Review

Cell PRESS

A genotype and phenotype database of genetically modified malaria-parasites

Chris J. Janse[1], Hans Kroeze[1], Auke van Wigcheren[1,2], Senad Mededovic[2], Jannik Fonager[1], Blandine Franke-Fayard[1], Andrew P. Waters[3] and Shahid M. Khan[1]

[1] Leiden Malaria Research Group, Parasitology, Leiden University Medical Centre (LUMC), Albinusdreef 2, 2333 ZA, Leiden, The Netherlands
[2] StudioDrie, Newtonstraat 6 g, 2562 AJ Den Haag, The Hague, The Netherlands
[3] Division of Infection and Immunity, Institute of Biomedical Life Sciences & Wellcome Centre for Molecular Parasitology, Glasgow Biomedical Research Centre, University of Glasgow, 120 University Place, Glasgow, Scotland, UK, G12 8TA

The RMgm database, www.pberghei.eu, is a web-based, manually curated, repository containing information on genetically modified rodent-malaria parasites. It provides easy and rapid access to information on the genotype and phenotype of mutant and reporter parasites. The database also contains information on unpublished mutants without a clear phenotype and negative trials to disrupt genes. Information can be searched using predefined key features, such as phenotype, life-cycle stage, gene model, gene-tags and mutations. The information relating to the mutants is reciprocally linked to PlasmoDB and GeneDB. Access to mutant-parasite information, and gene function/ontology inferred from mutant phenotypes provides a timely resource aimed at enhancing research into *Plasmodium* gene function and (systems) biology.

Genetically modified rodent-malaria parasites and their use in malaria research

The genomes of several different *Plasmodium* species have been almost completely sequenced including those of the most important human malaria parasites, *P. falciparum* and *P. vivax*, as well as three closely related rodent species: *P. chabaudi*, *P. yoelii* and *P. berghei* [1–4]. The increase in knowledge about malaria parasite genomes and genes not only helps us to unravel the molecular and cellular processes the parasite engages in during development through its life cycle but also permits the characterization of parasite–host interactions associated with infection and disease. Comparative analyses of *Plasmodium* genomes and genomes of other organisms have greatly improved the identification and assignation of putative functions to *Plasmodium* genes [5–9]. Although recognition of gene function by orthology is becoming increasingly sophisticated [10], comparative gene analysis has revealed that more 50% of malaria parasite genes cannot be assigned a function, and it is therefore probable that many of these genes perform functions that are unique to *Plasmodium*. Although classical genetics experiments have been performed with *Plasmodium* and yielded fundamental insights into parasite gene function, even these ultimately relied upon the reverse genetic technologies that have been implemented in the past 12 to 15 years in *Plasmodium* research. Such technologies are now the front line methodologies used to gain an understanding into the function of genes that are specific to malaria parasites [11–13]. Methods for targeted genetic modification have been developed for different *Plasmodium* species [14–20], specifically, *P. falciparum P. knowlesi*, *P. cynomolgi* and the three rodent parasites *P. chabaudi*, *P. yoelii* and *P. berghei*. The availability of efficient reverse genetic technologies for *P. berghei* and *P. yoelii* and the fact that these can be combined with analyses of parasites throughout their complete life cycle, both *in vitro* and *in vivo*, have made these parasites the most frequently used models for gene function analysis [14]. The relevance of such studies for the study of human disease stems from the fact that *Plasmodium* genomes are highly conserved and approximately 80% of the 5300 or so genes in the *Plasmodium* genomes are (probably functionally) orthologous. Targeted disruption (knock-out) or mutation of genes coupled with protein tagging has been extensively used to investigate gene function in rodent parasite species. Tagging of proteins, for example with fluorescent or other epitope tags, has provided insights into expression, localization and transport of parasite proteins and is being used to analyse protein–protein interactions. Reverse genetics has been applied extensively to loss-of-function studies and is now increasingly being used for the generation of so-called reporter parasites. These reporter parasites have heterologous genes introduced into their genome, specifically genes encoding fluorescent or luminescent proteins being driven by a variety of parasite stage-specific or constitutive promoters. Such reporter parasites have been instrumental in the visualization and analysis of parasite–host interactions in real-time *in vitro* and *in vivo* [21–26]. Another more recent development using reverse genetics in rodent parasites has been the generation and analysis of 'attenuated' parasites. These engineered attenuated parasites can become developmentally arrested subsequent to invasion of liver cells [27–29], or attenuation is associated with a marked decrease in the virulence in the host [30–32]. A number of these lines are now being tested and used in research aimed at developing malaria vaccines that consist of attenuated parasites.

The number of different rodent parasite mutants is increasing rapidly, year on year. In order to bring together essential information of all these rodent parasite lines, we

Corresponding author: Janse, C.J. (c.j.janse@lumc.nl).

1471-4922/$ – see front matter © 2010 Elsevier Ltd. All rights reserved. doi:10.1016/j.pt.2010.06.016

have developed a web-based repository (www.pberghei.eu) that contains genotypic and phenotypic data from mutant and reporter parasites.

What is the RMgm Database?

The RMgm (Rodent Malaria–parasites Genetically Modified) database contains genotype and phenotype information of mutant and transgenic reporter parasites and information on gene function inferred from the mutant phenotypes. Most information has been manually curated from scientific literature retrieved from Medline searches. In addition, the database contains unpublished data on the generation of mutants without a clear phenotype and on negative trials to either disrupt or mutate genes; this data has kindly been provided by many different laboratories (see Acknowledgements). The database is continually updated, and as of June 2010 it contained information on over 360 mutants (or attempts to generate mutants). The database can also be accessed via appropriate gene pages in the *Plasmodium* genome resource PlasmoDB (www.plasmodb.org) (Figure 1) and GeneDB (www.genedb.org). In PlasmoDB and GeneDB, the information of a mutant reported in one species is directly linked to the orthologous genes of the other *Plasmodium* species. In the RMgm database there are direct links to PlasmoDB, and Medline abstracts wherever there is a gene and/or paper(s) associated with the mutant or reporter parasite. The database is updated on a weekly basis and collates all the latest data from Medline searches and/or the introduction of unpublished data from different laboratories.

Why have an RMgm database?

The increase in efficiencies and simplified methodologies that permit the generation of rodent mutant parasites [33] has resulted in a rapid increase in the number of publications where gene function has been analyzed by genetic modification. In addition, many different transgenic rodent parasites have been generated that express either tagged proteins or reporter proteins under different regulatory sequences. Currently the standard review of the literature is incapable of providing a sufficiently comprehensive and up-to-date overview of all existing mutants. The RMgm database provides much easier access to this information and is far easier than searching literature databases directly because mutants and their relevant information has already been extracted from publications. In addition searches for the relevant data can be user defined which can be done simply by specifying key features of mutants such as, gene model, type of mutation, life cycle stage in which the phenotype is manifest, reporter proteins,

Figure 1. Direct links between genes in PlasmoDB/GeneDB and RMgm-mutants.
In both PlasmoDB and in GeneDB direct links now exist between *Plasmodium* gene pages and mutants in the RMgm database. For example, clicking on the database link 'phenotypes from genetically modified rodent malaria parasites' in the *P. falciparum* gene page of PFC0640W within PlasmoDB, directs the reader to all records within the RMgm database where the orthogous gene has been modified/deleted (i.e. mutants) in rodent-malaria parasites.

transgenes etc. We believe that the information in the database should provide *Plasmodium* researchers with a useful easy-to-use tool that should assist in the generation of novel hypotheses regarding the roles of genes and their protein products and how they are involved in infection and disease. In addition the database might enhance the exchange of reagents between researchers, specifically of mutants, reporter parasites, constructs and protocols. The hope is that the database also helps to prevent an unnecessary duplication of experiments in different laboratories, because it contains, among other information, unpublished data of unsuccessful trials to disrupt or mutate genes.

What does the RMgm database offer?

The information within the database relating to each mutant parasite line has been subdivided into three sections (Figure 2). The first section provides general information on the generation of the parasite mutant. This information includes links to the relevant Medline publications describing the mutant and information on the researchers and research groups who generated the mutant. If the parasite mutant is available from the Malaria Research and Reference Reagent Resource Center (www.mr4.org), the MR4 number is provided.

The second section provides information on the procedures used to generate and select the mutants, and information on the genotype of the mutants. Four different types of genetic modification have been characterized (Box 1): (i) mutants with disrupted genes (i.e. 'knock-out mutants'); (ii) mutants containing mutated genes; (iii) mutants with tagged genes; and (iv) mutants expressing transgenes. A single mutant parasite might contain multiple different genetic modifications. For each modification, details are provided of the targeted gene, such as gene model which is directly linked to the information on the gene in PlasmoDB and the 'gene product name' as provided by GeneDB. Searching the database for specific mutants can be performed by specifying the gene model or gene product name in combination with specifying the desired type of genetic modification (Figure 3). Certain mutants have been placed in categories which require some further explanation. For example, some *Plasmodium* genes have been targeted by conditional mutagenesis to silence gene expression using the Flp/FRT recombination system of yeast [34], and these genes are classified as mutated genes and not as disrupted genes. The yeast based Flp/FRT site-specific recombination system works by a recombinase (Flipase, Flp) mediated recognition and excision between two identically orientated 34 nucleotide FRT ('Flirted') sites; these sites having previously been

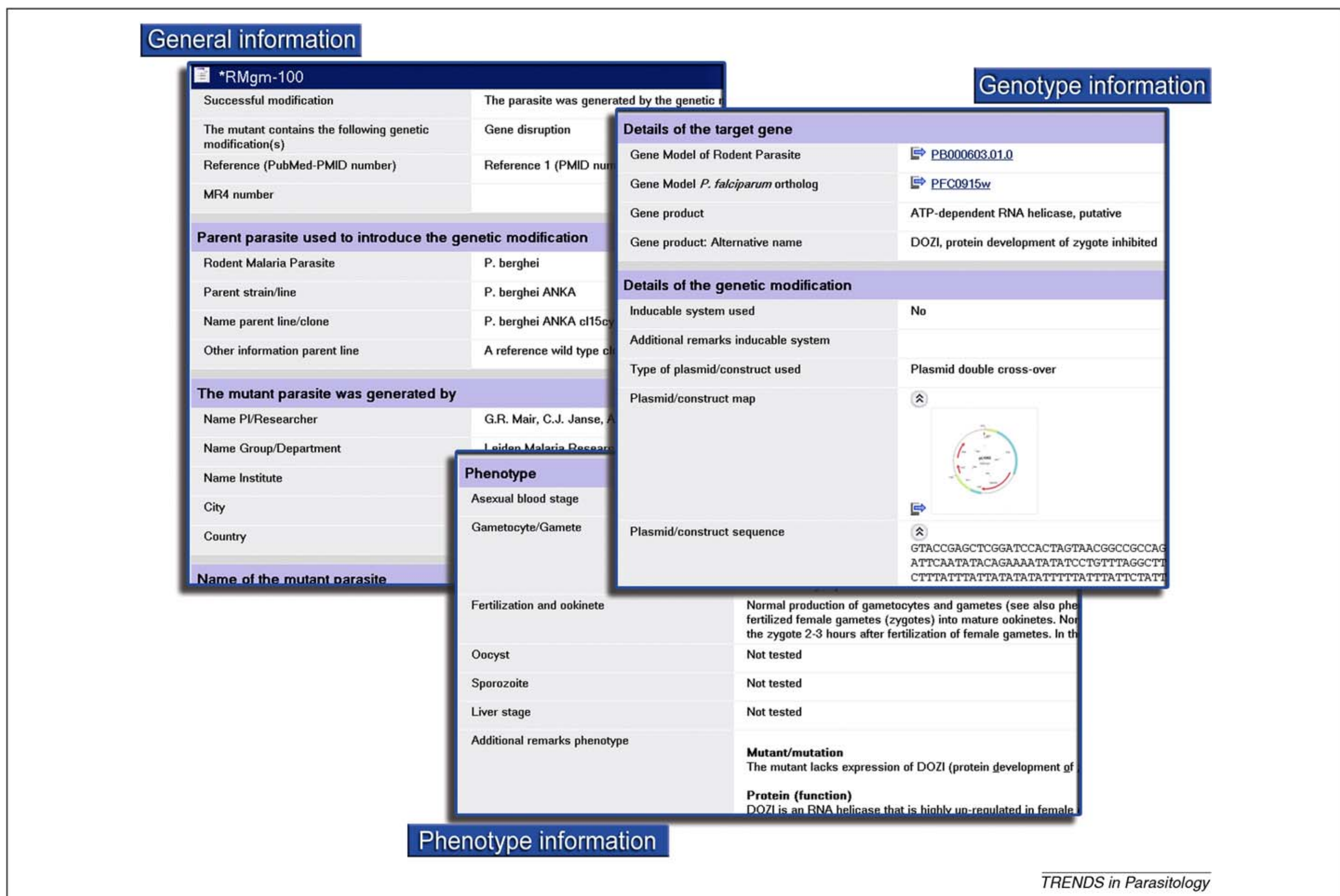

Figure 2. **Genotype and phenotype information in the RMgm database.**
For each mutant the information has been subdivided into three sections, consisting of: **(i)** general information; **(ii)** genotype information, including procedures used to generate and select mutants; and **(iii)** phenotype information, including details about the target gene/protein. Shown is (part of) the mutant information relating to the mutant where the gene encoding DOZI, an ATP dependent RNA helicase has been disrupted.

Box 1. Types of genetic modification and classification of parasite life-cycle stage phenotypes

Genetic Modifications

- ***Disrupted genes (knock-out)***
  Gene function affected by either complete or partial removal the target gene from the genome.
- ***Mutated genes (including conditional knock-out)***
  Gene function altered by replacing the endogenous gene by a mutated form (e.g. a homolog from another species, site directed mutagenesis, etc). This also includes introducing 'FLiRted' genes for conditional knock-out of genes using the Flp/FRT recombination system of yeast [34] and altering the 5'- and 3'-regulatory sequences of target genes.
- ***Tagged genes***
  Modifications of the endogenous *Plasmodium* gene to incorporate heterologous sequences such as c-Myc, TAP, genes encoding fluorescent proteins etc.
- ***Transgenes***
  Heterologous genes, either from other organisms (e.g. GFP, luciferase) or from other *Plasmodium* species, introduced on plasmids or into the genome at a 'silent locus' (disruption of this specific locus does not alter the wild-type phenotype of the parasite). **NOTE**: If genes from other organisms or *Plasmodium* species replace their ortholog in the rodent–parasite genome they are classified as mutated genes.
- ***Mutant parasites***
  Parasites containing any one or more of the following modifications: disrupted, deleted, mutated, transgenes or tagged genes (i.e. any of the above conditions). These include parasites containing modifications that might permit conditional gene knock-outs (e.g. FRT sites).
- ***Reporter parasites***
  Parasites expressing transgenes, such as fluorescent or luminescent proteins, which are often used to investigate parasite–host interactions or in the analysis of parasite growth and development.

Life cycle phenotypes

- ***Asexual blood stage***
  Cell cycle, merozoite egress, red blood cell invasion, growth/multiplication rate, pathology.
- ***Gametocyte/gamete***
  Gametocyte production, male/female gametocyte ratio, formation of gametes, exflagellation.
- ***Fertilization and ookinete***
  Fertilization, zygote formation, meiotic division, ookinete development, ookinete motility, infectivity.
- ***Oocyst***
  Oocyst development, formation of sporozoites inside oocysts.
- ***Sporozoite***
  Development and egress of sporozoites in/from the oocyst or salivary gland; motility, salivary gland invasion, hepatocyte traversal capacity.
- ***Liver stage***
  Sporozoite invasion of liver cells, liver schizont development, parasitophorous vacuole formation, merozoite egress, pre-patent period (time taken for a blood infection to become patent after sporozoite infection) examined in mice.

introduced into a contiguous region of the organism's genome (referred to as the FRTed sequence). In *Plasmodium* these genetic modifications have been achieved by introducing two flanking FRT ('FliRTed') around the gene of interest and gene removal then occurs only after expression of the yeast Flp–recombinase, which excises parasite DNA sequence contained between the FRT sites [34,35]. Also genes that have been replaced with an ortholog from another *Plasmodium* species or from another organism are classified as mutated genes; however, if orthologs are introduced as an additional copy, the orthologs are classified as transgenes. For all mutated genes, a detailed description of the mutation exists and, in addition, a short description of the type of mutation is reported in the overview of search results (Figure 3). The database offers the possibility to search for mutants expressing transgenes and also to specifying the regulatory regions (e.g. promoters, untranslated regions; Figure 3).

The third section of the mutant information provides descriptions of the phenotype of the mutants as well as some additional information on the gene or protein that has been targeted by the genetic modification. As of yet, no guidelines exist for the use of standardized phenotype assays in *Plasmodium* or for standardizing the description of *Plasmodium* phenotypes. Therefore, the assays and phenotypes are described in the database using the same terminology as provided in the corresponding publications. The phenotype description is subdivided according to different parasite life-cycle stages (Box 1). Searching the database for phenotypes or gene functions can be performed by using text term searches in combination with a specific genetic modification (Figure 3).

Importantly, the RMgm database also contains information on unsuccessful attempts to disrupt or mutate genes, and therefore these genes can be refractory to deletion or mutation. For these unsuccessful trials, we provide information on the DNA constructs used in modifying the genes, the selection procedure and the number of independent transfections attempted.

Standardization of mutant generation and describing mutant phenotypes

The full utility of this database for the research community is dependent on the quality of the information available in the database and how this information is presented and how it can be searched. In turn the accuracy of the genotype and phenotype information is principally dependent on the quality of the description of the mutants given in the corresponding scientific report. In *Plasmodium* research no guidelines have been defined for the generation of the mutants or for the analysis and description of their genotypes. Table 1 details a set of considerations, distilled from a number of publications, and what is shown might serve as a standardization template in the generation, analysis and description of mutants of rodent malaria parasites. These criteria are based on generally implemented and now increasingly widely accepted procedures used in a number of different laboratories. It is believed that such standardization might not only improve the quality of the generation and description of mutants but also can enhance the comparability of results from different studies. With the increasing use of genetic modification in *Plasmodium* research and the corresponding rise in genetically modified rodent-malaria parasites, now would seem an appropriate time for the malaria research community to define and accept a set of guidelines for generation and genotyping of mutants. One critical issue in standardization in *Plasmodium* mutant analysis is a definition of

Figure 3. Searching the database for specific transgenic parasites or mutants
(a) Mutants in the RMgm database can be searched for by using a gene model or text term, in combination with a specific genetic modification; for transgenes the regulatory regions can be specified in the search field. **(b)** For mutants containing mutated genes, a short description of the mutation has been provided in the overview of search results.

what constitutes the minimum set of requirements to support an observed phenotype. In truth, these probably differ for the different *Plasmodium* species as each species offers different advantages and limitations in malarial research. It is vital that a phenotype is demonstrably the result of the intended mutation or deletion of the target gene and is not due to unrelated mutations or alterations in the parasite genome. Optimally this would be demonstrated by the restoration of the wild type phenotype through 'genetic complementation' of the mutant by reintroducing a wild type copy of the target gene. This issue is discussed in more detail in a separate letter also in this issue [36]. Whereas complementation experiments are desirable, such experiments in rodent malaria parasites are often difficult to perform because of the limited set of selectable markers available. This is especially apparent when a reported phenotype is based on a set of different mutants or is the consequence of a set of successive transfections performed in a single parasite (e.g. deletions of more than one gene). In this situation an alternative requirement is that a mutant phenotype is confirmed through the analysis of at least a second, independently derived and genetically identical mutant (Table 1). Up to now, most studies have used this strategy and analyzed phenotypes of two independent mutants, and to our knowledge it has not been proven that an incorrect phenotype has been reported when this analytical strategy was applied; however, a few published studies have reported phenotypes based on studies of only a single mutant. When reporting these mutants in the RMgm database or mutants

Table 1. Considerations for generating and describing parasite mutants; suggestions for standardization

	Required information on the generation and/or description of mutants	Additional information
Parasite	Parasite species and strain used for generation of the mutant.	Further information if the mutant is made in a specific cloned or mutant line (e.g. in a reporter background)
Target gene	The PlasmoDB/GeneDB gene model Schematic representation of the gene and information necessary to understand the mutant genotype (e.g. restriction sites, location of probes etc.)	In the case where multiple mutants all containing disrupted genes are generated and analyzed using an identical approach it might be sufficient to provide the gene models in combination with the sequence of the primers used to amplify the target regions for homologous recombination (with a generalized schematic).
DNA construct	Schematic representation of the construct containing information necessary to understand the mutant genotype (e.g. restriction sites; location of probes) Sequence information for all sequences/primers used to disrupt or mutate the target gene Selectable marker cassette	
Mutated gene	Schematic representation containing information to understand the mutant genotype (e.g. restriction sites; location of probes, confirmation PCR primers, size/location of bands/chromosomes on a Southern blot etc.)	
Mutation – procedure for gene disruption	Indication of what type of disruption: Deletion of all or some of the open reading frame (ORF) of the target gene through construct integration by double cross-over (DXO) homologous recombination. Gene deletion by double cross-over integration is strongly preferred over single cross-over (SXO) gene disruption	In case of partial ORF disruption, Northern and/or Western analysis are needed to prove absence or truncated/reduced gene expression. In case of disruption by SXO the possibility exists of reversion to the wild type genotype by recombination and removal of the integrated DNA construct. Genotype and phenotype analyses are required to show an absence of wild type parasites.
Mutation – procedure for gene mutation or gene tagging	Indication that the mutation or tagging is through the integration of a construct by either DXO or SXO homologous recombination.	In case of mutation/tagging by SXO the possibility exists of reversion to the wt genotype by recombination and removal of the integrated DNA construct. In case of a gene mutation, the genotype and phenotype analyses are required to show an absence of wild type parasites.
Selection of mutants	Selection procedure, i.e. drug-selection, flow-sorting (FACS)	
Cloning of mutants	Cloning of selected mutant, e.g. by the method of limiting dilution	If mutants are not cloned, genotype and phenotype analyses need to show the absence of wild type parasites.
Genotype analysis	PCR and Southern analysis of genomic DNA (digested genomic DNA or separated chromosomes) to show correct disruption, mutation or tagging In case of gene disruption: Analyses of transcription and/or protein expression to show absence of protein expression In case of gene tagging: Demonstration of the presence/expression of the tagged-protein	If the mutation introduced into a line is complicated (e.g. use of negative selection to remove drug selectable markers), a genomic southern analysis identifying the correct mutation is necessary.
Phenotype analyses	Confirmation of the mutant phenotype (gene disruption or mutated) must include: (i) analysis of a second independently derived mutant or (ii) complementation of the mutant with a wild type copy of the gene	As of yet no guidelines exist for the use of standardized phenotype assays in *Plasmodium* research. Also no guidelines are currently available that facilitate the uniform reporting of mutants phenotypes or gene function inferred from analyses of mutants (i.e. standardized vocabularies).

where there might be insufficient information supporting the genotype or phenotype, that analysis is indicated on these RMgm entries as possibly incomplete.

As with the analysis of mutant genotypes, no guidelines exist in *Plasmodium* research for the use of standardized assays to analyze mutant phenotypes nor are there initiatives to standardize vocabularies for describing mutant phenotypes. Such initiatives has been initiated in scientific communities that study for example yeast, *Arabidopsis* or mice [37–39]. In the RMgm database, in its current form, the phenotypes and gene functions are provided as 'free text' using the same terminology as used in the

publications. The lack of standardized vocabularies for phenotypes and gene functions limits searches of the database, for example searching for mutants with a comparable phenotype or for genes with a similar cellular location or function. Moreover, it reduces possibilities of integrating information on gene function and protein location from the RMgm database with gene data from other resources and therefore limits the use of this information with larger scale analyses of gene function. The use of Gene Ontology (GO) has been widely adopted as a standard vocabulary for annotating gene function in many organisms [40–45] and is an important tool for structuring information on gene function. In general GO provides information on gene function with a set of defined biological processes, cellular components and molecular functions. These associations are taken from a variety of published sources and the link to this source is given via a defined GO evidence code. GeneDB has adopted GO to annotate *Plasmodium* gene function. It is our intention, in collaboration with the genome curators at GeneDB, to describe functions of *Plasmodium* genes through the use of GO terms, thereby contributing to the ongoing official annotation of *Plasmodium* gene function that is being propagated through GeneDB. These annotations will then be linked to genes in GeneDB using the standard GO evidence code 'inferred from mutant phenotype (IMP)'. Importantly, as *Plasmodium* parasites exhibit biological processes, cellular localizations and gene functions that are specific for malaria parasites (or other Apicomplexa parasites), it might become necessary to develop *Plasmodium (*and Apicomplexa) specific ontologies [42] that better describe the functions of *Plasmodium* genes. Examples of *Plasmodium* (Apicomplexa) specific genes that play a role in malaria parasite-specific biological processes are those that encode proteins involved in sequestration and invasion of red blood cells or in pathogenesis or are important for mosquito infection. The challenges for developing ontologies that define gene functions of pathogenic organisms are clearly different from free-living organisms (e.g. yeast), and to this end several initiatives exist where controlled vocabularies have been developed to describe, among other things, virulence factors and genes involved in pathogenesis [46–53]. One good example of this is the Plant-Associated Microbe Gene Ontology (PAMGO) consortium, which has developed several hundreds of new GO terms and relationships for gene products that help define the interactions between bacterial pathogens and plant hosts [46]. A large number of these GO terms would serve well in describing *Plasmodium* gene function. Using such tractable annotations that are relatable across species might help to further improve gene annotations that are often solely based on phenotype analyses of parasite mutants. In turn this information can then be used to establish 'guilt by association' links with information derived from other approaches such as proteome, microarray and metabolomics, as well as comparative genomics. Only through the use of defined vocabularies to describe gene function and protein location will databases, such as the RMgm database, be able to facilitate both the cross-linking to and mining of these different datasets (expression, metabolic, experimental, etc.), which then would allow for a fully integrative 'systems biology' program in *Plasmodium*. Such approaches will not only help to better understand gene function and gene regulatory networks in *Plasmodium* but also their involvement in infection and disease.

Box 2. Example cases of RMgm database searches

Mutants expressing mutated CSP proteins

Searching for the text term 'circumsporozoite protein' or gene model (PFC0210c; PB001026.00.0) AND 'gene mutated' retrieves >10 mutants that have been generated by different research groups; each mutant expresses a different mutated form of CSP. The different phenotype analyses provide insights into the different functions of (the domains of) CSP in sporozoite maturation, salivary gland invasion and liver infection.

Genes with a role in fertilization and ookinete development

Searching for the phenotype 'fertilization and ookinete' and both 'mutated gene' and 'disrupted gene' retrieves >45 mutants; these encompass >30 proteins, all of which play a role in *Plasmodium* sexual development.

Reporter parasites expressing GFP-Luciferase

Searching for the text term 'GFP Luciferase' AND 'transgene' retrieves >15 mutants expressing GFP-luciferase under the control of different promoters, including reporter lines generated in different *P. berghei* strains (ANKA, NK65, K173), lines that have been used *in vivo* imaging and drug-susceptibility testing.

DNA mismatch repair genes

Searching for the text term 'DNA mismatch repair' retrieves three different 'mismatch repair (MMR)' genes, two of which report 'negative disruption attempts', indicating that these genes might be essential for asexual parasite blood stages. One reports a disrupted gene mutant (i.e. MSH2-2). This mutant shows a wild type phenotype indicating redundancy of gene function or, as of yet, the methods of analysis are too insensitive to precisely examine DNA repair mechanisms in *Plasmodium.*

Mutants expressing tagged proteins

Searching for the text term 'GFP' OR 'c-Myc' AND 'gene tagged' retrieves 21 and 11 mutants expressing GFP- or c-Myc tagged proteins, respectively. The database provides information on the different methods and plasmids used for generating tagged *Plasmodium* proteins. Rapid access to information on mutants expressing tagged proteins might facilitate a more rapid exchange of these lines and should help enhance and standardize research using these tools.

How can you contribute to the RMgm-database?

For all mutants there is a 'comment box' available, where it is possible to submit additional information, corrections and comments as well as suggestions for improving the description of the mutants. Moreover, information on new mutants can be submitted either in excel format or uploaded directly onto the database (more information on how to submit data can be found on the database-website www.pberghei.eu). Researchers are encouraged to submit their information on unpublished mutants or negative attempts to disrupt or mutate genes for inclusion into the database. Such experiments are often not published in regular scientific publications but the knowledge on the existence of mutants without a phenotype or unsuccessful attempts to target *Plasmodium* genes can be of relevance for understanding gene function. For example, although the failure to disrupt or mutate a gene does not in itself prove that the target gene is not amenable to modification, it does provide indirect evidence that the gene might be essential for blood stage development. This is because *Plasmodium* is haploid during blood stage development and therefore disruption of an essential gene

results in parasites that can not be selected as the deletion is lethal. The existence of null-mutants without a distinct phenotype might similarly provide information about the functional redundancy of the target gene. The lack of an observable phenotype might also be the result of assays that are inadequate or, as yet, too insensitive to reveal a phenotypic effect of the genetic modification. Further analysis of such mutants in improved phenotype assays might reveal novel aspects of gene function.

Concluding remarks

The RMgm database provides information on the genotype and phenotype of mutant rodent malaria parasites and also on the function of *Plasmodium* genes inferred from these phenotypes. In Box 2, a few examples are shown of some of the searches and outputs that are possible in the RMgm database. As well as the details on gene function inferred from mutant phenotypes, it is envisaged that the information in the database might benefit *Plasmodium* research by improving the exchange of mutants and materials, such as plasmids between different laboratories. Reporter parasite exchange and the further analysis of existing knock-out mutants (either with or without a described phenotype) might reveal novel insights into gene function. It might be especially true for mutants that express tagged proteins as such mutants can be used to answer multiple research questions, for example, investigations into protein localization, protein transport and protein–protein interactions at different points of the parasite's life-cycle. By facilitating an easy and rapid access to this information it is expected that this resource will not only enhance research into *Plasmodium* gene function but also might contribute to the application and integration of systems biology approaches aimed at understanding biological processes that are unique to *Plasmodium* and play an important role in infection and disease.

Acknowledgements

We would like to thank all researchers who have contributed to RMgm database unpublished information on their genetically modified rodent malaria parasites (see http://www.pberghei.eu/contact.php for a list of contributors). We would also like to thank the following researchers for stimulating and fruitful discussions: David Roos, Omar Harb and Brian Brunk (all at PlasmoDB), and Matthew Berriman and Ulrike Boehme (at the GeneDB/Sanger Institute). Financial support has been received from the European Commission (FP6 BioMalPar Network of Excellence; FP7 Evimalar Network of Excellence) and the Leiden University Medical Center (LUMC, The Netherlands). JF was supported by The Netherlands Organization for Scientific Research (ZonMw TOP grant number 9120_6135).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.pt.2010.06.016.

References

1 Carlton, J.M. *et al.* (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 419, 512–519
2 Carlton, J.M. *et al.* (2008) Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* 455, 757–763
3 Gardner, M.J. *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511
4 Hall, N. *et al.* (2005) A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307, 82–86
5 Carlton, J. *et al.* (2005) The genome of model malaria parasites, and comparative genomics. *Curr. Issues Mol. Biol.* 7, 23–37
6 Dolinski, K. and Botstein, D. (2007) Orthology and functional conservation in eukaryotes. *Annu. Rev. Genet.* 41, 465–507
7 Hall, N. and Carlton, J. (2005) Comparative genomics of malaria parasites. *Curr. Opin. Genet. Dev.* 15, 609–613
8 Kooij, T.W. *et al.* (2005) A *Plasmodium* whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog.* 1, e44
9 Kooij, T.W. *et al.* (2006) *Plasmodium* post-genomics — better the bug you know? *Nature Rev. Microbiol.* 4, 344–357
10 Kuzniar, A. *et al.* (2008) The quest for orthologs: finding the corresponding gene across genomes. *Trends Genet.* 24, 539–551
11 van Dijk, M.R. *et al.* (1995) Stable transfection of malaria parasite blood stages. *Science* 268, 1358–1362
12 vanDijk, M.R. *et al.* (1996) Expression of a *Plasmodium* gene introduced into subtelomeric regions of *Plasmodium berghei* chromosomes. *Science* 271, 662–665
13 Wu, Y. *et al.* (1996) Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. *Proc. Natl. Acad. Sci. U. S. A.* 93, 1130–1134
14 Carvalho, T.G. and Menard, R. (2005) Manipulating the *Plasmodium* genome. *Curr. Issues Mol. Biol.* 7, 39–55
15 Waterkeyn, J.G. *et al.* (1999) Transfection of the human malaria parasite *Plasmodium falciparum*. *Int. J. Parasitol.* 29, 945–955
16 Gardiner, D.L. *et al.* (2003) Malaria transfection and transfection vectors. *Trends Parasitol.* 19, 381–383
17 Balu, B. and Adams, J.H. (2007) Advancements in transfection technologies for *Plasmodium*. *Int. J. Parasitol.* 37, 1–10
18 Menard, R. and Janse, C. (1997) Gene targeting in malaria parasites. *Methods* 13, 148–157
19 van der Wel, A.M. *et al.* (1997) Transfection of the primate malaria parasite *Plasmodium knowlesi* using entirely heterologous constructs. *J. Exp. Med.* 185, 1499–1503
20 Kocken, C.H. *et al.* (1999) *Plasmodium cynomolgi*: transfection of blood-stage parasites using heterologous DNA constructs. *Exp. Parasitol.* 93, 58–60
21 Amino, R. *et al.* (2005) In vivo imaging of malaria parasites–recent advances and future directions. *Curr. Opin. Microbiol.* 8, 407–414
22 Amino, R. *et al.* (2007) Imaging malaria sporozoites in the dermis of the mammalian host. *Nat. Protoc.* 2, 1705–1712
23 Campisi, L. *et al.* (2008) Imaging host–pathogen interactions. *Immunol. Rev.* 221, 188–199
24 Dube, A. *et al.* (2009) Reporter genes facilitating discovery of drugs targeting protozoan parasites. *Trends Parasitol.* 25, 432–439
25 Heussler, V. and Doerig, C. (2006) *In vivo* imaging enters parasitology. *Trends Parasitol.* 22, 192–195
26 Silvie, O. *et al.* (2008) Interactions of the malaria parasite and its mammalian host. *Curr. Opin. Microbiol.* 11, 352–359
27 Mueller, A.K. *et al.* (2005) *Plasmodium* liver stage developmental arrest by depletion of a protein at the parasite–host interface. *Proc. Natl. Acad. Sci. U. S. A.* 102, 3022–3027
28 Mueller, A.K. *et al.* (2005) Genetically modified *Plasmodium* parasites as a protective experimental malaria vaccine. *Nature* 433, 164–167
29 van Dijk, M.R. *et al.* (2005) Genetically attenuated, P36p-deficient malarial sporozoites induce protective immunity and apoptosis of infected liver cells. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12194–12199
30 Aly, A.S. *et al.* (2010) Subpatent infection with Nucleoside Transporter 1-deficient *Plasmodium* blood stage parasites confers sterile protection against lethal malaria in mice. *Cell Microbiol.* 12, 930–938
31 Spaccapelo, R. *et al.* (2010) Plasmepsin 4-deficient *Plasmodium berghei* are virulence attenuated and induce protective immunity against experimental malaria. *Am. J. Pathol.* 176, 205–217
32 Ting, L.M. *et al.* (2008) Attenuated *Plasmodium yoelii* lacking purine nucleoside phosphorylase confer protective immunity. *Nat. Med.* 14, 954–958
33 Janse, C.J. *et al.* (2006) High efficiency transfection of *Plasmodium berghei* facilitates novel selection procedures. *Molecular and Biochemical Parasitology* 145, 60–70

34 Combe, A. *et al.* (2009) Clonal conditional mutagenesis in malaria parasites. *Cell Host. Microbe* 5, 386–396
35 Falae, A. *et al.* (2010) Role of *Plasmodium berghei* cGMP-dependent protein kinase in late liver stage development. *J. Biol. Chem.* 285, 3282–3288
36 Goldberg, D.E. *et al.* Has the time come for us to complement our malaria parasites? *Trends Parasitol.* (in press)
37 Mungall, C.J. *et al.* (2010) Integrating phenotype ontologies across multiple species. *Genome Biol.* 11, R2
38 Costanzo, M.C. *et al.* (2009) New mutant phenotype data curation system in the Saccharomyces Genome Database. Database. (Oxford) 2009, bap001
39 Peng, Z.Y. *et al.* (2009) *Arabidopsis* hormone database: a comprehensive genetic and phenotypic information database for plant hormone research in *Arabidopsis*. *Nucleic Acids Res.* 37, D975–D982
40 The Gene Ontology Consortium (2010) The Gene Ontology in 2010: extensions and refinements. *Nucleic Acids Res.* 38, D331–D335
41 Ashburner, M. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29
42 Harris, M.A. *et al.* (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 32, D258–D261
43 Camon, E. *et al.* (2004) The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res.* 32, D262–D266
44 Rhee, S.Y. *et al.* (2008) Use and misuse of the gene ontology annotations. *Nat. Rev. Genet.* 9, 509–515
45 Aslett, M. and Wood, V. (2006) Gene Ontology annotation status of the fission yeast genome: preliminary coverage approaches 100%. *Yeast* 23, 913–919
46 Torto-Alalibo, T. *et al.* (2009) The Plant-Associated Microbe Gene Ontology (PAMGO) Consortium: community development of new Gene Ontology terms describing biological processes involved in microbe–host interactions. *BMC. Microbiol.* 9 (Suppl. 1), S1
47 Meng, S. *et al.* (2009) Gene Ontology annotation of the rice blast fungus, *Magnaporthe oryzae*. *BMC. Microbiol* 9 (Suppl. 1), S8
48 Meng, S. *et al.* (2009) Common processes in pathogenesis by fungal and oomycete plant pathogens, described with Gene Ontology terms. *BMC. Microbiol.* 9 (Suppl. 1), S7
49 Arnaud, M.B. *et al.* (2009) Gene Ontology and the annotation of pathogen genomes: the case of Candida albicans. *Trends Microbiol.* 17, 295–303
50 McCarthy, F.M. *et al.* (2009) Understanding animal viruses using the Gene Ontology. *Trends Microbiol.* 17, 328–335
51 Giglio, M.G. *et al.* (2009) Applying the Gene Ontology in microbial annotation. *Trends Microbiol.* 17, 262–268
52 Lindeberg, M. and Collmer, A. (2009) Gene Ontology for type III effectors: capturing processes at the host-pathogen interface. *Trends Microbiol.* 17, 304–311
53 Tseng, T.T. *et al.* (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC. Microbiol.* 9 (Suppl. 1), S2