

Supplementary data:

Chromosome-wide analysis of gene function by RNAi in the African trypanosome

Chandra Subramaniam^{1,3*}, Paul Veazey^{2,4*}, Seth Redmond^{2*}, Jamie Hayes-Sinclair¹, Emma Chambers^{1,5}, Mark Carrington⁶, Keith Gull^{1,3}, Keith Matthews^{1,7,+}, David Horn^{4,+} and Mark C. Field ^{2,8,+}

¹School of Biological Sciences, University of Manchester, Oxford Road, Manchester, UK, ²Department of Biochemistry, Imperial College of Science Technology and Medicine, London, UK, ³Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford, UK, ⁴London School of Hygiene and Tropical Medicine, Keppel Street, London, UK, ⁵Department of Biological Sciences, Lancaster University, Lancaster, UK, ⁶Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK, ⁷Institute of Immunology and Infection Research, University of Edinburgh, Edinburgh, UK and ⁸Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, UK

*These authors contributed equally to this work

+ **Correspondence:** MCF (data and website) Tel: +44 1223 333734, email: mcf34@cam.ac.uk, DH (RNAi technology) Tel: +44 20 7927 2352, email: david.horn@lshtm.ac.uk, KM (transfection and phenotyping) Tel: +44 131 651 3639 email: keith.matthews@ed.ac.uk

Legends to supplementary tables

Table S1: Compiled phenotyping data from analysis of *T. brucei* chromosome one and control ORFs by RNAi knockdown. The excel spreadsheet lists all of the ORFs analysed by RNAi together with details of annotation and phenotypic analysis where available. Predicted products are colour coded as defined by Sanger Institute annotation (note that the annotation here may be in advance of that presently displayed at geneDB). Phenotypes are indicated as follows; mild (Yes+, orange), severe (Yes++, red) or none/wild type (No, green). Conserved hypothetical protein 'k' indicates that detectable homologues are limited to within the sequenced kinetoplastid group. TFN1.184 represents a chromosome one sequence annotated as an ORF in an earlier assembly but not in the current assembly where it is noncoding. Asterisk; these genes did not give the expected phenotype in the main analysis but see Table S3 and text for further experimental details. The spreadsheet format facilitates searching and resorting of the data and is available in tab-delimited text format from the authors on request.

Table S2: Frequency of phenotypes amongst genes annotated as conserved hypothetical and with function inferred from homology. Phenotype frequency amongst the conserved hypothetical and function inferred from homology categories, together with subdivision into kinetoplastida-restricted ORFs and widely represented ORFs. These data are presented graphically in Figure 2 and are derived from Table S1.

Table S3: Validation RNAi set. The ORFs selected to validate the various phenotype assays are shown, together with the positive and negative controls used in the analysis. Also shown is the geneDB and trypanoFAN accession, together with the predicted phenotype and annotation of the obtained phenotype.