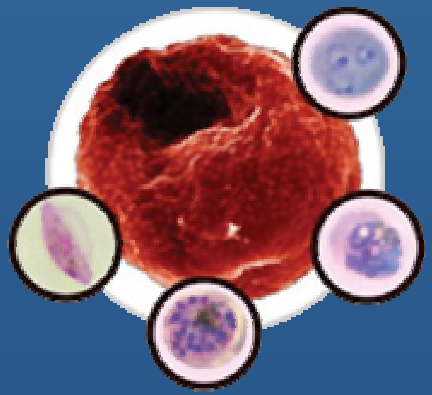
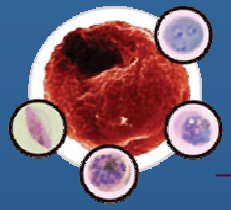


A proteomic view of the *Plasmodium falciparum* life cycle

LAURENCE FLORENS, MICHAEL P. WASHBURN, J. DALE RAINE, ROBERT M. ANTHONY, MUNIRA GRAINGER, J. DAVID HAYNES, J. KATHLEEN MOCH, NEMONE MUSTER, JOHN B. SACCI, DAVID L. TABB, ADAM A. WITNEY, DIRK WOLTERS, YIMIN WU, MALCOLM J. GARDNER, ANTHONY A. HOLDER , ROBERT E. SINDEN, JOHN R. YATES & DANIEL J. CARUCCI

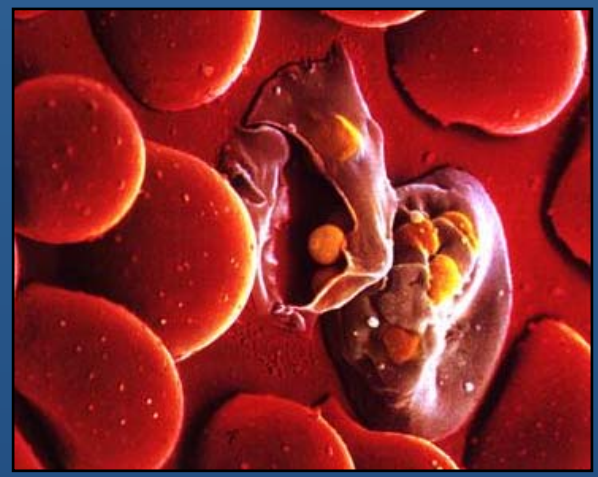
Nature 419, 520-526 (2002)

Image from Nature's website, http://www.nature.com/nature/malaria/



What is "*P. falicparum*", and why do I care?

- *Plasmodium falciparum* is a protozoan parasite which can infect humans, causing the disease **malaria**. It is spread primarily by the bite of infected *Anopheles* mosquitos (especially *A. gambiae*).



http://www.medinfo.ufl.edu/year2/mmid/bms5300/bugs/plasfalc.html

- Malaria causes over 1 million deaths per year. The disease is most prevalent in Africa, where it kills one child every 30 seconds. It is estimated that approximately 40% of the world's population is at risk.

- Malaria can be treated with drugs if caught early, but it is quickly developing drug resistance, and most areas at risk are poor and without access to new (or any) treatments.

Information on the disease & current statistics available at: http://www.who.int/inf-fs/en/fact094.html

The Life Cycle of Malaria

- The life cycle of the malarial parasite requires specialized protein expression for life in mammals and insects, intracellular and extracellular environments, and evasion of host immune response.

- Differential protein expression results in difficulties for creation of vaccines and anti-malarial drugs: the metabolic pathways (anti-malarials) and surface coatings (vaccines) change drastically during the organism's life cycle!

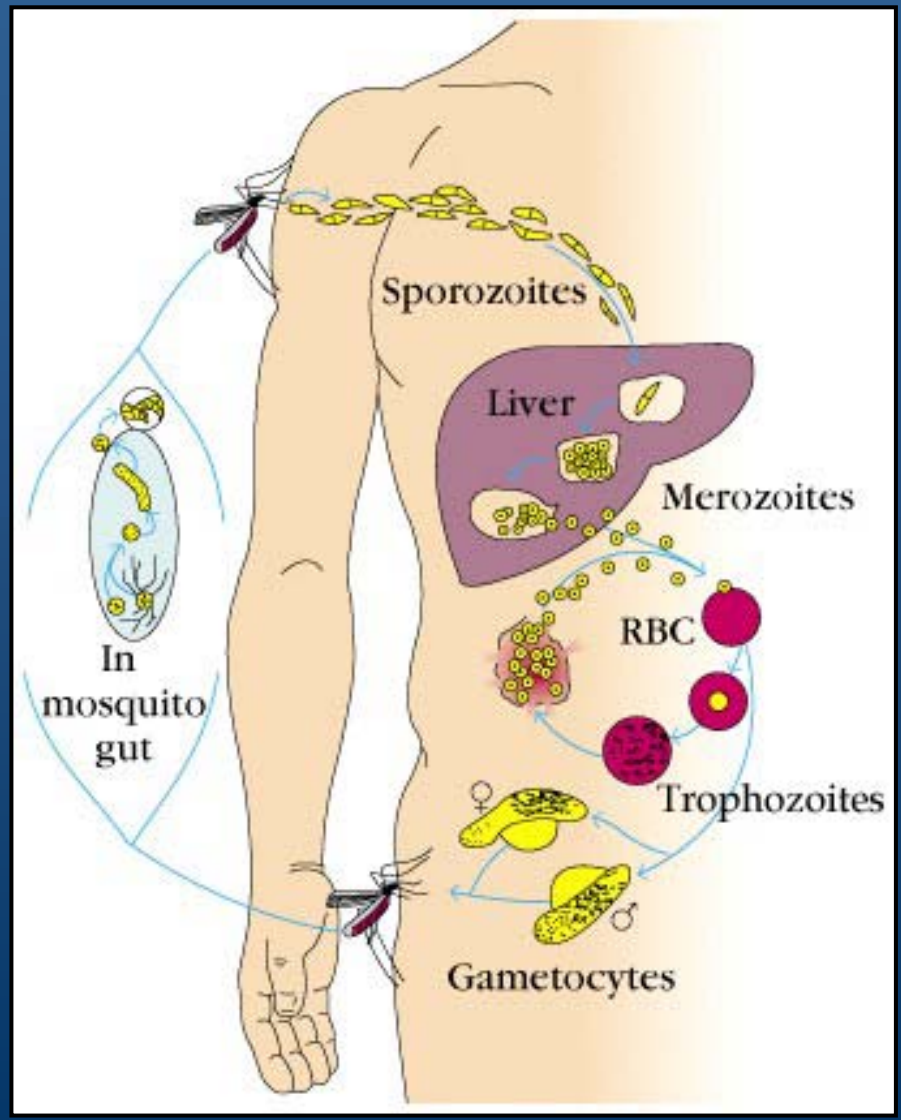
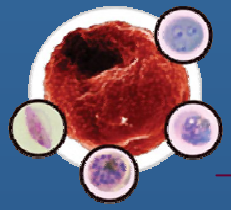


Image from Kuby Immunology, 4th edition, figure 17.11

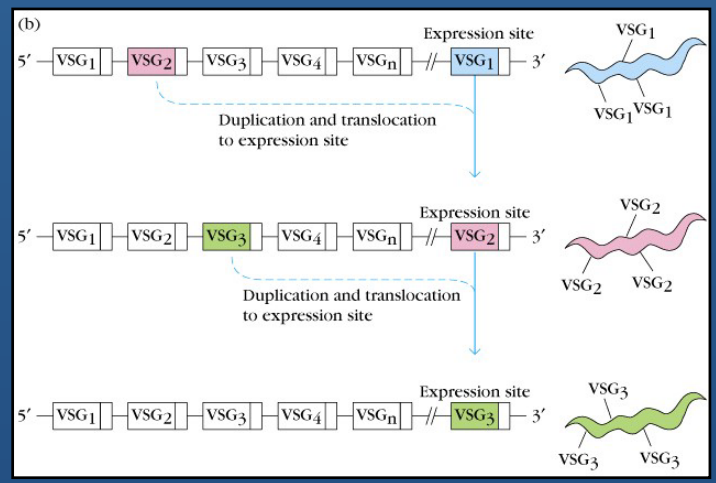


Protozoan parasites & host immune response

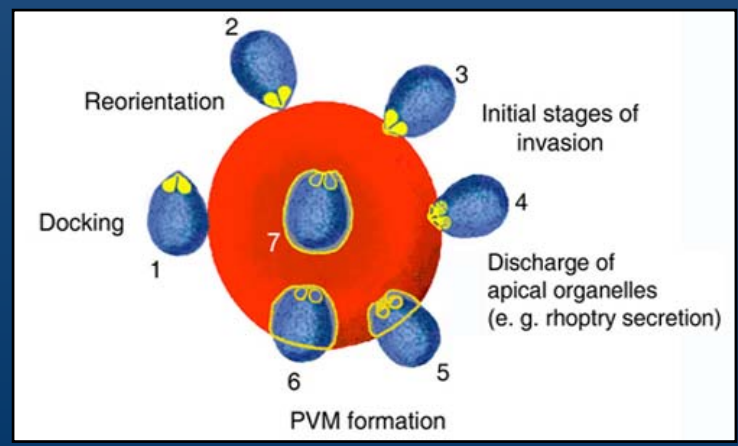
- **Antigenic shift** is the ability of a parasite to change the amino acid sequence of certain surface proteins which would otherwise be recognized by the host's immune response.

- **Rhoptry organelles** contain proteins ("rhoptry proteins") which are released into a red blood cell during invasion; they are thought to be involved in host cell conversion.

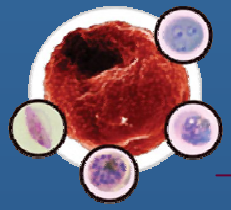
- The **parasitophorous vacuole** is a membrane-bound compartment formed around the parasite when it enters the cell.



Antigenic shift in *Trypanosoma sp.*
From Kuby Immunology, 4th ed., fig17-12



Rhoptry organelles and the parasitophorous vacuole
http://www.uni-wuerzburg.de/infektionsbiologie/ML-projects.htm



Abstract

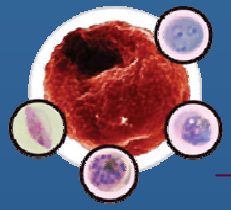
- A high-throughput proteomics approach was used to identify new potential drug and vaccine targets and to better understand the biology of the complex protozoan parasite *P. falciparum*.

- Four stages of the parasite life cycle (sporozoites, merozoites, trophozoites and gametocytes) were characterized by multidimensional protein identification technology (MudPIT).

- Functional profiling of over 2,400 proteins agreed with the physiology of each stage.

- Unexpectedly, the antigenically variant proteins of var and rif genes, defined as molecules on the surface of infected erythrocytes, were also largely expressed in sporozoites.

- The detection of chromosomal clusters encoding co-expressed proteins suggested a potential mechanism forcontrolling gene expression.



Methods

- The *Plasmodium falciparum* clone 3D7 was used throughout.
- Sporozoite proteins were isolated from the salivary glands of *Anopheles stephansii* mosquitos, 14 days after infection.
- Trophozoite proteins were isolated from trophozoites removed from synchronized infected erythrocyte (red blood cell) cultures.
- Merozoite proteins were isolated from merozoites removed from highly synchronized schizonts (a parasite stage which is basically a cluster of merozoites).
- Gametocytes were cultured; there was minor contamination (< 3%) by mixed asexual stage parasites.
- Cellular debris from parasite-free *A. stephansii* and non-infected human erythrocytes were used as controls for the sporozoite and trophozoite samples, respectively.

MudPIT: How the proteome was analyzed

- **MudPIT** (Multidimensional Protein Identification Technology) is a version of **DALPC** (Direct Analysis of Large Protein Complexes).

- **How DALPC Works:**
  - Proteins are fragmented into peptides.
  - Peptides are separated by two-dimensional liquid column chromatography.
  - Each peptide, once separated, is eluted directly from the column into a mass spectrometer.
  - The peptides, "sequenced" by the mass spec, are mapped to the genome using the program SEQUEST: http://fields.scripps.edu/sequest/

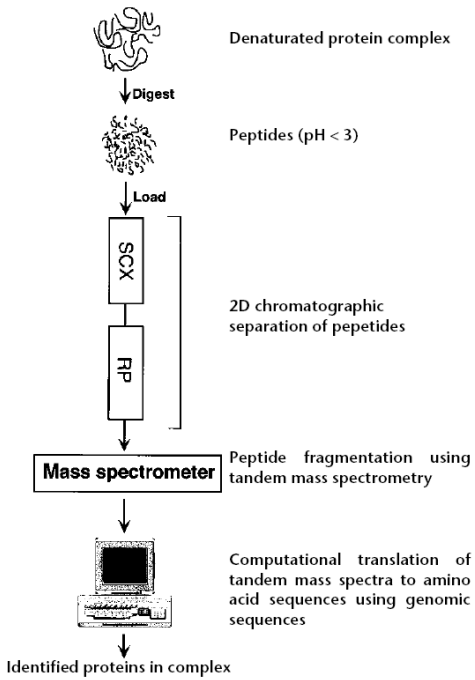
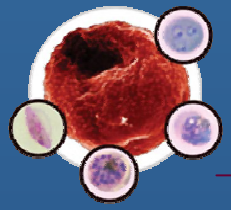


Figure 1. Direct analysis of large protein complexes (DALPC). In the flow diagram, the rectangles represent a strong cation exchange (SCX) and a reversed-phase (RP) liquid chromatography column. Typically, a denatured protein complex is digested with trypsin. The acidified peptide mixture is loaded onto the SCX column. A discrete fraction of peptides is displaced from the SCX column to the RP column. This fraction is eluted from the RP column into the mass spectrometer. This iterative process is repeated, obtaining the fragmentation patterns of peptides in the original peptide mixture. The program SEQUEST is used to correlate the tandem mass spectra of fragmented peptides to amino acid sequences using nucleotide databases[6]. The filtered outputs from the program are used to identify the proteins in the original protein complex.

Image from: Link AJ et al. Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol. 1999 Jul;17(7):676-82.



MudPIT, part II: what the data looks like

Supplementary Table 1: *Plasmodium falciparum* Proteins Expressed in Sporozoites, Merozoites, Trophozoites and Gametocytes as determined by MudPIT analysis.

[a] Protein sequences used for this analysis can be found at: http://fields.scripps.edu/Pf_proteomics/
[b] In the rows showing the locus names, the sequence coverage (percentage of the protein sequence covered by identified peptides) measured for a particular stage are reported. When comparing the same protein across different stages, the percentage of a protein sequence covered by identified peptides and/or the number of peptides can be considered empirical parameters for protein abundance. In the rows showing peptide sequences, the cross-correlation scores (Xcorr) measured by SEQUEST for the peptide/spectrum matches are reported for each of the *P. falciparum* stages analysed: sporozoite (Spz), merozoite (Mrz), trophozoite (Tpz), and gametocyte (Gmt). The dots (.) in the peptide sequence represent the protease cleavage sites. The charge state of the precursor ion is reported as +1, +2 or +3 next to the peptide sequence.

Locus[a]	Spz[b]	Mrz[b]	Tpz[b]	Gmt[b]	Description (as of August 20th, 2002)
PFL0590c	2.7	3.6	1.5	2.2	p-type tpase, putative
E.KNDDNNSIVK.V +1	1.8264				
K.EGLNNINGEK.N +1	2.5746				
K.GKVENINPLSDIVHPELR.L +2		3.2326	3.7884		
K.NANVRKLPAVETLG.C +1	2.2319				
K.TSDGYNDELEGILK.K +2		4.1825			
K.YVNYDENGGFIPMGYVASFDPPRPGVK.E +3				4.1491	
R.LDDLEVTFNSSR.K +2		3.29			
PFL0670c	2.1	5.1	2.1	2.5	Bi-functional aminoacyl-tRNA synthetase, putative
K.KLNVENSYFPLFVTK.N +2		3.7486			
K.LVFDILDLYR.R +2		3.397			
K.NMQEVIIHPLHNYSSLYIK.T +2				3.0627	
K.NMQEVIIHPLHNYSSLYIK.T +3				4.0252	
R.AIQAATSHYLGTNFAK.M +2	5.1825		3.6307		
R.RWYEEYLAVPIIK.G +2		2.6531			
R.WYEEYLAVPIIK.G +2		3.7741			

Comparative proteomics of the malarial life cycle

- **Only 152 proteins (6%) were common to all stages.** These were mainly "house-keeping" proteins, such as ribosomal proteins, transcription factors, histones, and cytoskeletal proteins.
- The sporozoite proteome was markedly different from that of the blood stages, with 49% of its proteins unique to this stage.
- The blood stages had between 22% and 33% unique proteins, and shared between 39% and 56% of their proteins.

Table 1 **Comparative summary of the protein lists for each stage**

Protein count	Sporozoites	Merozoites	Trophozoites	Gametocytes
152	X	X	X	X
197	–	X	X	X
53	X	–	X	X
28	X	X	–	X
36	X	X	X	–
148	–	–	X	X
73	–	X	–	X
120	X	–	–	X
84	–	X	X	–
80	X	–	X	–
65	X	X	–	–
376	–	–	–	X
286	–	–	X	–
204	–	X	–	–
513	X	–	–	–
2,415	1,049	839	1,036	1,147

Whole-cell protein lysates were obtained from, on average, 17×10^6 sporozoites, 4.5×10^9 trophozoites, 2.75×10^9 merozoites, and 6.5×10^9 gametocytes.

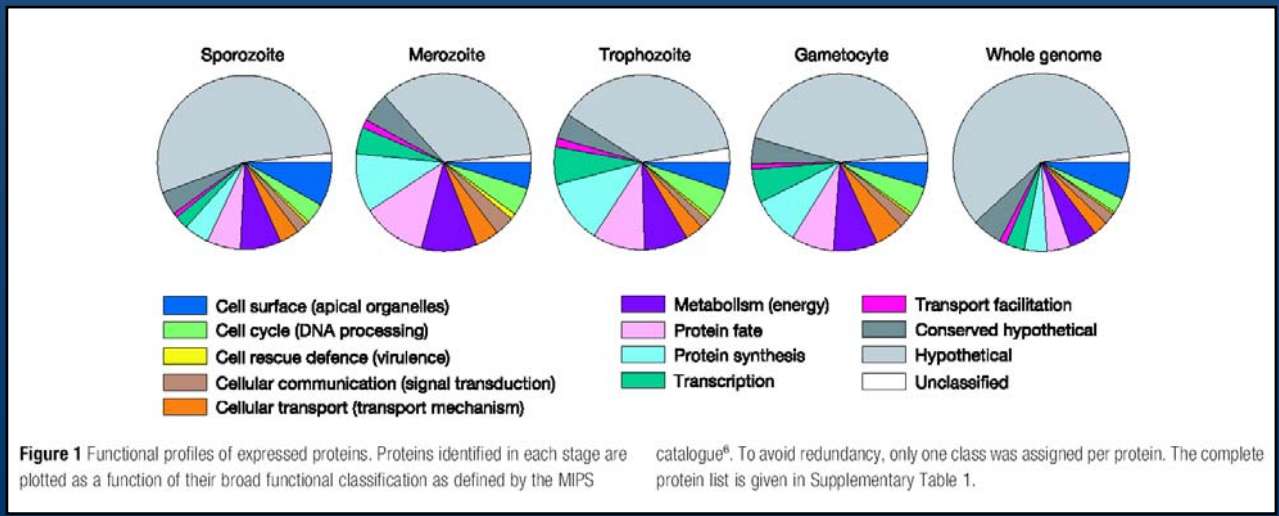
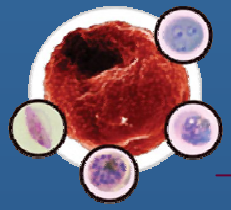
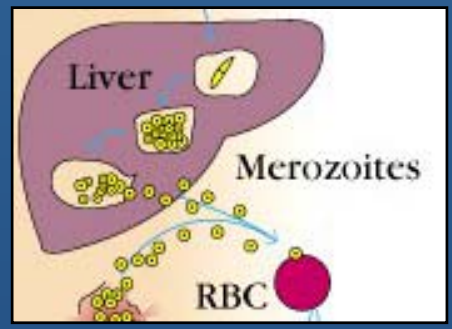


Figure 1 Functional profiles of expressed proteins. Proteins identified in each stage are plotted as a function of their broad functional classification as defined by the MIPS catalogue[6]. To avoid redundancy, only one class was assigned per protein. The complete protein list is given in Supplementary Table 1.

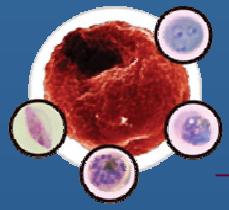


The merozoite proteome

- **Merozoites** infect red blood cells (RBCs). They are generated either by sporozoites in the liver, or trophozoite division in RBCs. Eventually, they will develop into male or female gametocytes.

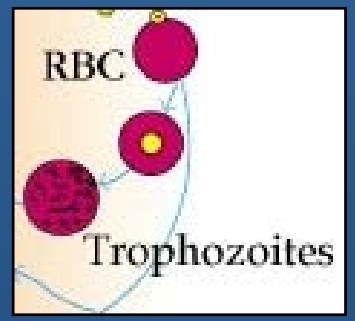


- **Proteins unique or unusually abundant in this stage:**
  - Cell-surface proteins known to be involved in the recognition & invasion of RBCs, including 4 proteins hypothesized to be involved in a specialized actin-myosin "invasion motor"
  - Rhoptry proteins. Merozoites have "rhoptry" organelles, which are thought to be related to host cell remodeling. Only merozoites and sporozoites expressed these proteins.



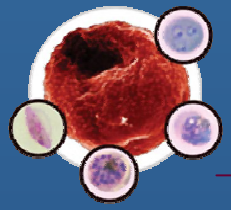
The trophozoite proteome

- **Trophozoites** live within red blood cells (RBCs), feeding on hemoglobin. They eventually undergo division into multiple merozoites, which invade more RBCs.

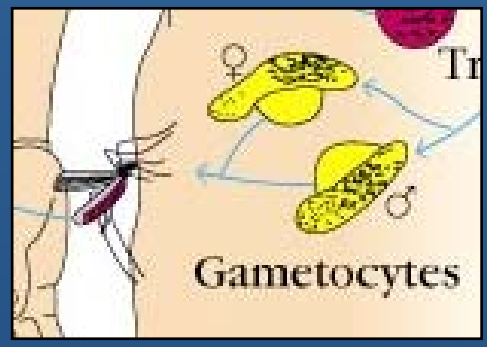


- **Proteins unique or unusually abundant in this stage:**
  - Hemoglobin digestion proteins
  - Proteases not involved in hemoglobin digestion; perhaps involved in RBC invasion, release from RBCs, or merozoite protein processing.
  - Proteins involved in the formation of the parasitophorous vacuole membrane, which is a special membrane that surrounds the parasite while it lives in the RBC; it is unique to the intracelluar trophozoites.
  - Proteins which associate with the RBC membrane, and may be involved in transport of necessary solutes from the RBC surface to the parasite.

- No rifins (proteins derived from the antigenic-shift *rif* genes) were found in trophozoites.

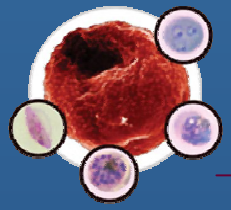


The gametocyte proteome



- **Gametocytes** are produced by differentiation of merozoites. They are ingested by mosquitos and mature into gametes in the mosquito's gut. The gametes will combine to form sporozoites, which can infect humans.

- **Proteins unique or unusually abundant in this stage:**
  - Proteins involved in transcription: transcription factors, RNA-binding proteins, splicing factors, RNA helicases, ribonucleoproteins (RNPs), small nuclear RNPs (snRNPs)
  - Ribosomal proteins – only in the female gametocytes. (for the increase in protein synthesis required during gametogenesis and early zygote development ?).
  - Cell cycle & DNA processing proteins. (Gametocytes are arrested in G0 of the cell cycle until they receive gametogenesis stimuli from the mosquito's gut.)
  - Metabolic proteins associated with respiration.



The sporozoite proteome

- **Sporozoites** are transmitted into humans through the bite of an infected mosquito. They travel to the liver, where they mature and divide into merozoites. A single sporozoite can produce 5,000-10,000 merozoites.

- **Proteins unique or unusually abundant in this stage:**
  - Proteins known to be involved in host invasion and motility, including actin, myosin, rhoptry, and cell-adhesion proteins
  - Known cell-surface markers associated with this stage.
  - Metabolic proteins associated with respiration.
  - Proteins from **Multiple** *var* and *rif* genes. It was previously thought that only one of these "antigenic shift" genes was expressed at a time. The sporozoite, which is unable to undergo antigenic shift during its time in the bloodstream, multiple expression may help it evade recognition.
    NOTE: *var* and *rif* expression overlap with other stages of only of 2 *rif* and 10 *var*.

Chromosomal clusters encoding co-expressed proteins

- Distinct proteomes of each stage suggest a highly coordinated expression of genes involved in common processes..
- 98 clusters containing 3 or more consecutive loci co-expressed in a single stage were found, indicating that, as in *H. sapiens* and *S. cervisea*, functionally related genes cluster in *P. falciparum*.
- Some examples of functionally related clusters:
  - A cluster on chr14 contained 4 aspartic acid proteases (involved in hemoglobin degradation). This cluster was found in all the blood stages, but not in sporozoites.
  - Of the 10 sporozoite-specific clusters, 5 involved *var* and *rif* genes.
  - A cluster on chr11 highly specific to trophozoites contained known surface proteins and three hypothetical proteins, which may be involved in adherance.

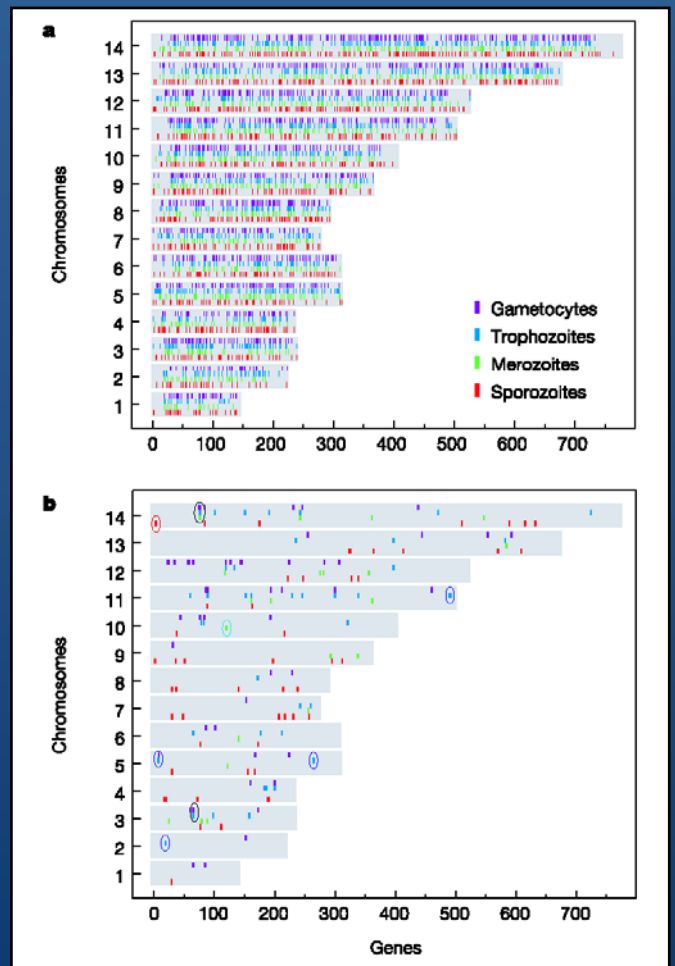
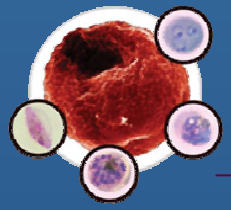


Figure 3 Distribution of expressed proteins by chromosome. **a**, For each stage, genes whose products were detected (coloured vertical bars) are plotted in the order they appear on their chromosome (grey boxes). **b**, Groups of at least three consecutive expressed genes are defined as chromosomal clusters of co-expressed proteins. Examples of such clusters, circled in **b**, are specified in Table 3 and the complete description of the 138 clusters can be found in Supplementary Table 3.



A probasbilistic analysis of chromosomal clustering

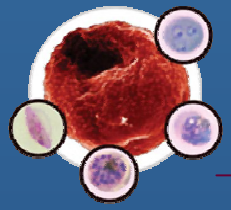
- The probability of finding *m* consecutive genes by chance alone is the number of favorable cases (the ways of having *m* adjacent genes in a sequence of *n* genes) divided by the number of ways of selecting *m* out of *n* genes:

  m = # of consecutive genes
  n = total #of genes on chromosome

$$p_m = \frac{n - m + 1}{\binom{n}{m}}$$

- To compare P_m to the absolute frequency within a chromosome, P_m was multiplied by the number of proteins found on the chromosome. The χ^2 value for each chromosome was calculated for both absolute frequency and absolute probability using 3 degrees of freedom (3 stages –1).

- In all cases, χ^2 values were greater than the 99.9% confidence interval threshold, and therefore the clusters were detected with greater frequency than would occur by random chance.



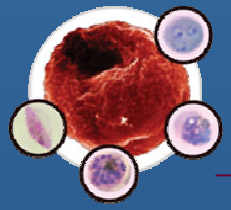
Discussion

- Validation of the dataset is very important when 51% of of the 2,415 proteins identified are annotated as hypothetical. Expression of known proteins was consistent with previously published expression profiles, which provides a reasonable validation.

- Over half of the secreted and integral membrane proteins detected were annotated as hypothetical, and provide new possible targets for drug and vaccine development.

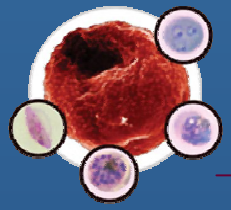
- It has been shown that the sporozoite stage expresses large and varied numbers of *var* and *rif* protein products, which were previously thought to be limited to the blood stages of the parasite.

- Chromosomal clustering of genes encoding co-expressed proteins at specific stages demonstrates a high order of organization. Such clusters may aid in understanding *Plasmodium* gene expression regulation.



Conclusion

The malaria parasite is a complex multi-stage organism, which has co-evolved in mosquitoes and vertebrates for millions of years. Designing drugs or vaccines that substantially and persistently interrupt the life cycle of this complex parasite will require a comprehensive understanding of its biology. The *P. falciparum* genome sequence and comparative proteomics approaches may initiate new strategies for controlling the devastating disease caused by this parasite.



How does being a heterozygote for sickle-cell anemia block malaria?

When oxygen levels drop in the bloodstreams of homozygotes, their red blood cells change from round to sickle-shaped. This shape causes them to get tangled up and clog the spleen and blood vessels.

When the mosquito bites a human, it passes the malaria microbe into the human's bloodstream. The microbe enters the blood cell, where it uses up oxygen which causes the cell to change to the sickle shape. Because of their shape, the sickle cells are more easily filtered out of the bloodstream by the spleen and are soon killed by white blood cells, which gather in the spleen. This keeps the infected cell from bursting open and infecting other cells!

Answer by "Dr. Universe",
http://www.wsu.edu/DrUniverse/evol2.html