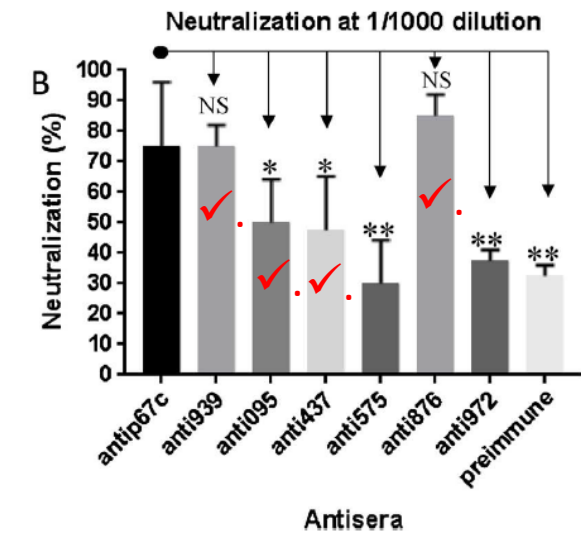
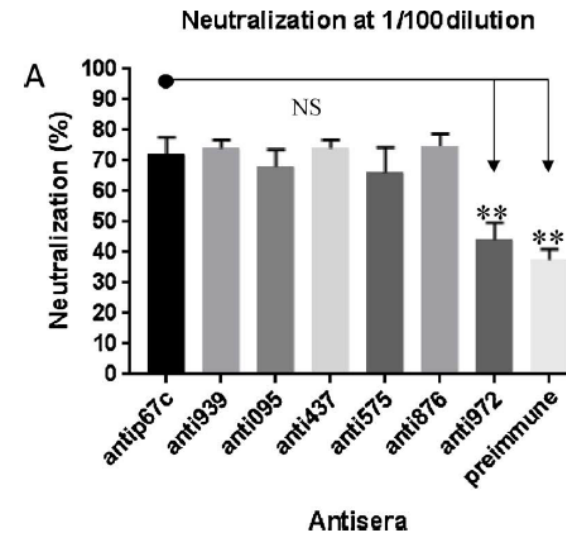


Purified recombinant proteins



Antigen + adjuvant used to immunize mice and sera obtained

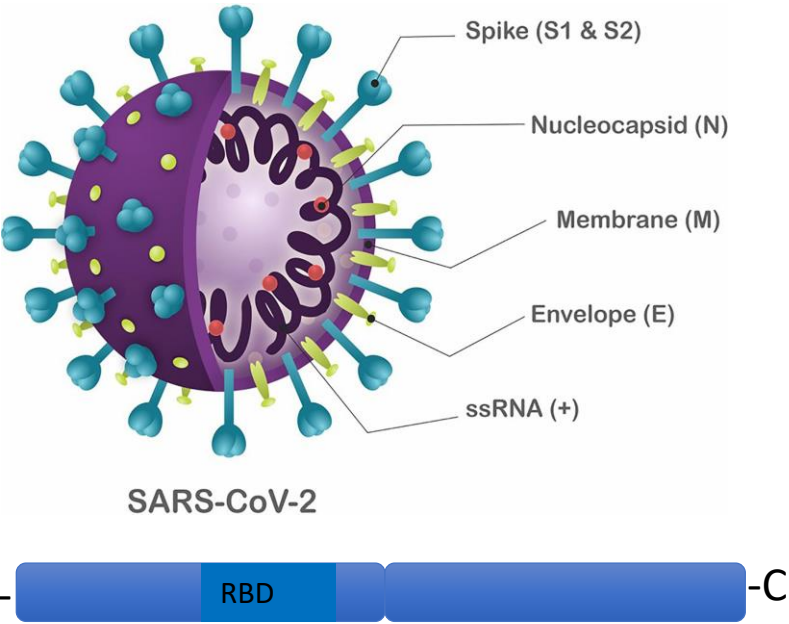
Sporozoite neutralization assay



✓. 4 promising vaccine candidate antigens identified

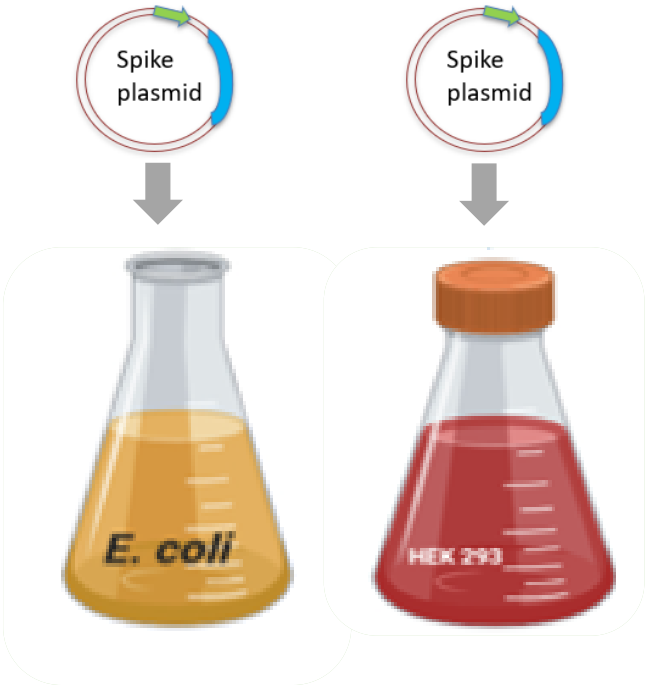
Recent times; pandemic response

1. Spike protein selection

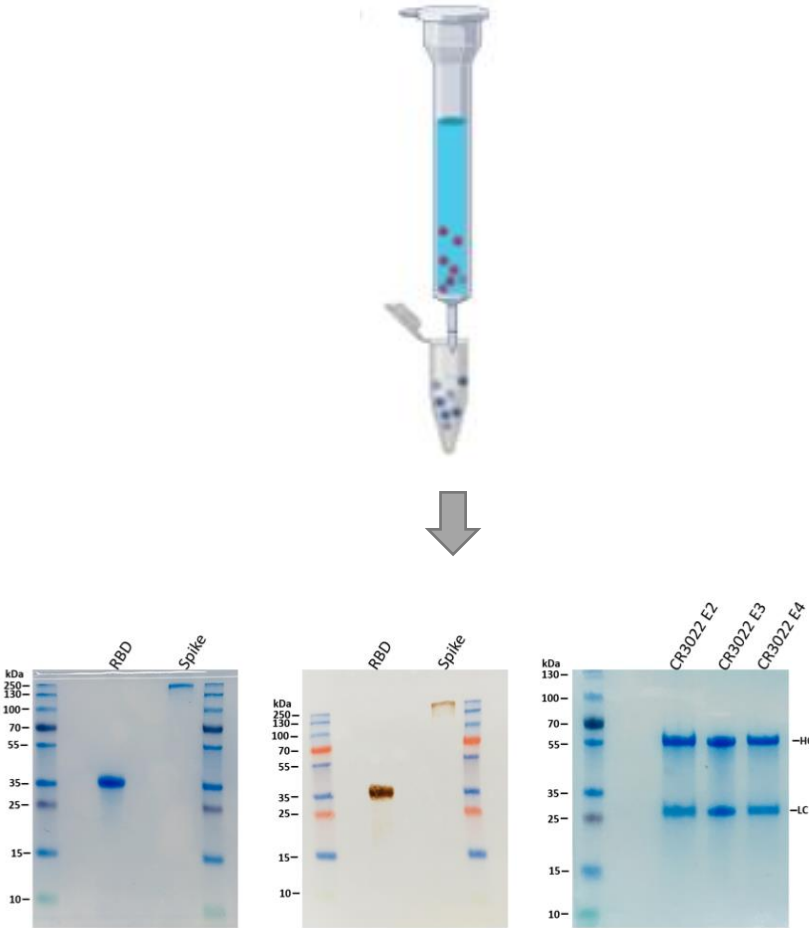


Santos et al., 2020

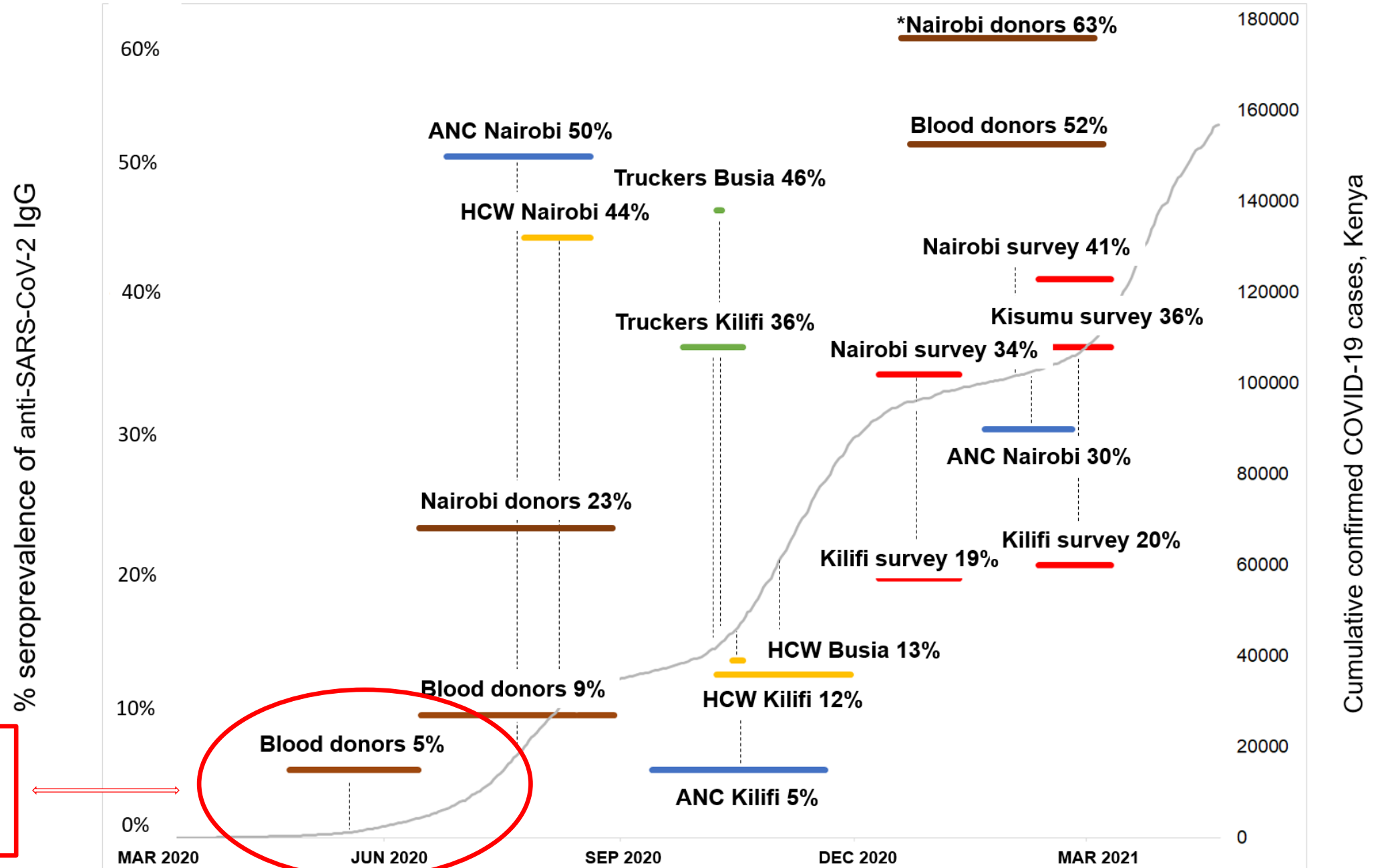
2. Cloning and expression of Spike protein



3. Purification of proteins



Seroprevalence IgG antibodies in Kenya



Summary

- **Proteomics is useful in identifying targets for vaccines, therapeutics, and diagnosis**
- **Proteomics useful for elucidating protein structures and quantifying expression levels**
- **Preceding wet lab is as important as bioinformatic analysis**
- **You need genomics/transcriptomic databases to make sense of your data**

- **Thanks for listening**