Ann Ist Super Sanità 2005;41(4):469-477

Functional genomics, new tools in malaria research

Francesco DI GIROLAMO, Carla RAGGI, Emanuele BULTRINI,
Alessandra LANFRANCOTTI, Francesco SILVESTRINI, Massimo SARGIACOMO,
Cecilia BIRAGO, Elisabetta PIZZI, Pietro ALANO and Marta PONZI

Dipartimento di Malattie Infettive, Parassitarie ed Immunomediate,
Istituto Superiore di Sanità, Rome, Italy

Summary. - The mosquito-transmitted unicellular parasite *Plasmodium falciparum*, the agent of malaria disease, still causes more than one million deaths every year in the tropical and subtropical areas of the world. New intervention strategies are needed to contrast the insurgence of resistance to effective drugs and insecticides. The complete annotated genomes of the human parasite *P. falciparum* and the rodent model *P. yoelii* is now available thus providing a prediction of their possible gene products. This makes feasible the application of functional genomics to malaria research with the final goal of providing a complete survey of *Plasmodium* life cycle. Genome-wide approaches to the study of transcriptome or proteome were successfully applied to malaria parasite with the promise for new drug and vaccine candidates in the next future.

Key words: *plasmodium*, transcriptome, proteome, bioinformatics.

Riassunto (*Genomica funzionale, nuovi strumenti nella ricerca in malaria*). - Il parassita unicellulare *Plasmodium falciparum*, trasmesso attraverso la zanzara, è l'agente eziologico della malaria, malattia che ancora causa più di un milione di morti l'anno nelle aree tropicali e subtropicali del mondo. L'insorgenza di resistenze a farmaci ed insetticidi comunemente utilizzati contribuiscono alla diffusione di questa malattia e rendono sempre più urgente il disegno di nuove strategie di intervento. La disponibilità della sequenza completa del genoma del parassita umano *P. falciparum* e del modello dei roditori *P. yoelii* e la conseguente predizione dei potenziali prodotti genici rendono possibile l'applicazione degli approcci metodologici della genomica funzionale alla ricerca in malaria. Lo studio globale del trascrittoma e del proteoma del parassita fornirà un quadro completo del suo ciclo vitale e nuovi candidati per il disegno di farmaci o vaccini.

Parole chiave: *plasmodium*, trascrittoma, proteoma, bioinformatica.

Introduction

Despite the efforts to eradicate or control malaria, nearly ten per cent of the global population suffer from this disease. With 300-500 million cases each year and up to 2.7 million of deaths [1], malaria still remains a major threat of countries in the tropical and subtropical areas. *Plasmodium*, the causal agent, is a unicellular parasite, haploid for most of the complex life cycle, which sequentially involves a vertebrate host, and a mosquito vector.

Of the species that infect humans, *Plasmodium falciparum* is responsible for the most severe form of malaria. The main factors contributing to resurgence of this disease are the development of parasites resistant to effective and inexpensive drugs, the appearance of mosquitoes resistant to insecticides, the lack of a licensed malaria vaccine and the complexity of parasite life cycle. The asexual stages, which multiply within red blood cells of a vertebrate host, are responsible of the symptoms of malaria disease, while a fraction of parasite committed to differentiate into sexual stages (gametocytes) are responsible of transmission to the vector. Once taken up by mosquito, maturation into male and female gametes is triggered followed by fertilization and meiosis. Mosquito asexual cycle leads to the formation of infective stages (sporozoites), which are accumulated in salivary glands and successively injected in the bloodstream of a new vertebrate host.

New tools are needed to combat malaria and the most critical challenge is to better understand the function of *P. falciparum* genes and their biological role to support the development of new antimalarial strategies. In 1996 an international consortium was

Indirizzo per la corrispondenza (Address for correspondence): Marta Ponzi, Dipartimento di Malattie Infettive, Parassitarie e Immunomediate, Istituto Superiore di Sanità, V.le Regina Elena 299, 00161 Roma. E-mail: mponzi@iss.it.

established to determine the sequence of the *P. falciparum* genome resulting in its publication in 2002 [2]. Several additional *Plasmodium* genome sequencing projects are now under way which will allow comparative and phylogenetic analysis. Further, the availability of the genome sequence of key species in the life cycle of malaria parasites: human, mouse and the most important mosquito vector, the *Anopheles gambiae* [3], will be of great value for coevolution studies of these organisms.

Current state of *Plasmodium* genome sequencing projects

In addition to the complete and ordered sequence of the laboratory-maintained *P. falciparum* line 3D7 and the complete sequence of one of the rodent malaria parasites, *P. yoelii* [4] performed by a consortium of laboratories at the Welcome Trust Sanger Institute (WTSI), the Institute for Genomics research (TIGR) and Stanford University, WTSI and TIGR are now sequencing several additional *Plasmodium* parasites: the rodent *P. chabaudi* and *P. berghei*; two non-human primate parasites, *P. knowlesi* and *P. reichenowi*; an avian parasite, *P. gallinaceum*; a second human parasite *P. vivax*. There is also a project to sequence a *P. falciparum* isolate obtained directly from a patient to fill the gap of sequence information due to chromosomal deletion observed in laboratory-maintained lines. Comparative genomic analysis revealed that approximately 60% of genes are shared across all *Plasmodium* species and their arrangement along chromosomes is maintained [4]. This synteny will assist in the functional characterization of the encoded proteins of unknown function, which represent the majority of the predicted genes. Conservation among species breaks down at the ends of chromosomes where species-specific genes, often involved in host-parasite interaction are clustered. The analysis of these terminal regions will allow identifying factor conferring host and tissue specificity, such as those implicated in immune evasion and host cell adhesion [2, 5].

Novel human genes implicated in malaria pathogenesis or in host responses might be identified e.g. through the surveys of single nucleotide polymorphism (SNP) specific for humans living in endemic areas. While parasite genes that interact with the mosquito host factor or novel genes important for parasite survival within the vector might contribute to disegn transmission blocking drugs or vaccines.

The availability of *Plasmodium* genome sequence has the potential to reveal a high number of new drug targets and genes important to parasite biology and pathogenesis. However, only a third of the predicted genes code for proteins the function of which can be identified. Due to the limited genetic tools and the complexity of parasite life cycle conventional methods cannot be routinely applied for functional studies. Recent advances in genome-wide transcription profiling through microarray technology [6, 7], proteome analysis of different developmental stages [8, 9] and subcellular compartments as well as computational biology, will allow exploitation of genome sequence data in a systematic way leading to a more comprehensive survey of the *Plasmodium* life cycle.

The transcriptome

The word "transcriptome" is gaining increasing popularity since genome sequence data and functional genomic approaches became constitutive elements of modern biology. Transcriptome is the complement of mature messenger RNAs produced in a given cell in a given moment of its life. Although the transcriptome is not rigidly translated in the proteome of that cell, nevertheless it is in most cases largely representative of the cell protein population, and, importantly, it represents a key step in executing the genetic program of differentiation ongoing in that cell.

From the disclosure of the genome sequence of *P. falciparum*, and, more recently, with the availability of genome data from rodent Plasmodia, transcriptome analysis became a key tool in investigating the biology of the malaria parasite. A recent technology of major importance for these studies is represented by DNA microarrays. DNA sequences specific for all predicted malaria genes, in the form of cDNAs or long oligonucleotides, can be placed in ordered arrays in minute portions of a glass slide, where they can be hybridised by fluorescent cDNAs representing the mRNA complement of the parasite. Our laboratory utilises the type of microarrays shown in Fig. 1, panel A, in which about 7500 long oligonucleotides representative of about 4500 *P. falciparum* genes are spotted in an area of approximately 2 cm^2. This microarray is hybridised with cDNA from the parasite sample under study, usually labelled with the red fluorescent Cyanine-5, and simultaneously with a control or reference cDNA, labelled with the green fluorescent Cyanine-3. The ratio of red versus green fluorescence for each spot of the microarray provides a quantitative measure of the relative expression level of the gene corresponding to that spot. Simple direct comparisons between two parasite stages or treatments, or more complex time course analyses can thus be performed in order to study how the entire repertoire of Plasmodium genes are regulated under the conditions of an experiment.

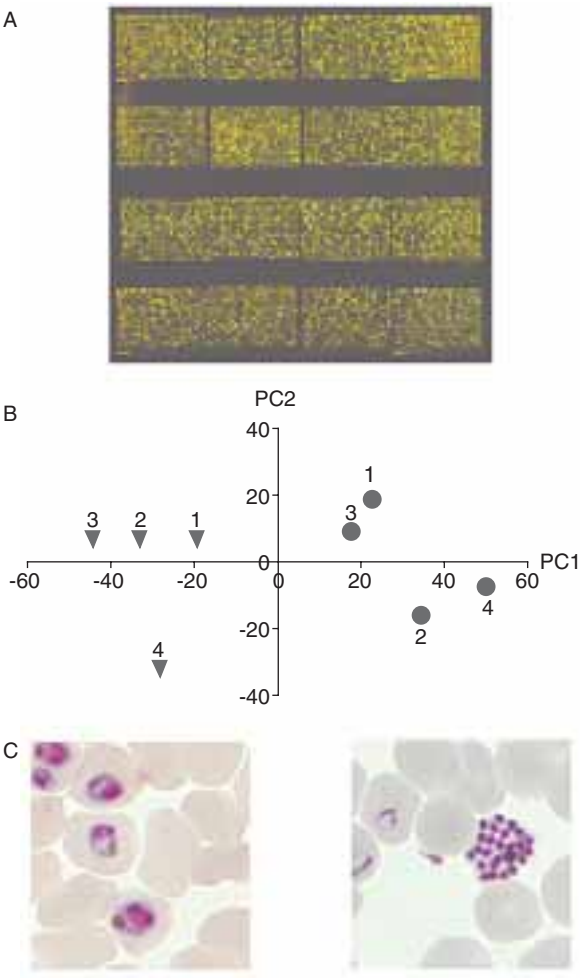

Fig. 1. - *P. falciparum* genome-wide expression analysis with DNA microarrays.

Panel A. A microarray with 7462 long oligonucleotides (70-mers), representing 4488 predicted genes of *P. falciparum*, spotted on a glass slide in approximately 2 cm^2, and hybridised with Cyanine-5 and Cyanine-3 parasite cDNA samples. **Panel B**. Principal Component Analysis (PCA) of expression data from eight cDNA samples of *P. falciparum*: parasites from 4 independent synchronous cultures (1-4) were sampled at 30-34 h and at 40-44 h post erythrocyte invasion, and the eight cDNAs obtained were hybridised against an identical control cDNA on eight microarray slides. After hybridisation quality control, expression values (relative transcript abundance) were obtained for 1938 genes. As each expression data set could thus be described by the 1938 gene-specific variables, PCA was used to identify the major sources of variance between the data sets. The analysis determined that Principal Component (PC) 1 and PC2 were responsible for 63% and 15% of total variance, respectively. Using the new bidimensional space defined by the values of PC1 and PC2, data sets from parasites obtained at 30-34 h (triangles) were thus grouped together, and were clearly separated by those from parasites obtained at 40-44 h (circles). **Panel C**. Morphological differences between *P. falciparum* parasites at 30-34 h post invasion (trophozoites), on the left, and at 40-44 h post-invasion (schizont), on the right.

A few reference studies have been conducted so far with the technology of microarray on human and rodent malaria parasites, and provided relevant cues to understanding *Plasmodium* biology and development. One of them revealed that in the 48 hours of *P. falciparum* asexual cycle, the parasite activates sequential waves of gene expression in which almost 80% of its gene complement is positively regulated [7]. Other studies on *P. falciparum*, *P. berghei* and *P. yoelii* confirmed that the different developmental stages of the parasite are accompanied by the expression of large sets of stage-specific transcripts, and indicated that the physiology of each stage is a highly elaborate interaction of constituve and specialised parasite molecules [6, 7].

Our group utilised microarrays specific for *P. falciparum* with the objective to study a very particular moment of the parasite life cycle: the early phases of parasite transformation in the gametocyte, the *Plasmodium* sexual cell. The gametocyte is responsible for the transmission of the parasite from the bloodstream of an infected individual to the mosquito, and it therefore represents a key step in the ability of *Plasmodium* to spread malaria. While gametocytes in course of maturation, after about 5-6 days of differentiation, have been shown to express specific sets of mRNAs in transcriptome and proteome analyses [6, 8, 9], the early gametocytes have been comparatively much less studied, and only two specific molecules - Pfs16 and Pfg27- were described in that stage of differentiation. Since early gametocytes cannot be distinguished from small asexual forms, apart from expression of the mentioned early markers, we designed a time course experiment in which parasites producing young gametocytes in the first 40 hours of differentiation were compared with isogenic parasites unable to undergo sexual differentiation. In this experiment RNAs from various time points were hybridised on microarrays, and genes were identified which had an expression profile significantly correlated to the profiles of the genes for the early markers Pfs16 and Pfg27. Further characterisation of this group of genes led so far to identify two novel proteins specifically expressed from early stages of differentiation [10]. This analysis and similar studies conducted in other laboratories [11, 12] were able to significantly increase in few years our knowledge of the specific molecular changes occurring in this underexplored moment of parasite differentiation, and raised new hopes to find effective ways to inhibit specific mechanisms active in early gametocytes with the aim to block parasite transmission.

Transcriptome analysis however provides the opportunity to investigate several other aspects of Plasmodium biology. Understanding parasite response

to drug treatment, or identification of key specific molecules required for essential steps of its development are potentially relevant to identification of novel targets for pharmacological and immunological attack. This is a much needed activity given that the fast spread of parasite drug resistance is leaving little hope that current available antimalarial armoury will be adequate to combat the disease in the span of few years. Beside such applicative aspects, transcriptome analysis can also be utilised to enquire fundamental issues of *Plasmodium* biology, such as how the parasite controls the expression of its genes. The little knowledge available to date on this issue indicates that mechanisms controlling gene expression in *Plasmodium* recur to molecular components highly specific to the parasite, which strongly suggests that these mechanisms are likely to be attackable without significant interference and damages to homologous molecular components of the human host.

Another example of how the parasite transcriptome can be analysed is presented in Fig. 1. Four independent *in-vitro* cultures of *P. falciparum* were synchronised, parasites were sampled and their RNA extracted at two time points, at 30-34 hours and at 40-44 hours after invasion of the red blood cell. The eight cDNA samples thus obtained were all labelled with Cyanine-5, and hybridised on eight microarrays against identical aliquots of a reference cDNA from asynchronous parasites, labelled with Cyanine-3, producing eight gene expression data sets. Between the analytical tools for analysis of microarray expression data, we present here the analysis conducted with the method of Principal Component Analysis (PCA). According to this statistical technique, each of the eight experiments is viewed in a multidimensional space whose coordinates were the expression values of the 1938 genes which gave quality-controlled signals in the eight hybridisations. By applying PCA, relationships between coordinates were captured in order to find a new space in which diversities between experiments (variance) are highlighted. As it is evident from the distribution of points (experiments) in the graphic (Fig. 2, panel B), the analysis clearly distinguishes the transcriptome of parasites sampled at 30-34 hours (triangles) from that of parasite sampled ten hours later (circles). This molecular difference is clearly reflected in the morphological differences exhibited by the parasites in the two time points, i.e. the trophozoite (T) and the schizont (S), respectively (Fig. 1, panel C). In conclusion, this simply designed experiment shows that transcriptome analysis is a sentitive tool to monitor molecular differences associated to parasite development. Moreover, a closer analysis of the transcripts upregulated in one or in the other parasite stage can provide useful information on molecules active at different moment of the parasite asexual cycle. In more complex approaches, genome sequence information can be integrated in the analysis of the trancriptome in order to address questions on the regulatory mechanisms governing gene expression in *P. falciparum*. Transcriptome data such as those obtained in the experiment presented can for instance be utilised for inspecting genomic sequences upstream of large number of genes showing concomitant upregulation.

Table 1. - Malaria resources on the Web

Welcome Trust Sanger Institute (WTSI; http://www.sanger.ac.uk)
The Institute for Genomics research (TIGR; http://www.tigr.org)
Stanford University (http://www.stanford.edu)
Malaria Trascriptome Database (http://malaria.ucsf.edu/)
Malaria full-length cDNA database (http://fullman.ims.u-tokyo.ac.jp/).
PlasmoDB (http://www.plasmodb.org)
GeneDB (http://www.genedb.org)

Bioinformatics

Analyses of the huge amount of data produced by wide-throughput approaches for studying genome, trascriptome and proteome required to use statistical, mathematical and informatic tools specifically designed for these biological data. Even if it is not a recent science, today bioinformatics is considered an essential body of knowledge to manage and interpret experimental results, to approach and solve biological problems.

After completion of a genome sequencing a crucial step is genome annotation. This implies the use of collections of fast and powerful bioinformatic software in order to identify genes and to perform reliable predictions on the function and/or structure of their protein products. Although many progresses have been done on this topic, a complete genome annotation should also include identification of other essential parts of a genome such as regulatory regions (promoters). This is one of the most difficult tasks for bioinformaticians, due to the complex combinatorial nature of these regions and to the fact that sequences, among regulatory elements (transcription factor binding sites), are weakly conserved. Several methods have been, recently, developed to approach this problem. Some of them are based on the knowledge of gene expression data, the basic idea being that genes sharing a similar expression profile probably share regulatory elements too [13].

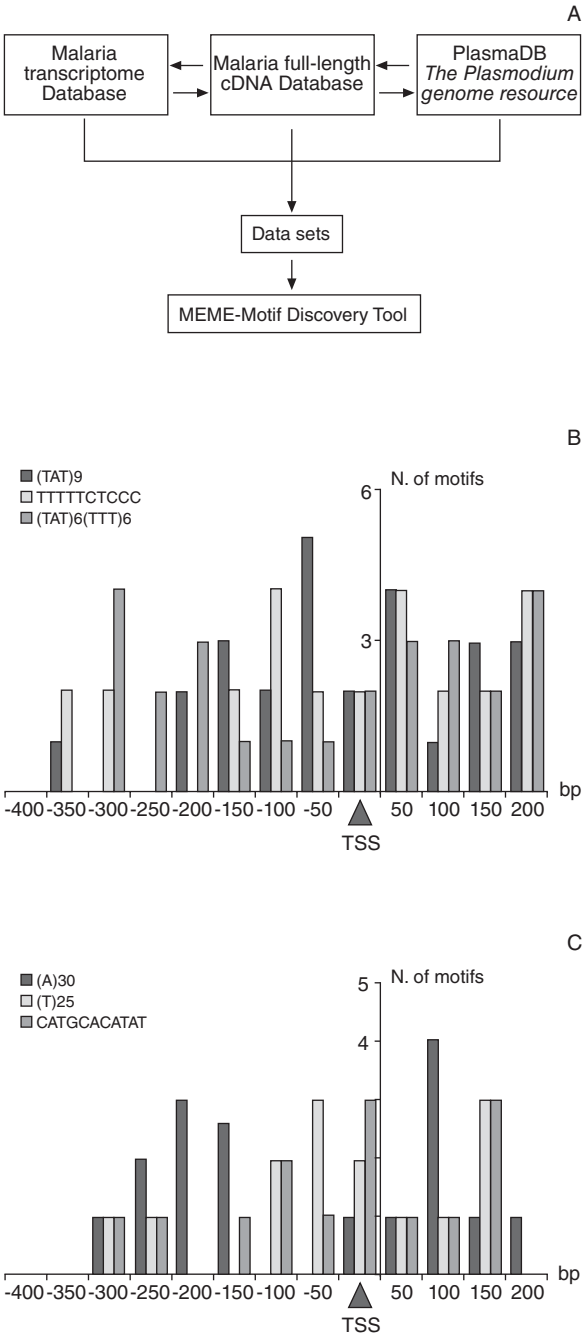

Fig. 2. - A bioinformatic approach to the identification of conserved motifs with a potential biological relevance. **A)** Data stored in three specialized databases (Malaria Trascriptome Database; Malaria full-length cDNA database; PlasmoDB) can be integrated to select suitable sets of sequences to be submitted to a motif discovery tool (MEME). **B)** Distribution of statistically significant motifs around TSSs for two groups of genes up-regulated either in the early phases (1) or in the late phases (2) of intraerythrocyte cycle of *P. falciparum*. Each motif is represented by a bar, the height of which corresponds to the number of occurrences (calculated in windows of 50 bp) along the sequences.

We approached the search for regulatory elements in *P. falciparum* by integrating collections of experimental data coming from genome annotation, full-length cDNA sequencing and gene expression microarrays results (the rationale of this approach is reported in Fig. 2, panel A). In the case of *P. falciparum* this is made possible by the availability of specialized databases in which all these data are organized (Table 1) [14]. We proceeded following successive steps: first, we used Malaria Trascriptome Database (see Table 1, for references) to select two groups of genes differently expressed during the intraerythrocytic stages of the parasite. The first one comprises 135 genes whose expression profile peaks at the 12th hour of the cycle; the second group is formed by 58 genes that are up-regulated after the 30th hour. In a second step, we retrieved the transcriptional start sites (TSS) available from Malaria full-length cDNA database (26 out of 135 in the first group of genes and 14 out of 58 in the second). Sequences surrounding the TSS (from -400 bp to +200 bp) were extracted from PlasmoDB (Table 1), and submitted to a motif discovery tool (MEME version 3.0) [15]. This software looks for sequence patterns that occur repeatedly, with a statistical significance, in a set of unaligned DNA sequences. In our case the presence of conserved patterns might be related to transcription factor binding sites, thus providing hints on the location of potential regulatory regions. The two groups of selected sequences were analysed separately for patterns occurring at least once for sequence. The identified patterns and their distribution around TSS sequences are shown in Fig. 2 panels B and C. The upstream regions of the first group of genes (up-regulated during the early intraerythrocytic stages) contain TAT and TTT repeats that accumulate from -300 bp to +200 bp around the TSS and a C-rich pattern distributed in two peaks at -100 bp and +200 bp. The upstream regions of the second group of genes (upregulated during the late phases of intraerythrocytic development) contain long runs of A (with a peak at +100 bp) and runs of T with two maxima at -50 bp and -150 bp with a GC-rich element associated in the latter position.

It is well known that runs of A, T or TA repeats reflect conformational changes in DNA molecule and it was suggested that TSS are often associated to them [16]. Recently it was demonstrated that long tracts of A and T affect the activity of calmodulin promoter in *P. falciparum* [17]. Our results indicate that similar motifs are present around TSS of a number of *P. falciparum* genes, suggesting for them a role in promoter activity. Experimental data are required to validate the selected candidate sequences.

As shown in the study presented here, bioinformatic tools are invaluable to extend analysis from single genes, proteins, or intergenic regions to an entire genome in order to predict candidate sequences that share structural/functional features.

Two studies [18, 19] have been, recently, conducted on *P. falciparum*, in which a unique motif was identified and functionally characterized as a sequence which allows translocation of proteins to the host cell compartment (HCT, host cell-targeting). A great value of this finding is the possibility to predict through a search in *Plasmodium* databases novel proteins containing this consensus sequence and hence potentially exported beyond parasite membrane. More than three hundred *P. falciparum* proteins contain this targeting signal suggesting that an extensive remodelling of the host cell occurs during intraerythrocytic cycle. These studies again confirm how bioinformatics provides a valide support to researchers for disegning new experimental approaches in order to characterize molecules implicated in processes crucial for parasite development or host/parasite interplay.

The proteome

The identification of a large number of proteins in complex biological samples was made feasible by the application of mass spectrometry to generate peptide fingerprints to be compared with peptides predicted from genome databases. Two general approaches can be adopted in proteomic studies: protein samples can be first separated through one or bidimensional gel electrophoresis and successively submitted to protease digestion or, in alternative, protein mixtures can be directly treated with proteases and the generated peptides successively separated using liquid chromatography (LC) (for review see ref [20]).

However, one of the key challenges in the study of the proteome of endocellular parasites, as the case of *Plasmodium*, is to obtain a sufficient material with a low level of contamination from host cell. This limits the analysis of particular stages (e.g. hepatic or sexual stages) or subcellular fractions. To overcome these problems selection methods such as magnetic beads coated with specific monoclonal antibody were successfully used to purify invasive mosquito stage (sporozoite) [8]. Male and female gametocytes were, instead, separated by flow cytometry using transgenic lines expressing the green fluorescent protein (GFP) under the control of sex-specific promoters [21].

Data on both rodent [22] and human [8, 9] malaria parasite proteome of invasive, replicative or sexual stages were obtained and showed a high degree of specialization between the different developmental stages, revealed by the presence of a high number of stage-specific proteins.

Further, the complete proteome analysis of human erythrocyte [23] led to the identification of 181 unique proteins half of which reside in the plasma membrane and half in the cytoplasm, most of them were also categorized according to their function. This might provides a basis for the study of metabolic and structural changes induced by malaria infection, thus contributing to unravel the complex host-parasite relationships.

Combined transcriptomic and proteomic data allowed also to uncover regulatory mechanisms of gene expression. It was observed for a number of genes involved in gametocyte differentiation that transcriptional and translational processes are not coupled but occur at different stages of sexual development, suggesting for them a posttranscriptional control of gene expression [22]. This confirms bioinformatic analysis which predicts from whole genome annotation an abundance of proteins modulating mRNA decay and translation rates [24] that might indicate that protein levels in malaria parasite are significantly controlled through mRNA processing and translation [25].

Proteomics of subcellular parasite compartments

Malaria parasite has the peculiar ability to invade highly specialized red blood cells, and to develop within a parasitophorous vacuole, which represents the interface between parasite and the host cell environment. The infected erythrocytes have no intracellular organelles and are devoid of protein synthesis and trafficking. In order to ensure its own survival *Plasmodium* remodels the host cell by exporting proteins into the cytoplasm and plasma membrane of the erythrocyte and generating new subcellular compartments within the infected erythrocytes which allow the parasite to establish mechanisms of immune evasion and create new permeation pathways for nutrient uptake. A complex tubovescicular membrane network is formed [26] that extends from the parasitophorous vacuole throughout the erythrocyte cytoplasm which most probably associates with flattened vesicular structures beneath the erythrocyte membrane, a Golgi like compartments called Maurer's clefts, involved in translocation of parasite-encoded proteins to the erythrocyte surface (see ref [27] for review). The infected erythrocytes acquire adhesive properties and sequester in capillary of internal organs. This allows the parasite to escape passage to the spleen but causes severe pathology to the host [28]. Proteome analysis of subcellular components of *Plasmodium* is at the early days also for the difficulties encountered in their purification. Components of rhoptries, apical secretory organelles of merozoites, most probably involved in parasitophorous vacuole biogenesis, have been identified through proteomic combined with bioinformatic approaches [29]. Among

proteins predicted to be located in this compartment are proteases possibly involved in processing merozoite proteins targeted to the rhoptries, enzymes of lipid metabolism that might be implicated in the establishment of the vacuolar membrane, as well as, proteins known to localize to the vacuolar membrane, finding that supports the involvement of rhoptry components in the genesis of parasitophorous vacuole.

Two complementary studies have been reported: the one on Maurer's cleft proteomics [30] that shed new light on the important biological functions of this parasite-derived compartment. This study confirms that Maurer's cleft have characteristics of a secretory compartment addressing parasite proteins to the red cell surface but also suggests that these structures are not involved only in protein trafficking as indicated by the presence of enzymes, and proteins implicated in signal transduction. The other study [31] applied high throughput proteomics to identify antigens on the surface of infected erythrocyte using a multidimensional protein identification technology (MudPIT), a two-dimensional liquid chromatography coupled with tandem mass spectrometry. The advantage of this approach is its ability to analyse complex protein mixture that are difficult to resolve by gel-based protein separation systems, such as membrane proteins. 36 candidates were selected with the support of bioinformatics for the presence of a predicted signal peptide and/or of transmembrane domain(s). The surface location was confirmed by immunolocalization experiments for two of the selected proteins.

Membrane microdomain proteomics, future perspectives

The traditional view that the plasma membrane is a uniform lipid bilayer containing randomly distributed membrane proteins has given way to a more complex model in which glycosphingolipids associate with cholesterol to form organized "liquid ordered" structures or "lipid rafts" [32]. The physical properties of these microdomains impose barriers to the free diffusion of membrane proteins. They act as molecular sieves admitting certain proteins while excluding others. The ability to dynamically compartmentalize membrane proteins allows synchronizing a wide array of cellular responses linked to signal transduction, transcytosis, cholesterol transport or internalization of pathogens [33]. In addition, evidence is growing that lipid rafts are present not only at the plasma membrane but also internally where they could assist trafficking between membrane compartments.

Microdomains, most probably present in all cells, have been also identified at the erythrocytes membrane and at membrane compartments generated by malaria parasite. As detailed below, lipid rafts are involved in the establishment and maintenance of malaria parasite infection.

It was shown that host proteins associated to lipid rafts can be drawn into the *P. falciparum* vacuole during invasion and intraerythrocytic growth [34]. Raft-cholesterol depletion from infected red blood cells destabilizes the vacuolar membrane and causes parasite expulsion from host cell [34] indicating that the presence of these microdomains is critical for maintaining parasitophorous vacuole. When raft-cholesterol depletion is performed on uninfected red blood cells, a reduction of invasion by 80-90% was observed [35], suggesting that parasite entry might occur via lipid rafts.

A number of parasite proteins have been identified as stably or transiently associated to lipid rafts [36, 37] but a complete analysis of raft protein composition in malaria parasite have not been undertaken. This would be of help to elucidate the raft-dependent functions in the course of parasite infection and possibly the interplay between host and parasite components.

A limited number of characterized raft proteins is available to date (examples are given in refs [38] and [39]), although the isolation of microdomain components exploiting their physico-chemical features (i.e. insolubility in nonionic detergents and low buoyant density) is a well established procedure. Our group isolated proteins associated with lipid rafts from *P. berghei* in order to provide candidate components of this crucial membrane compartment through a comprensive mass spectrometry analysis. Here we present a MALDI-MS analysis aimed at identifying in the complex mixture of raft components PbSEP1, a *P. berghei* protein dinamically associated to this membrane compartment. This protein belongs to a family (SEP) of small exported proteins (13-16 Kda), conserved within *Plasmodium* genus. SEPs are located at the membrane of parasitophorous vacuole and/or at the Maurer's clefts [40]. They are characterised by the presence of a signal peptide, a hydrophobyc region and a highly charged region at the carboxy-terminus. PbSEP 1 was identified in rhoptry proteome of free merozoites [29] while localizes at the vacuolar membrane early after invasion [40], suggesting a role for this protein in the genesis of this parasite compartment.

Lipid rafts were isolated, by buoyant density-gradient centrifugation, from purified *P. berghei* parasites with an intact vacuolar membrane (Fig. 3A). An immunoblot of the collected fractions probed with PbSEP 1 is shown in Fig. 3B. Raft-associated proteins were separated by one dimension gel electrophoresis and stained with Comassie Blue (Fig. 3C). MALDI-MS analysis of a gel area containing proteins ranging

from 13 to 16 KDa confirmed the presence of peptides derived from PbSEP 1 (Fig. 3D). This analysis indicated that integral membrane proteins which are minor components of this specialized proteome, as the case of PbSEP1, can be identified with the applied procedures and that a more comprehensive analysis of raft proteome by mass spectrometry is feasible. We expect that plasmodial rafts will contain proteins involved in the generic structural and functional tasks of membrane microdomains along with a subset of proteins unique to the parasite which might constitute potential targets for intervention.

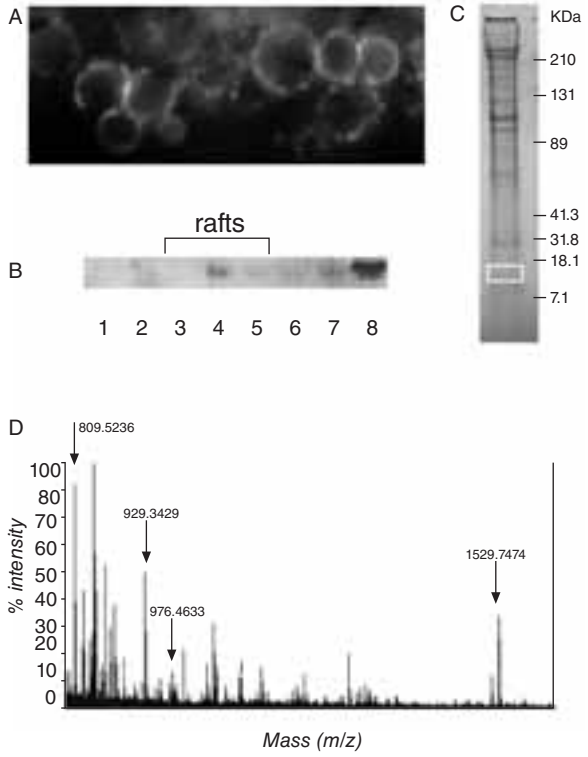

Fig. 3. - *Plasmodium* lipid rafts purification and analysis. *P. berghei* parasites freed from host erythrocytes, but still maintaining parasitophorous vacuole, were used to isolate lipid raft components. Parasite preparation, immmunostained with anti-PbSEP 1 antibodies which specifically decorate the vacuolar membrane is shown in **A**). After parasite lysis with cold Triton X-100, rafts proteins were obtained by buoyant-density fractionation over a discontinuous sucrose-density gradient (5%, 30%, 40%). Lipid rafts could be identified as insoluble material at the 5% - 30% sucrose interface (fractions 3-5), whereas soluble proteins were present within 40% sucrose layer (fractions 8-12) and insoluble nuclear and cytoskeletal material remained beneath the 40% cushion. Immunoblot of gel-separated fractions 1-8 probed with PbSEP 1 immune serum, (shown in **B**) revealed that only a small amount of SEP 1 is associated to lipid rafts while the majority associates with Triton-soluble membranes (fraction 8). Triton-insoluble raft proteins (fractions 3-5) separated by polyacrylamide gel electrophoresis and stained with Comassie Blue are shown in **C**. A gel slice containing proteins between 13 and 16 KDa (white box) was excised and subjected to trypsin digestion to obtain a peptide mixture for MALDI-MS analysis. Peptide mass map of the processed protein band is presented in **D**). Peaks indicated by an arrow are derived from PbSEP 1 digestion.

Acknowledgements

The authors wish to thank the Italian Ministero della Salute and the European Commission (FP6, Network of Excellence BioMalPar, contract N° LSHP CT-2004 503578) for finantial support.

Received on 5 September 2005.
Accepted on 21 September 2005.

REFERENCES

1. Breman JG, Egan A, Keusch GT. The intolerable burden of malaria: a new look at the numbers. *Am J Trop Med Hyg* 2001;64(1-2 Suppl):iv-vii.

2. Gardner MJ *et al*. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 2002;419:498-511.

3. Holt RA *et al*. The genome sequence of the malaria mosquito *Anopheles gambiae. Science* 2002;298:129-49.

4. Carlton JM *et al*. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 2002;419:512-9.

5. del Portillo HA, Fernandez-Becerra C, Bowman S, Oliver K, Preuss M, Sanchez CP, Schneider NK, Villalobos JM, Rajandream MA, Harris D, Pereira da Silva LH, Barrell B, Lanzer M. A superfamily of variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature* 2001;410:839-42.

6. Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 2003;301:1503-8.

7. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. *PLoS Biol* 2003;1:E5.

8. Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL, Witney AA, Wolters D, Wu Y, Gardner MJ, Holder AA, Sinden RE, Yates JR, Carucci DJ A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* 2002;419:520-6.

9. Lasonder E, Ishihama Y, Andersen JS, Vermunt AM, Pain A, Sauerwein RW, Eling WM, Hall N, Waters AP, Stunnenberg HG, Mann M. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* 2002;419:537-42.

10. Silvestrini F, Bozdech Z, Lanfrancotti A, Di Giulio E, Bultrini E, Picci L, deRisi JL, Pizzi E, Alano P. Genome-wide identification of genes upregulated at the onset of gametocytogenesis in *Plasmodium falciparum*. *Mol Biochem Parasitol* 2005;143:100-10.

11. Eksi S, Haile Y, Furuya T, Ma L, Su X, Williamson KC. Identification of a subtelomeric gene family expressed during the asexual-sexual stage transition in *Plasmodium falciparum*. *Mol Biochem Parasitol* 2005;143:90-9.

12. Young JA, Fivelman QL, Blair PL, de la Vega P, Le Roch KG, Zhou Y, Carucci DJ, Baker DA, Winzeler EA. The *Plasmodium falciparum* sexual development transcriptome: A microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol* 2005;143:67-79.

13. Hampson S, Kibler D, Baldi P. Distribution patterns of over-represented k-mers in non-coding yeast DNA. *Bioinformatics* 2002;18:513-28.

14. Kissinger JC, Brunk BP, Crabtree J, Fraunholz MJ, Gajria B, Milgram AJ, Pearson DS, Schug J, Bahl A, Diskin SJ, Ginsburg H, Grant GR, Gupta D, Labo P, Li L, Mailman MD, McWeeney SK, Whetzel P, Stoeckert Jr CJ, Roos DS. The *Plasmodium* genome database. *Nature* 2002;419:490-2.

15. Bailey TL, Gribskov M. Methods and statistics for combining motif match scores. *Journal of Computational Biology* 1998;5:211-21.

16. Pedersen AG, Baldi P, Chauvin Y, Brunak S. The biology of eukaryotic promoter prediction-a review. *Comput Chem* 1999;23:191-207.

17. Polson HE, Blackman MJ. A role for poly(dA)poly(dT) tracts in directing activity of the Plasmodium falciparum calmodulin gene promoter. *Mol Biochem Parasitol* 2005;141:179-89.

18. Marti M, Good RT, Rug M, Knuepfer E, Cowman AF. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* 2004;306:1930-3.

19. Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estrano C, Haldar K. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 2004;306:1934-7.

20. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003;422:198-207.

21. Khan SM, Franke-Fayard B, Mair GR, Lasonder E, Janse CJ, Mann M, Waters AP. Proteome analysis of separated male and female gametocytes reveals novel sex-specific Plasmodium biology. *Cell* 2005;121:675-87.

22. Hall N *et al*. A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 2005;307:82-6.

23. Kakhniashvili DG, Bulla LA Jr, Goodman SR. The human erythrocyte proteome: analysis by ion trap mass spectrometry. *Mol Cell Proteomics* 2004;3:501-9.

24. Coulson RM, Hall N, Ouzounis CA. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res* 2004;14:1548-54.

25. Wirth DF. Biological revelations. *Nature* 2002;419:495-6.

26. Wickert H, Wissing F, Andrews KT, Stich A, Krohne G, Lanzer M. Evidence for trafficking of PfEMP1 to the surface of *P. falciparum*-infected erythrocytes via a complex membrane network. *Eur J Cell Biol* 2003;82:271-84.

27. Przyborski JM, Wickert H, Krohne G, Lanzer M. Maurer's clefts-a novel secretory organelle? *Mol Biochem Parasitol* 2003;132:17-26.

28. Miller LH, Baruch DI, Marsh K, Doumbo OK. The pathogenic basis of malaria. *Nature* 2002;415:673-9.

29. Sam-Yellowe TY, Florens L, Wang T, Raine JD, Carucci DJ, Sinden R, Yates JR 3rd. Proteome analysis of rhoptry-enriched fractions isolated from *Plasmodium* merozoites. *J Proteome Res* 2004;3:995-1001.

30. Vincensini L, Richert S, Blisnick T, Van Dorsselaer A, Leize-Wagner E, Rabilloud T, Braun Breton C. Proteomic analysis identifies novel proteins of the Maurer's clefts, a secretory compartment delivering *Plasmodium falciparum* proteins to the surface of its host cell. *Mol Cell Proteomics* 2005;4:582-93.

31. Florens L, Liu X, Wang Y, Yang S, Schwartz O, Peglar M, Carucci DJ, Yates III JR, Wu Y. Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Mol biochem Parasitol* 2004;135:1-11.

32. Brown DA, London E. Structure and function of sphingolipid- and cholesterol- rich membrane rafts. *J Biol Chem* 2000;275:17221-4.

33. Rosenberger CM, Brumel JH, Finlay BB. Microbial pathogenesis: lipid rafts as pathogens portals. *Curr Biol* 2000;10:R823-5.

34. Lauer S, VanWye J, Harrison T, McManus H, Samuel BU, Hiller NL, Mohandas N, Haldar K. Vacuolar uptake of host components, and a role for cholesterol and sphingomyelin in malarial infection. *EMBO J* 2000;19:3556-64.

35. Samuel BU, Mohandas N, Harrison T, McManus H, Rosse W, Reid M, Haldar K. The role of cholesterol and glycosylphosphatidylinositol-anchored proteins of erythrocyte rafts in regulating raft protein content and malarial infection. *J Biol Chem* 2001;276:29319-29.

36. Wang L, Mohandas N, Thomas A, Coppel RL. Detection of detergent-resistant membranes in asexual blood-stage parasites of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2003:130: 149-53.

37. Hiller NL, Akompong T, Morrow JS, Holder AA, Haldar K. Identification of a stomatin orthologue in vacuoles induced in human erythrocytes by malaria parasites. A role for microbial raft proteins in apicomplexan vacuole biogenesis. *J Biol Chem* 2003;278:48413-21.

38. Blonder J, Hale ML, Lucas DA, Schaefer CF, Yu L-R, Conrads TP, Issaq HJ, Stiles BG, Veenstra TD. Proteomic analysis of detergent-resistant membrane rafts. *Electrophoresis* 2004;25: 1307-18.

39. Foster LJ, de Hoog CL, Mann M. Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Cell Biol* 2003;10:5813-8.

40. Birago C, Albanesi V, Silvestrini F, Picci L, Pizzi E, Alano P, Pace T, Ponzi M. A gene-family encoding small exported proteins is conserved across *Plasmodium* genus. *Mol Biochem Parasitol* 2003;126:209-18.