SUPPLEMENTARY INFORMATION

Supplementary Table 1. A current list of links to other online resources provided within MaGnET gene fact sheets.

Resource (home page)	Summary
PlasmoDB (http://plasmodb.org/plasmo/)	*Plasmodium* Genomics Resource. Provides access to publically available genomic, functional genomic and expression datasets and community annotation for all sequenced *Plasmodium* species.
GeneDB (http://www.genedb.org/Homepage/Pfalciparum)	Sequence and annotation information for all *Plasmodium* species sequenced at the Wellcome Trust Sanger Institute.
UCSC Malarial Genome Browser (http://exon.ucsc.edu/cgi-bin/hgGateway?db=pf5)	*P. falciparum* genome browser allowing upload and display of custom data as well as publically available datasets.
ApiCyc (http://apicyc.apidb.org/PLASMO/)	*P. falciparum* metabolic pathways, computationally generated.
Malaria IDC Strain Comparison Database (http://malaria.ucsf.edu/comparison/index.php)	Database of information from the microarray studies published by the DeRisi laboratory.
TDR Targets Database (http://tdrtargets.org/)	Chemogenomics resource for neglected tropical diseases, with information on potential targets and compounds.
Malaria Drug Target Discovery System (http://malport.bi.up.ac.za:8150/)	Database of functional genomics data and predictions for various *Plasmodium* species, as well as the host and vector species.
Malaria Literature Database (http://chemlims.com/MalariaGenePaper/MalariaGenePaperServlet.ChemInfo?module=Overview)	Database of literature references and genes for *Plasmodium* species.
ModBase (http://modbase.compbio.ucsf.edu/)	Database of comparative protein structure models, calculated by a modeling pipeline.
OrthoMCL (http://orthomcl.org)	Database of ortholog groups of sequences from multiple species.
UniProtKB (http://www.uniprot.org/)	Comprehensive protein knowledgebase containing protein sequences and functional information.

Supplementary Table 2. A current list of datasets accessible through MaGnET, their sources, and planned additions/changes in the near future.

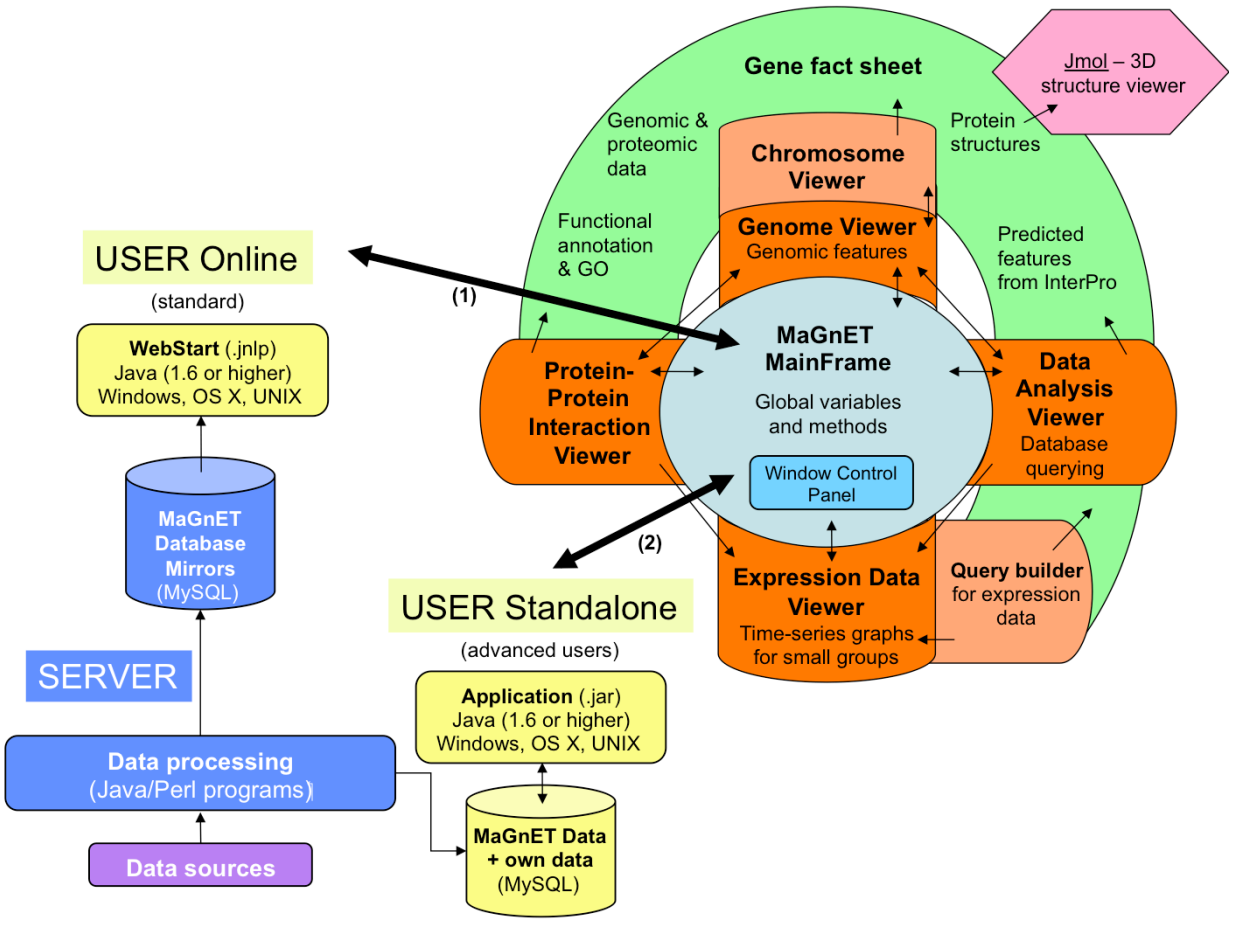
Dataset	**Source**
Chromosome, gene and protein sequences and curated gene annotation	Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/Projects/P_falciparum/) PlasmoDB (http://www.plasmodb.org/plasmo/) Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., et al. 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498-511.
Gene Ontology annotation	PlasmoDB (http://www.plasmodb.org/plasmo/) Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25: 25-29.
InterPro sequence features	PlasmoDB (http://www.plasmodb.org/plasmo/) Hunter, S., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., et al. 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* 37: D211-215.
Ortholog and paralog groupings	PlasmoDB (http://www.plasmodb.org/plasmo/) Li, L., Stoeckert, C.J., Jr., and Roos, D.S. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13: 2178-2189.
Protein-protein interactions	LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., et al. 2005. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438: 103-107.
Experimentally-solved 3D protein structures	RCSB Protein Data Bank (http://www.rcsb.org/pdb) Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. 2000. The Protein Data Bank. *Nucleic Acids Res* 28: 235-242.
Comparatively-modelled 3D protein structures (see below for quality-filtering criteria)	ModBase (http://modbase.compbio.ucsf.edu/)* Pieper, U., Eswar, N., Webb, B.M., Eramian, D., Kelly, L., Barkan, D.T., Carter, H., Mankoo, P., Karchin, R., Marti-Renom, M.A., et al. 2009. MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res* 37: D347-354. *A comprehensive updated comparative model set for MaGnET will be added in the autumn of 2013 through collaboration with the Protein Model Portal (http://www.proteinmodelportal.org).
mRNA time series expression data for several life cycle stages	Llinas, M., Bozdech, Z., Wong, E.D., Adai, A.T., and DeRisi, J.L. 2006. Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Res* 34: 1166-1173. Young, J.A., Fivelman, Q.L., Blair, P.L., de la Vega, P., Le Roch, K.G., Zhou, Y., Carucci, D.J., Baker, D.A., and Winzeler, E.A. 2005. The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol* 143: 67-79. Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., et al. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301: 1503-1508.
RNA-Seq transcriptome data	Otto, T.D., Wilinski, D., Assefa, S., Keane, T.M., Sarry, L.R., Böhme, U., Lemieux, J., Barrell, B., Pain, A., Berriman, M., et al. 2010. New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol Microbiol* 76: 12-24. Bártfai, R., Hoeijmakers, W.A., Salcedo-Amaya, A.M., Smits, A.H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T.W., Françoijs, K.J., Stunnenberg, H.G. 2010. H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog* 6: e1001223.
Protein time series expression data for several life cycle stages	Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., et al. 2004. Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res* **14:** 2308-2318. Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., et al. 2002. A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* 419: 520-526. Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M., Pain, A., Sauerwein, R.W., Eling, W.M., Hall, N., Waters, A.P., Stunnenberg, H.G., et al. 2002. Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* 419: 537-542.

Supplementary Information: **Criteria used for filtering out low quality comparative models downloaded from ModBase.**

In order to help the targeted non-expert usership of MaGnET, we try to apply conservative quality/confidence cut-offs when offering predicted information, and make a distinction in the visual display style between experimental and predicted information (e.g. by colour, or separation into known/predicted sections). For example, we currently supplement the few known protein structures for *P. falciparum* with comparative 3-D models that we extract from the comprehensive sets available through ModBase (Pieper, et al., 2006). ModBase delivers a set ranging from high to low confidence models for each modellable gene, so that experts can select based on their criteria of choice. To avoid confusing non-experts with models that stand very little chance to be accurate, MaGnET only displays a sub-selection of medium-to-high confident ModBase models, pre-filtering based on ModBase quality indices:

Parameter	Condition
E value of match to template sequence	Less than or equal to 10^{-6}
Sequence identity to template sequence	Greater than or equal to 20%
Model length	Greater than or equal to 45 residues
ModBase model score	Greater than or equal to 0.7

See also Suppl. Table 2, above, for a planned transition of our 3-D model source to the multi-method portal.



Supplementary Fig. 1. MaGnET organisational diagram. MaGnET consists of four main viewers: a Genome Viewer for visualising genomic features; a Protein-Protein Interaction Viewer for visualising protein interaction networks; an Expression Data Viewer for exploring time-series data and viewing expression profiles; and a Data Analysis Viewer for querying the database. The Window Control Panel enables one-click navigation between windows. At any point fact sheets listing textual information about individual genes are available. The Jmol viewer is included as an add-on program for visualising 3-D protein structures. MaGnET is available to use "online" via a Java Web Start application that accesses data at our data servers (which are updated synchronously with new releases of its data sources), or "standalone" with the Java program connecting to a local database user-side.