The RNA Structurome in the Asexual Blood Stages of Malaria Pathogen *Plasmodium falciparum*

Diana Renteria Alvarez[1a], Alejandra Ospina[1a], Abhishek Dey[1], Tiffany Barwell[1], Bo Zheng[1], Chong Li[2], Shrabani Basu[3], Xinghua Shi[2], Sabah Kadri[4]* and Kausik Chakrabarti[1]*

Affiliation:

[1]Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, North Carolina 28223

[2]Department of Computer & Information Sciences, College of Science and Technology Temple University, Philadelphia, PA 19122

[3]Division of Medical genetics, Children's Hospital of Pittsburgh of UPMC, 4401 Penn Avenue Pittsburgh, PA 15224

[4]Division of Heath and Biomedical Informatics, Northwestern University Feinberg School of Medicine and Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, IL 60611

[a] **Contributed equally**

*** Corresponding authors**
E-mail: k.chakrabarti@uncc.edu
Affiliation: Department of Biological Sciences, University of North Carolina at Charlotte, Charlotte, North Carolina 28223

E-mail: sabah.kadri@northwestern.edu
Affiliation: Northwestern University Feinberg School of Medicine and Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago

Abstract

RNA as an effector of biological functions often adopts secondary and tertiary structural folds. *Plasmodium falciparum* is a deadly human pathogen responsible for the devastating disease called malaria. In this study, we measured the differential accumulation of RNA secondary structures in coding and noncoding transcripts from the asexual developmental cycle in *P. falciparum* in human red blood cells. Our comprehensive analysis, combining high-throughput nuclease mapping of RNA structures by duplex RNA-seq, immunoaffinity purification and RNA analysis, collectively measured differentially base-paired RNA regions during the parasite development. Our mapping data not only aligned to a diverse pool of RNAs with known structures but also enabled us to identify new structural RNA regions in the malaria genome. On average, ~71% of the genes with secondary structures are found to be protein coding mRNAs. Mapping pattern of these base-paired RNAs corresponded to all parts of protein-coding mRNAs, including 5' UTR, CDS and 3' UTR. In addition to histone family genes which are known to form secondary structures in their mRNAs, transcripts from genes which are important for transcriptional and post-transcriptional control, such as unique plant-like transcription factor family, *ApiAP2*, DNA/RNA binding protein family, *Alba*, ribosomal proteins and eukaryotic initiation factors involved in translational control and the ones important for RBC invasion and cytoadherence also show strong accumulation of duplex RNA reads in various asexual stages. Intriguingly, our study determined a positive relationship between mRNA structural contents and translation efficiency in *P. falciparum* asexual blood stages, suggesting an essential role of RNA structural changes in malaria gene expression programs.

Introduction

RNA is innately structured. The dynamic intra- and inter-molecular interactions of RNA govern almost all steps of gene expression programs, from transcriptional activation to RNA processing and translation (Sharp 2009). Changes in RNA folding affect biological functions in a cell in a measurable fashion and are facilitated by several factors, such as temperature, ion homeostasis, internal modifications and interactions with proteins. For example, melting of RNA base pairs in response to temperature changes can lead to RNA structural transitions (Nikolova and Al-Hashimi 2010), which can be directly linked to translational regulation as a result of start codon accessibility in mRNAs and ribosomal recruitment for translation initiation process (Narberhaus 2010; Kramer and Gregory 2018; Mustoe et al. 2018). Likewise, drug or metabolite sensing of RNA structures are known to effect mRNA stability or translation efficiency in both prokaryotes and eukaryotes (Barrick and Breaker 2007; Cheah et al. 2007; Henkin 2008; Breaker 2018; Warner et al. 2018). In addition to sequence-specific recognition of RNA by RNA binding proteins (RBPs) (Jones et al. 2001; Sanchez de Groot et al. 2019), chemical modification of RNA nucleotides, such as N6-methyladenosine (m6A) modification (Roundtree et al. 2017) provides an added layer of regulation which can significantly impact gene function (Sanchez de Groot et al. 2019; Corley et al. 2020). Recent developments in high throughput sequencing technologies and their coupling with RNA structure-probing approaches now provide a comprehensive platform to map the secondary structure of the whole transcriptome of yeast, plants, metazoans and mammalian cells and functionally relevant RNA motifs (Kertesz et al. 2010; Li et al. 2012; Spitale et al. 2013; Ding et al. 2015; Kwok et al. 2015; McGinnis et al. 2015; Smola et al. 2016; Smola and Weeks 2018). However, our knowledge of these RNA structural motifs in lower eukaryotes, especially in protists, many of which cause deadly human diseases, are limited due to lack of knowledge on their genome-wide abundance of RNA secondary structures and their influence in gene regulatory processes.

Plasmodium falciparum is a parasitic protist which is responsible for the deadly disease, malaria (Chen et al. 2000). More than 3 billion people live in areas that are at high risk for malaria transmission (Val et al. 2019). Malaria, a febrile illness, is often characterized by cyclical changes in body temperature (Oakley et al. 2011), which can lead to increased adhesive properties of *P. falciparum* infected red blood cells (iRBCs) (Zhang et al. 2018), a critical contributor to malaria pathology. Also, biophysical studies have recently highlighted temperature dependence of receptor-ligand interactions in *P. falciparum* pathogenesis (Lim et al. 2020). Since temperature fluctuations can greatly affect RNA structures and folding energies, which can ultimately influence RNA stability, translation (Wan et al. 2012; Qi and Frishman 2017) and potentially affect virulence properties of pathogenic organisms (Brower-Sinning et al. 2009; Leach and Cowen 2013; Meyer et al. 2017), it is crucial to get a clearer view of how RNA folding can affect gene function in a lethal disease like malaria. Since the sequencing of the malaria genome, numerous studies have shown that transcription in malaria parasite *P. falciparum* is developmentally regulated (Coulson et al. 2004; Le Roch et al. 2004; Llinas and DeRisi 2004; Hughes et al. 2010). Transcriptome wide studies revealed that the *P. falciparum* asexual stages in human RBCs have a cyclic pattern of steady-state mRNA expression, with more than 75% of the genes achieving high abundance of mRNAs at only one time-point of their 48 hours life

cycle (Bozdech et al. 2003; Le Roch et al. 2003; Llinas et al. 2006). Additionally, widespread translational delay was reported in malaria (Bunnik et al. 2013). However, gap remains in our current understanding on how this steady state, timely mRNA expression relates to delay in translation, as transcription and translation of malaria genes appear to be tightly linked processes (Caro et al. 2014). Therefore transcriptome-wide control of gene expression and translation remains an active area of exploration in malaria research. In addition, several studies have indicated that post-transcriptional control mechanisms are major means of gene expression regulation in malaria parasites in both sexual and asexual stages. Global ribosome profiling, mechanistic studies on RNA binding proteins (RBPs) (Mair et al. 2006; Sims et al. 2009; Mair et al. 2010; Otto et al. 2010; Caro et al. 2014; Zhang et al. 2014b; Lu et al. 2017; Rios and Lindner 2019) and differential gene expression studies have shown that a wide array of genes, including virulence genes in malaria parasites, are regulated at both post-transcriptional and translational level (Amulic et al. 2009; Amit-Avraham et al. 2015; Chan et al. 2017). RNA-binding proteins (RBPs) are abundant in *P. falciparum* and could particularly influence post -transcriptional genetic control cascades in malaria (Bunnik et al. 2013; Reddy et al. 2015). A recent survey of m6A modifications, which affect translation by resolving mRNA secondary structures (Liu et al. 2015), suggest that these modifications can dynamically calibrate transcriptional and post-transcriptional processes in *P. falciparum* asexual cycle (Baumgarten et al. 2019). Malaria parasites also express antisense RNAs (Militello et al. 2005; Lopez-Barragan et al. 2011; Siegel et al. 2014) that are considered gene regulatory in nature (Gardiner et al. 2000; Epp et al. 2009; Amit-Avraham et al. 2015) and several other noncoding RNAs (Chakrabarti et al. 2007; Mourier et al. 2008; Vembar et al. 2014; Broadbent et al. 2015), many of which fold into characteristic secondary structures (Chakrabarti et al. 2007) that could account for their diverse functional activities. Therefore, like human, yeast and plants (Wan et al. 2011; Rouskin et al. 2014; Wan et al. 2014; Foley et al. 2015), structural contents of RNA could act as a major force of post-transcriptional and translational regulation in malaria. Indeed, early evidence from differences in structural content of large subunit ribosomal RNAs (rRNAs), between asexual and sexual developmental stages, indicated dynamic changes in these rRNAs related to stage-specific ribosome function in *P. falciparum* (Rogers et al. 1996; Hughes et al. 2010).

Given the plethora of observations on the functional importance of the RNA secondary structures in general and RNA-mediated genetic regulation in malaria parasites, we sought to globally monitor changes in RNA structures in distinct asexual developmental stages of *P. falciparum* by combining nuclease mapping with high throughput sequencing. Additionally, we were able to purify biologically relevant antisense RNA regions from total RNA by immunoprecipitation with antibodies that recognize the double-stranded nature of these RNA molecules. Our RNA mapping data not only aligns with known noncoding RNAs that have defined cellular function, it also provides the first genome-wide view of the structural landscape of messenger RNAs (mRNAs) and novel noncoding transcripts which could play important roles in the biology and pathogenesis of malaria. Thus, our findings provide an extensive collection of RNA secondary structures in *P. falciparum* and a valuable new resource for the analysis of RNA structure-function in this organism.

Results and Discussion

Transcript assembly and analysis of duplex RNA-seq from *P.falciparum* developmental stages

When studying this pathogen, the primary focus for this project was to look into the RNA secondary structures that were formed in the coding and noncoding segments of transcripts at the genome-wide level. Since conventional RNA interference (RNAi) machinery that produce double-stranded RNAs (dsRNAs) (via siRNA and miRNA mediated pathways) for gene silencing (Carthew and Sontheimer 2009) is absent in malaria pathogen *P. falciparum* (Baum et al. 2009), to avoid any confusion we refer to the double-stranded regions identified in our experiments in *P.falciparum* as 'duplex RNAs' or 'RNA duplexes', similar to how they are denoted in previous studies (Piao et al. 2017; Lu et al. 2018). To identify RNA duplexes, i.e. secondary structures that are results of inter- and intra-molecular base pairing interactions by deep sequencing, we optimized a protocol (referred here as 'duplex RNA-seq') that was originally developed for double-stranded RNA Sequencing (dsRNA-seq) in plants (Zheng et al. 2010; Li et al. 2012; Foley et al. 2015) and obtained base-paired RNAs from the asexual blood stage cultures of malaria parasite *P. falciparum*. Synchronized cultures of *P. falciparum* were collected from the 48 hours asexual RBC cycle and labeled as follows: Early or Ring stage collection: 2-12 hours, Mid or Trophozoite stage collection: 18-26 hours and Late or Schizont stage collection: 30-42 hours post-invasion of parasite in human RBCs. Three samples were collected and pooled for each stage during *P. falciparum* asexual blood stage development (**Figure 1A**). Replicates of these synchronized (>90%) parasite cultures were subjected to RNA isolation by TRIZOL (Invitrogen) method. Total RNA isolated was subjected to two rounds of rRNA depletion using custom made oligos against *P. falciparum* rRNAs (see 'Materials and Methods' and **Suppl. Table-1**). These RNA samples were then treated by RNase One (Promega), a single-stranded RNA (ssRNA) -specific endoribonuclease to destroy single-stranded regions of the transcriptome while keeping the base-paired regions intact. Following nuclease treatment, *P. falciparum* duplex RNAs were purified and stage-specific libraries were subjected to RNA sequencing to obtain the first assembly of *P. falciparum* strand-specific transcripts, representing duplex RNA abundance from sequencing data. Two biological replicates of each stage-specific sample were sequenced (see 'Materials and Methods' for details) and the sequenced reads obtained from duplex RNA-seq were aligned to the 3D7 reference genome (PlasmoDB, Release 34) using STAR aligner (Dobin et al. 2013) (see 'Materials and Methods' for details about the analysis). We used two different Illumina sequencing platforms for two separate rounds of duplex RNA-seq, generating the largest number of raw reads of 67,970,256 and 437,153,310 from HiSeq 2500 and NextSeq500 runs respectively (**Suppl. Table 2**). The read lengths varied from 75bp to 125bp. Correlation analysis between the biological replicates of the structural RNA-seq normalized gene counts (REC) (see Materials and Methods) across the three time point collections of *P. falciparum* reveals high correlation of gene counts between the replicates across these two platforms (**Suppl. Fig 1**). It is worth noting that the duplex RNA-seq measured abundance of base-paired RNA regions among developmental stages and thus, may not be quantitated the same way as total RNA-seq data. Despite this, and differences in sequencing instruments, we see a high correlation of gene counts between biological replicates of the same

stage (**Suppl. Fig 1**). Assessing sequencing accuracy, we found on average 83.2% of these total reads showed Phred quality score of Q≥ 30. The distribution of GC content in these reads were between 40-50% which is much higher than average GC content of *P. falciparum* genome (~19%) (**Suppl. Table 2**). This GC-richness of duplex RNA-seq reads suggests that these reads are aligning mostly to the euchromatic coding sequences or functionally important noncoding regions in *P. falciparum* genome which are usually high in GC content (Gardner et al. 2002; Chakrabarti et al. 2007).

Overview of duplex RNA distribution and dynamics in *P. falciparum* RBC cycle

As anticipated, a large proportion of global coverage of base-paired RNA reads aligned to genomic regions which produce structural RNAs of known and unknown functions in *P. falciparum* (Chakrabarti et al. 2007; Mourier et al. 2008; Broadbent et al. 2015). Reads mapping to known tRNAs and residual rRNAs were bioinformatically removed before differential expression analysis of duplex RNAs (see 'Materials and Methods'). Interestingly, a larger proportion of snoRNA, snRNA and known noncoding RNA -specific read accumulations were observed in the trophozoite stage compared to ring and schizont stages of *P. falciparum* (**Figure 1B**). In contrast, ring and schizont stages contained >50% reads that correspond to protein-coding transcripts, representing highly abundant secondary structures within mRNA molecules in these two stages of the *P.falciparum* asexual cycle. On average, the ring and schizont stages have ~23% more "expressed" genes (see expression count calculations in Materials and Methods) containing duplex RNA reads than the trophozoite stage. This data corroborates with stage-specific transcript stabilization that was observed during developmental mRNA dynamics in *P. falciparum* (Painter et al. 2018), probably because RNA stability plays a critical role in mRNA structure driven translational regulation (Mauger et al. 2019).

To better understand *P.falciparum* RNA structurome, we decided to split the coverage by chromosome and determine comparative distribution of the base-paired RNA reads. Thus, we produced a plot of duplex RNA expression from Ring, Trophozoite and Schizont stages on all 14 chromosomes (**Figure 2A**) which revealed that most highly expressed protein-coding genes producing RNA secondary structures in the coding transcripts are residing in chromosomes 6, 11, 13 and 14. Interestingly, many of the antigenic gene families in *P. falciparum* which are involved in host immune evasion and pathogenesis are transcribed from these chromosomes, including multiple copies of *var* and *rif* genes (Gardner et al. 2002). It is noteworthy that the duplex RNA-seq data showed significant enrichment for these mRNAs involved in antigenic variation in *P. falciparum* (see below).

For known noncoding RNAs, such as snoRNAs, snRNAs, ncRNAs, transcripts from chromosomes 11, 12 and 14 showed highest accumulation of duplex RNA reads and thus, indicate an abundance of structured regions (**Suppl Fig. 2**). On the other hand, duplex RNA enrichment on other chromosomes were very minimal. tRNA and rRNA genes were not included in this analysis. As **Suppl. Fig 2** shows, this observed trend is not simply a result a higher number of non-coding genes on these chromosomes (black dots). The size of non-coding genes also follows a similar pattern as the number of genes, with chromosomes 8 and 11 containing

the most area with non-coding genes (~10kB) and chromosome 14, which has the most enrichment for duplex RNA in non-coding regions covers ~6kB (data not shown).

Because the duplex RNA-seq data retained strand information of the molecules, we were able to determine the strand bias for different classes of RNAs. DNA replication and transcription can exert a strand bias on the DNA repair process (Liu et al. 2002; Brachman and Kmiec 2004), particularly transcription levels affecting the degree of bias (Strick and Savery 2017). This in turn could affect an organism's gene expression program. In *P. falciparum*, strong enrichment of duplex RNA reads transcribed from the forward (plus) strands were observed for mRNAs and snoRNAs from all three stages. Particularly, mRNAs and snoRNAs were ~1.5 fold and ~5 fold enriched respectively for the forward strand compared to reverse (minus) strand across all chromosomes. Expression of mRNAs with secondary structures were higher in ring and schizont stages compared to trophozoite stages, whereas snoRNAs were mostly expressed in the trophozoite stage, irrespective of strand bias (**Figure 2B**). On the contrary, snRNAs and other known noncoding RNAs showed a strong reverse strand bias in the duplex RNA-seq data.

Genomic distribution of duplex RNA hotspots in *P. falciparum.*

The evidence of highly base-paired regions of RNAs, referred hereby as duplex RNA 'hotspots' were identified across the coding and noncoding regions of *P. falciparum* genome. From previous studies, it is known that structured regions in the mRNA have the potential to act as regulators in cellular processes, such as RNA stability, turnover and translation in yeast and mammalian cells (Jacobs et al. 2012; Leppek et al. 2018; Mauger et al. 2019). Our duplex RNA-seq data mapped to a total of 3829 protein coding (mRNAs) and non-coding transcripts with relative expression counts ('REC') above set thresholds (REC>10) and showed high enrichment (>100 REC) for 969 mRNAs in two or more samples across any of the stages (See Materials and Methods). This suggests that these 3829 coding and noncoding regions could potentially fold into unique, stable secondary structures. A threshold of 10 was set based on empirical observations of peak data being unreliable below this threshold (see Materials and Methods for additional information). In addition, we identified known structural RNAs and novel, structured RNA molecules throughout the intronic and intergenic regions of *P. falciparum* genome.

Structural RNAs

Our duplex RNA-seq data is in agreement with known RNA structures in *P. falciparum.* We corroborated our duplex RNA sequencing data with 73 highly base-paired segments of known, conserved structural RNAs (**Suppl. Table 3**). To characterize the constituents of duplex RNA-seq reads, we determined the secondary structure of the *P. falciparum* U3 snoRNA *in vitro* by in-gel RNA-SHAPE analysis using NAI (**Figure 3A**) and mapped our duplex RNA-seq to the U3 secondary structure model. As anticipated, base-paired stem regions originating from intra-molecular interactions in PfU3 RNA corresponded to more duplex RNA-seq reads (represented by peaks in **Figure 3B**) than the unpaired regions (**Figure 3C**). Thus, our analysis revealed with confidence that the majority of our data obtained from duplex RNA-seq are related to biologically relevant structures in the malaria genome. Among all these noncoding RNAs, the highest enrichment of reads (> 333,000 normalized read counts) mapped to the *P. falciparum*

Signal Recognition Particle RNA (SRP RNA, PF3D7_1418800) (Chakrabarti et al. 2007), which is a conserved RNA universally required for co-translational protein targeting. The three-dimensional structure of mammalian SRP RNA is already known (Halic et al. 2004) and the secondary structure prediction of *P. falciparum* SRP RNA was previously reported based on comparative genomics and RNA analysis (Chakrabarti et al. 2007), which corroborates strongly with the duplex RNA-seq data (**Figure 3D and 3F**). When stage-specific PfSRP expression was compared to the base-paired transcript abundance in our duplex RNA-seq data (**Figure 3E and 3F**), we observed lower accumulation of SRP RNA domain-specific reads in the ring stage compared to the schizont stage, although expression levels were similar in early ring and schizonts stages as evident from northern analysis. (read accumulations are represented by 'peaks' in **Figure 3F** and the missing peak in the ring stage is shown by an arrow in **Figure 3F** with it's corresponding origin in **Figure 3D)**, This indicated that our duplex RNA-seq was able to capture stage specific RNA structural rearrangements during *P. falciparum* development. Among other highly structured classes of RNA molecules that were identified in *P. falciparum* genome, secondary structure and RNA modification sites for *P. falciparum* C/D box snoRNA 15 (PF3D7_1222200) was previously established by *in vivo* DMS probing (Chakrabarti et al. 2007). Our duplex RNA-seq accurately mapped to those structured regions of *P. falciparum* snoR15 (**Suppl. Figure 3**), suggesting that we are able to reliably capture intramolecular base-pairing interactions by duplex RNA seq and afforded the first genome-wide view of the structural landscape of *P. falciparum* RNAs, many of which are involved in conserved biological functions. In addition, RNAs of Unknown Function (RUFs) (Chakrabarti et al. 2007), RUF1 (PF3D7_1328700) and RUF 6 (PF3D7_0711800, PF3D7_0712100, PF3D7_0712700), all expressed from chromosome 7, showed enrichment of base-paired RNAs.

Coding Sequence (CDS) of mRNA

Our duplex RNA-seq readout allowed global analysis of RNA secondary structures in the coding regions of *P. falciparum* mRNAs. Recent reports suggest that highly expressed mRNAs usually have extensively structured coding sequence (CDS) (Wan et al. 2014; Mauger et al. 2019), which could positively correlate with transcript-specific translation efficiency by increasing the mRNA half-life (Mauger et al. 2019) and the ribosome (Beaudoin et al. 2018) or m6A modifications (Mao et al. 2019) in CDS or UTRs could be the major remodeler of the mRNA structures for translation. In malaria parasites, mRNA translation facilitates developmental stage transitions (Cui et al. 2015; Painter et al. 2018), suggesting that malaria disease progression is dependent on stage-specific protein output. Additionally, *in vivo* measurements of mRNA cellular dynamics by biosynthetic labeling recaptured the 'just-in-time' transcriptional program of *P. falciparum* mRNAs with variable degrees of mRNA stabilization profile (Painter et al. 2018), which could be due to differential presence of structure between these mRNA populations. Therefore, it is plausible that a great proportion of post-transcriptional gene regulatory program is driven by mRNA secondary structures in this parasite. We found that nearly 71.4% of the protein coding genes in *P. falciparum* aligns to duplex RNA-seq reads exceeding our set threshold (REC>10), and as a result, define its secondary structure content. Out of the uniquely "expressed" 3773 mRNA regions containing secondary structures, approximately a quarter (26.1%) show high cumulative expression (>100 REC) of transcripts with structured regions in at least 1 stage. Based on the average expression seen across all the dsRNA-Seq time points,

we conclude that approximately 71% (3773) of all mRNAs make secondary structures in at least one of the sequenced stages, with expression above set thresholds. When comparing the distribution of expressed genes (REC) between the stages, we found significant difference among Ring and Trophozoite or Trophozoite and Schizont stages (**Suppl. Figure 4**) (two-sided Wilcoxon test, $p < 0.001$). However, no significant difference was observed between the distributions of the expressed genes of Ring and Schizont stages in the *P. falciparum* RBC cycle.

In order to study the correlation between the gene expression of protein coding transcripts in *P. falciparum* total RNA-seq (Otto et al. 2010) and the duplex RNA-seq data, we normalized each dataset separately using DESeq2 (Love et al. 2014) and plotted the average normalized gene expression counts for each stage across the total RNA-seq and duplex RNA-seq datasets. As expected, the protein coding genes have more expression in the mRNA-seq dataset compared to the duplex RNA-seq as the latter represents a subset of the former, containing expressed structure contents of the protein coding genes (**Figure 4A, Suppl. Figure 5A** and **5B**). The overall degree of correlation between the two datasets is quite high (R^2 values are 0.66, 0.65, 0.70 across the three stages). Few genes show higher expression in duplex RNA-seq compared to mRNA-seq and few of these are highlighted in the figures.

To understand the mRNA structure dynamics during *P. falciparum* asexual stage transitions in human RBCs, we characterized differential expression of structured contents within mRNA coding sequences (CDS), which is depicted as a heatmap in **Figure 4B**. Results from the top 100 most expressed mRNAs with secondary structures (out of all the differentially expressed ones) during the 48 hours intra-erythrocytic developmental cycle (IDC) of this parasite suggest that highly folded RNA coding regions can potentially influence stability and translational outputs of individual transcripts in malaria gene expression programs. Among the highly expressed base-paired mRNA molecules in *P.falciparum*, histone mRNAs (Histone H2A (PF3D7_0617800), H2B (PF3D7_1105100) and H3 (PF3D7_0610400)) showed highest accumulation (within top 1%) of reads by duplex RNA-seq. Histone mRNAs are one of the most extensively validated functional mRNA structures present not only in metazoans, but also in many different protozoa, including alveolates (e.g. *P. falciparum)* (Davila Lopez and Samuelsson 2008). The histone mRNA forms a stable stem-loop structure at the 3′ UTR structure, which is involved in nucleocytoplasmic transport and regulation of translation efficiency (Zhang et al. 2014a). In addition to strong enrichments of duplex RNA reads at the 3' UTR of *P. falciparum* histone mRNAs, read abundance was also observed in the CDS, as shown for histone H2A (**Figure 4C**). In addition, the list of highly expressed duplex RNA containing transcripts in *P. falciparum* included mRNAs of chaperon proteins *HSP70* and *HSP90*, which are also known to be highly structured (Ahmed and Duncan 2004; Sanchez de Groot et al. 2019). These findings substantiated that our duplex RNA-seq interrogated the desired regions of mRNAs on a genome-wide scale, which are highly enriched in base-pairing interactions.

One emerging theme about *P. falciparum* gene regulation is the central role played by DNA and RNA binding proteins in transcriptional and post-transcriptional control of gene expression

(Vembar et al. 2016; Painter et al. 2018; Rios and Lindner 2019). For example, a plant-like transcription factor family consisting of 27-members, known as Apicomplexan AP2 (*ApiAP2*) protein family, expressed in different stages during malaria parasite development, is critically involved in parasite development and differentiation processes (Balaji et al. 2005; Painter et al. 2018). Expression of two members of this gene family, PF3D7_0420300 and PF3D7_1143100 (**Figure 4B**), were among the top 100 gene candidates that produced high levels of differentially expressed duplex RNAs in their coding regions and thus considered in our study as prominent 'hotspot' candidates. Among those, *ApiAP2* (PF3D7_0420300) expressing protein is known to functionally interact with ACACACAT DNA motif for *P. falciparum* gene expression regulation (Toenhake et al. 2018) and its expression is correlated with drug resistance transporter, *pfCRT*, indicating potential role in resistance evolution to chemotherapeutic molecules in malaria (Siwo et al. 2015). Similarly, for RNA mediated post transcriptional gene regulation, RNA binding proteins (RBPs) in *P. falciparum* plays pivotal role in translational regulation (Bunnik et al. 2016; Rios and Lindner 2019). Our duplex RNA-seq data revealed strong accumulation of base-paired reads mapping to DNA/RNA binding factor *PfAlba3* (PF3D7_1006200) (**Figure 4C**). *PfAlba3* protein interacts with *P. falciparum* histone deacetylase, *Sir2A* (Goyal et al. 2012), associates with polysome fractions during the *P. falciparum* IDC (Bunnik et al. 2016) and therefore, may play important regulatory roles in transcription as well as in translation. In addition to mRNAs that produce regulatory proteins, transcripts encoding antigens and adhesion molecules involved in host-parasite interaction in *P. falciparum* are the most highly base-paired classes of RNA molecules, including Erythrocyte binding antigen 175 (*EBA-175*) (PF3D7_0731500), which is involved in erythrocyte invasion (Mamillapalli et al. 2006), *SERA5* (PF3D7_0207600), important for parasite egress (Collins et al. 2017) and ring-infected erythrocyte surface antigen, *RESA-1* (PF3D7_0102200) and *RESA-3* (PF3D7_1149200) (Favaloro et al. 1986), which may prevent *P. falciparum* infected erythrocytes from reinfection through interactions with spectrin family proteins (Pei et al. 2007).

Genetic control of cellular metabolism often involves sensing of mRNA secondary structures by nutrients and metabolites or interactions with RBPs (Sudarsan et al. 2003; Clingman and Ryder 2013). These mRNA structures are widespread in prokaryotes and also identified in handful of eukaryotes. In prokaryotes, these are referred to as RNA genetic switches which are typically located in noncoding regions of mRNAs (such as UTR regions) of many metabolic enzymes and can modulate gene expression. In eukaryotes, metabolic enzymes can interact with mRNAs to affect mRNA stability or translation efficiency (Clingman and Ryder 2013). We have identified base-paired mRNA elements in the coding region of *P. falciparum* metabolic enzymes, with most enrichment of duplex RNAs in the Hypoxanthine-guanine phosphoribosyltransferase (*HGPRT*) (PF3D7_1012400) and Adenosylhomocysteinase (*SAHH*) (PF3D7_0520900). Since most parasitic protozoa do not have the *de novo* purine nucleotide biosynthesis pathway and rely mostly on the salvage pathway for purine metabolism where hypoxanthine is the key precursor, *HGPRT* in this pathway makes a promising drug target for *P. falciparum* infection (Cheviet et al. 2019). However, the cellular interactions and importance of these novel mRNA structural elements in the coding regions of metabolic enzymes remain to be analyzed.

The importance of sequence and structural motifs in mRNAs are vastly appreciated in studying post-transcriptional regulatory processes (Rabani et al. 2008; Dominguez et al. 2018). In order to determine if there were any sequence motifs located within the structured regions of mRNAs across highly expressed pool of transcripts, we obtained sequences from 50 highly expressed mRNAs representing a defined structured region (structured regions represented by peaks shown in **Figure 4C**). We then used the pattern search program, Improbizer (Ao et al. 2004) that searched through duplex RNA read sequences and looked for reoccurring short sequence motifs and found six motifs which could be potential regulatory region or an RBP interaction site on mRNA structures. The highest scoring motif sequence is shown as sequence logo in **Figure 4D**. In summary, the overall abundance (**Figure 1B, 2B**) and dynamic changes (**Figures 4A, 4B**) in mRNA structures during *P. falciparum* blood stage development indicate that the extensive stage-specific RNA folding events can be potentially linked to fine-tuning of protein synthesis in malaria gene expression program.

Noncoding regions of mRNAs - 5' and 3' UTRs

To investigate the global changes in *P. falciparum* mRNA structure during asexual blood stage developments, especially in the untranslated regions (UTRs), we analyzed top 20% differentially structured regions over development in the UTR regions. Since the UTR lengths have a very broad range in the *P. falciparum* genome, the UTRs were binned into 50 and 500 nucleotides (5' and 3' UTRs: defined by (Chappell et al. 2020); CDS: 100 nt downstream of the start codon (for 5'UTR) and 100 nt upstream of the stop codon (for 3'UTR), calculated average expression from ring, trophozoite and schizont stages (see 'Materials and Methods' for details about the analysis). **Figure 5A** and **5B** show the distribution of normalized expression ratios of the duplex RNA-seq reads at the 5' and 3' gene ends for the top 20% genes with structures in the UTRs respectively. The ratio is calculated against the maximum expression of the structures in the region being investigated in order to study the relative enrichment (See Methods). For example, the Schizont structures at the UTRs and splice junctions show lower ratios and thus, lower relative enrichment than the Ring and Schizont stages (**Figure 5A-C**) (t test, $p < 0.001$). Having secondary structure near areas of translation initiation sites or coding regions have been seen to play a role in regulating translation efficiency. For non-protein coding regions, i. e. 5' UTRs and 3' UTRs of mRNAs molecules, there is a significant enrichment in longer UTRs in the Ring and Schizont stages compared to the Trophozite stage (**Suppl. Fig 6**) (Wilcoxon test p-value <0.01).

Antisense RNAs

Antisense RNA transcripts, originated from the opposite strand to a protein or RNA-coding strand, have been ascribed roles in gene expression regulation by degradation of the corresponding sense transcripts or by silencing genes at the chromatin level (Faghihi and Wahlestedt 2009). Therefore, pervasive antisense transcription in *P. falciparum* fueled the speculation that these RNAs regulate malaria gene expression programs either at the epigenetic or post-transcriptional level (Gunasekera et al. 2004; Lopez-Barragan et al. 2011; Jiang et al. 2013; Siegel et al. 2014). At the transcriptome level, widespread antisense transcription in *P. falciparum* corresponded to transcription initiation in an antisense orientation

to coding genes(Adjalley et al. 2016), transcription start site-associated RNAs (TSS-RNAs) expression (Chappell et al. 2020) as well as antisense transcription in the context of pre-mRNA splicing events(Sorber et al. 2011). At the gene level, natural antisense transcripts (NATs), longer than 200 nt and acting as antisense long RNAs can potentially regulate *P. falciparum* virulence gene expression (Epp et al. 2009; Amit-Avraham et al. 2015), however, the origin and the mechanism of action of these long antisense RNAs transcribed from either 'silent' or 'active' var genes, remains unclear (Ralph et al. 2005). To investigate if our duplex RNA reads constitute potential antisense transcripts in *P.falciparum* genome, we counted the antisense reads for all coding and non-coding transcripts and normalized these counts by depth of sequencing (REC). The same threshold described earlier (REC>10.0 in at least two samples across any of the three RBC stages of *P. falciparum*) was applied to the antisense data as this was the empirically set threshold for more reliable duplex RNAs. Since antisense reads are in low abundance (after applying the above stringent cutoff), we found a total of 159 genes (154 mRNAs and 5 non-coding RNAs) with antisense reads above set thresholds. To better understand the relationship between gene expression and the abundance of duplex antisense reads, we looked at the average expression of the top 50 mRNA genes with antisense expression (**Figure 6A)**. As shown in the **Figure 6A**, the gene expression from mRNA-seq and duplex antisense expression are not correlated but show an approximately negative relationship (Pearson correlation R^2 = -0.11). This means that for certain genes, such as *PF3D7_1425200, PF3D7_0112800* etc, the lower mRNA expression might be explained by antisense transcriptional regulation, although this remains to be validated by functional assays. We saw a few interesting types of antisense expression, especially in the 3' UTR regions (36.5% antisense structures were found in the 3'UTR). **Figure 6B** represents the three types of antisense peaks seen in the 3' UTRs (i) antisense peak without any overlapping genes (ii) antisense peak with overlapping UTR of another protein coding gene, and (iii) antisense peak from overlapping non-coding gene(s). An example of (iii) can be seen in **Figure 6C**. To validate antisense RNA expression, we immunoprecipitated rRNA depleted *P. falciparum* total RNA with dsRNA specific J2 monoclonal antibody (SCICONS) and probed this RNA with sense and antisense RNA specific complementary oligonucleotides. In this process we confirmed expression of ten (10) sense/antisense RNA pairs from asynchronous, mix stages of RBC cycle. Representative northern blots validating presence of antisense RNAs opposite three (3) genes are shown in **Figure 6D.** Although our deep-sequencing data and Northern blots correspond to the overlapping regions of the RNAs, it remains unclear whether mRNA stability was affected by the expression of these antisense RNAs or their effect in resolving secondary structures of translatable mRNAs in *P. falciparum*.

Genome-wide relationship between RNA secondary structures and biological processes in *P. falciparum*

Next we investigated genome-wide relationships between mRNA structures and biological functions of the encoded proteins. For this purpose, we used differential expression data of secondary structures from the three *P. falciparum* developmental stages within RBCs, i.e. ring, trophozoite and schizont stages (**Figure 4B**). Differential expression of structured mRNA can be measured through the read count in the different stages where a higher read count in one

stage versus the other is representative of stage -specific differential accumulation. A GO enrichment analysis of the differentially structured transcripts shows that the genes are enriched for molecular functions that are highly relevant for the developmental stages. These processes include host cell surface binding ($Q < 1 \times 10^{-3}$), DNA/RNA binding, which includes mRNA binding (GO:0003729) and sequence-specific DNA binding (GO:0043565) and multiple GO terms that belong to processes involved in enzyme activity. Among these genes, are transcription factors like Plant-like transcription factor *ApiAP2* and *ALBA3* (PF3D7_1006200) ApiAP2 protein *SIP2* (0604100) ($P<0.05$ Ring vs Trophozoite stage). In addition to most significantly enriched molecular functions corresponding to proteins involved transcription and RNA processing, our duplex RNA-seq data showed GO enrichment for processes, such as red blood cell invasion and host immunity (for example, *EBA-181* and EBA-175, $P<0.05$, Ring vs Trophozoite stage). Thus, systematic analysis of mRNA secondary structures in gene expression and molecular function in *P. falciparum* identified candidates that are substantially enriched for functions relevant to malaria biology and pathogenesis.

Analysis of mRNA structure remodeling and translational control

Dynamics of mRNA unwinding can significantly affect the rate of translation in eukaryotic systems. Previous studies have suggested that stable secondary structures in the 5'UTR or coding regions (CDS) of mRNAs can lead to ribosomal pausing our translational arrests (Mao et al. 2014; Rodnina 2016). Given the critical role mRNA structures can play in translation elongation, we next sought to determine the correlation between structural contents of coding regions and translation efficiency in *P. falciparum*. Using existing ribosome profiling data sets (Caro et al. 2014), we used the ribosomal footprint RPKM as a representation of translation efficiency in *P. falciparum* and compared it to the duplex RNA-seq structure levels. While looking at the top 50% structure- containing transcripts in the *P.falciparum* genome, our developmental stage-specific data, represented by the expression of the genes in the duplex RNA-seq datasets and the translational efficiency of these genes revealed an overall positive correlation for most of the mRNA coding regions ($R = 0.49, 0.51, 0.53$ for Ring, Trophozoite and Schizont $p<0.001$) (**Figure 7**). However, how individual transcript-specific differences in RNA structures relate to translation efficiency still needs to be determined for *P. falciparum* mRNA translation.

Conclusions

Our analysis of *P. falciparum* RNA structurome elucidated stage-specific RNA structural changes and therefore has a potential to yield significant insights into structure–function relationships in post-transcriptional control. Our comparative analysis with other eukaryotic RNA structuromes revealed that *P. falciparum* contains substantially high secondary structures (~71.4%) in protein coding regions compared to human, yeast and metazoan species. Given that *P. falciparum* genome has the highest AT content (~80%) of any organism sequenced so far, it is striking that *P. falciparum* transcriptome contains substantial stable, dynamic secondary structures compared to other organisms with higher GC content. Three membrane layers of

host and the parasite have posed a challenge for in-cell probing of malaria structure by high-throughput methods. By overcoming this using a robust *in vitro* assay, we show that reliable detection of structural contents in malaria transcriptome is feasible by duplex RNA-seq. Many of these structural regions in coding sequences are phylogenetically conserved at the sequence level, providing an opportunity to study RNA-structure-dependent functions in malaria animal models. Detailed interrogation of how these structures in malaria coding regions relate to function should provide mechanistic details for developing future therapeutics affecting mRNA structure and stability differently, depending on how they inhibit malaria translation.

Materials and Methods

Plasmodium falciparum Culture, Synchronization and Treatments

P. falciparum laboratory adapted parasite line 3D7 was cultured at 2% hematocrit in purified human erythrocytes/RBCs using standard culturing techniques (Trager and Jensen 1976; Duffy and Avery 2018) using RPMI 1640 medium (supplemented with 37.5 mM HEPES, 7 mM D-glucose, 6 mM NaOH, 25 µg of gentamicin sulfate/ml, 2 mM L-glutamine, and 10% Albumax II (Life Technologies) at a pH of 7.2 in a 3% carbon dioxide environment. The asexual blood stage parasite density was calculated routinely using thick blood films. Baseline RBC count was used to calculate the parasitaemia (parasites/µl). P.falciparum Infected Red Blood Cells (iRBCs) were treated with a pre-warm aliquot of 5% D-sorbitol in RPMI at 37°C to synchronize parasite cultures in order to isolate homogeneous stages of parasites. Multiple sorbitol treatments were made to achieve >90% synchrony of parasites. Parasites were collected from three different developmental stages, Ring, Trophozoite and Schizont, after stringent synchronizations.

RNA extraction, rRNA depletion and nuclease mapping

P. falciparum parasites were isolated after saponin lysis, washed with incomplete medium (without Albumax) and RNA from each stage were isolated by Trizol (Thermofisher) method followed by Turbo DNAse (Thermofisher) treatment. Total RNA was then subjected to two rounds of rRNA depletion with custom biotinylated oligos designed against *P.falciparum* rRNAs (**Suppl. Table 1**). Briefly, 500 pmol of biotinylated capture probes were used in binding buffer containing 10 mM Tris-Cl pH 7.4, 3 mM MgCl2, 300 mM NaCl and 0.5% Igepal-CA 630. Hybridization reaction was prepared on ice in presence of 50 ug of total RNA in each tube and 2ul of RNase inhibitor. Oligo probe and total RNA mixture were heated at 70 °C for 5 min followed by gradually reducing the heat until the oligonucleotides have reached room temperature. Finally, rRNA bound oligos were separated by Streptavidin conjugated magnetic Dynabeads M-280 (Invitrogen) followed by ethanol precipitation of rRNA depleted total RNAs. Next, RNA samples were treated with a single-strand specific ribonuclease, RNase One (Promega) (5 units per ug of RNA) according to manufacturer's instructions. After nuclease treatment, each RNA sample was fragmented with Fragmentation Reagents (Ambion) and then end corrected with T4PNK (Thermofisher), phenol-chloroform extracted and ethanol precipitated before using them for RNA-seq library preparation.

Library preparation and high-throughput sequencing of duplex RNAs (Duplex RNA-seq)

P. falciparum duplex RNA sequencing libraries were prepared from three developmental stages: Ring, Trophozoite and Schizont (see results section for details). Two independent sets of libraries were prepared (2 replicates for each developmental stage). Sequencing library was constructed using the Small RNA Sequencing Kit v3 (Illumina, San Diego, CA) as per manufacturer's instructions. Library quality was checked by Bioanalyzer and then pair-end sequencing was done on two different Illumina high-throughput sequencing platforms, HiSeq2500 for RNA-seq-1 (Otogenetics Corp., Atlanta, GA) and NextSeq500 for RNA-seq-2 (ACGT Inc, Germantown, MD) (**Table-1**).

Bioinformatics analysis of duplex RNA-seq data

The paired end reads from the duplex RNA-Seq data were first trimmed to remove any adapter sequences using fastp v0.20.0 (Chen et al. 2018). STAR v2.5.2 (Dobin et al. 2013) was used to align the trimmed data to the *P. falciparum genome* (v34) with the parameters --genomeSAindexNbases 11 and --alignIntronMax 500. The maximum splice junction length was set to a conservative 500bp as a larger value of this parameter yielded many false splice junctions and thus, noisy structural peaks. This is mainly due to the short insert sizes in the libraries (**Suppl. Table 2**) and thus, relatively shorter length of the trimmed reads post adapter trimming. A custom filter was then used to remove low mapping quality (mapping quality =0) and secondary or supplementary alignments. Finally, featurecounts v2.0.0 (Liao et al. 2014) was used to perform gene counting using the latest gene annotations from (Chappell et al. 2020). Reads that mapped to mitochondrial or apicoplast transcripts were removed from this gene expression analysis. Normalization of gene counts (as seen in Fig. 4) and differential gene expression was performed using DESeq2 v1.26.0 in R v3.6.2. Expression count calculations to estimate the expressed genes are described below.

Expression count calculations to flag potential stable secondary structures

The raw counts from 'featurecounts' including all known coding and non-coding genes were normalized by the relative sequencing depth based on the number of mapped reads to calculate the relative expression counts (REC). Only genes with 'REC' values greater than 10.0 across 2 or more samples were considered for this analysis. This is so that outlier noise in single samples will be removed from the analysis. Finally, from the filtered genes, all genes with 'REC' values >10.0 filter are considered genes with potential stable structures for each sample. A threshold of 10.0 was selected based on manual inspection of the data in Integrative Genomics Viewer (IGV) (Robinson et al. 2011) as genes with lower values were typically non-specific noise. We did not normalize data by gene length (similar to the popular RPKM i.e. reads per kilobase per million values used in total RNA-seq) as structural data is not as quantitative and does not necessarily contain reads across the entire gene. A longer gene might contain the

same number of reads as a similarly expressed shorter gene due to structural similarities and thus, using RPKM would penalize the longer genes. We use the 'REC' values to get a global picture about the genes that form duplex RNA structures across the three developmental stages.

UTR and splice junction analysis

For each coding transcript in the *P.falciparum* genome, we calculated the counts for the duplex RNA-seq at each position in the 5'UTR, 3'UTR as well as 100bp downstream of the start codon and 100bp upstream of the stop codon. Since the UTRs vary drastically in size (data not shown), we normalized the UTR lengths into two bins of 50 and 500. All genes with UTRs shorter than 50bp were not considered in this analysis. The threshold of 500 was selected as it is the median UTR length. All UTRs shorter than 500bp were normalized to 50 bins while all UTRs longer than 500bp were normalized to 500bp. We then calculated the normalized expression ratio calculcated against the maximum duplex RNA-seq expression that came from each position in the binned UTR and the 100bp down/upstream coding regions. This was done to study the relative distribution of duplex RNA-seq reads at the 5' and 3' gene ends. As low gene expression can make the trends noisy due to their stochastic distributions, we only considered the top 20% genes with UTR expression (~850-1200 genes across the six samples). For each stage, the replicates were averaged for the final analysis. A similar analysis was done for the splice junctions by studying 50bp flanking the top 20% expressed splice junctions.

Differential expression analysis (GO/KEGG)

Count data was generated for all duplex RNA-seq datasets using the annotated transcripts from (Otto et al. 2010) using featurecounts as described above. Differential gene expression was performed using DESeq2 v1.26.0 in R v3.6.2. The differentially expressed genes across all three stages (adjusted p-value<0.05) were used for GO analysis (198 genes), which was performed using GO terms from PlasmoDB and topGO v 2.38.1 package in R. The 'Molecular Function" ontology was used.

***P. falciparum* gene expression (RNA-seq) data**

Fastq files of mRNA-seq data was downloaded from the previously published (Otto et al. 2010) dataset. The data was run through the same bioinformatics processing as described about for duplex RNA-seq. The resulting gene counts were normalized using DESeq2 and used for comparison.

Antisense RNA immunoprecipitation with J2 antibody

For immunoprecipitation of total RNA, 50 µg of RNA from asynchronous, mix stage *P. falciparum* culture was incubated overnight at 4 °C in the reaction buffer containing 15 mM Tris·HCl (pH 7.5),150mM KCl, 5mM $MgCl_2$ and 0.5% Triton X-in presence of RNase inhibitor with 10 µg J2 antibody (Scicons). Then, 50 µL of washed (three times in reaction buffer) protein G- Dynabeads magnetic beads were added to the solution of RNA and J2 antibody

(ThermoFisher) and then incubated for 4 hours at 4 °C. Complexes were washed four times in reaction buffer and RNA was recovered by phenol-chloroform extraction and ethanol precipitation. RNA integrity, quality and quantity were measured by Qubit 4 fluorometer (ThermoFisher).

Northern blotting

2ug of J2 immunoprecipitated RNA for each sample was resolved on 6% polyacrylamide gel containing 7 M urea, 90 mM Tris-Borate, and 2.5 mM EDTA (TBE) (pH 8.3) after heating at 95°C for 3 min and instantly chilling on ice. RNA from the gel was subsequently transferred onto N+ nylon membranes (Amersham) using 300 mA constant current for 4 h at 4°C and then cross-linked onto the membrane by UV light. Northern hybridization was done in 0.5 M sodium phosphate buffer (pH 7.6) with gamma-32P-ATP labeled specific oligodeoxyribonucleotides at ~5×10^6 cpm/mL. Following hybridization, membranes were washed twice in 2× SSC, 0.1% SDS at hybridization temperature and once at 0.1× SSC, 0.1% SDS at room temperature for 15 min each and exposed to a PhosphorImager.

RNA-SHAPE analysis

For SHAPE modification of *P. falciparum* U3 RNA, 5ug of *in vitro* transcribed RNA was denatured by heating at 95^0C for 2 min followed by immediate incubation in ice for 3 min. Modification was carried out by incubating the RNA in 3X SHAPE buffer consisting of 333mM of NaCl, 333mM of HEPES and 20mM of MgCl 2 at 37^0C for 30 minutes prior to the addition of 100mM of 2-methylnicotinic acid imidazolide (NAI) for 15 min at 37^0C. As a negative control, equivalent volume of DMSO was added to the RNA. Based on the modifications obtained from our denaturing gel experiments, we derived secondary structure of *P. falciparum* U3 RNA, based on previously established model (Chakrabarti et al. 2007) using RNAstructure (Xu and Mathews 2016).

Supplemental Material

Supplemental material is available for this article.

Acknowledgements

We thank Rania Elbarki, Dipendra Gautam, Arundhati Mohanta and other past and current members of the Chakrabarti laboratory for help with experiments and data analysis. This work was part of DRA's undergraduate honors thesis project, as part of UNC Charlotte Honors Program in Biological Sciences. This work was supported by UNC Charlotte Faculty Development Award and a consortium award by DSF Charitable foundation to KC.

Author Contributions: Conceived and designed the experiments: DRA, AO, SK, KC., Performed the experiments: DRA, AO, TB, BZ, AD, SB, and KC. Analyzed the data: DRA, TB, SK, KC. Wrote the paper: SK and KC.

References

Adjalley SH, Chabbert CD, Klaus B, Pelechano V, Steinmetz LM. 2016. Landscape and Dynamics of Transcription Initiation in the Malaria Parasite Plasmodium falciparum. *Cell Rep* **14**: 2463-2475.

Ahmed R, Duncan RF. 2004. Translational regulation of Hsp90 mRNA. AUG-proximal 5'-untranslated region elements essential for preferential heat shock translation. *J Biol Chem* **279**: 49919-49930.

Amit-Avraham I, Pozner G, Eshar S, Fastman Y, Kolevzon N, Yavin E, Dzikowski R. 2015. Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum. *Proc Natl Acad Sci U S A* **112**: E982-991.

Amulic B, Salanti A, Lavstsen T, Nielsen MA, Deitsch KW. 2009. An upstream open reading frame controls translation of var2csa, a gene implicated in placental malaria. *PLoS Pathog* **5**: e1000256.

Ao W, Gaudet J, Kent WJ, Muttumu S, Mango SE. 2004. Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. *Science* **305**: 1743-1746.

Balaji S, Babu MM, Iyer LM, Aravind L. 2005. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* **33**: 3994-4006.

Barrick JE, Breaker RR. 2007. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* **8**: R239.

Baum J, Papenfuss AT, Mair GR, Janse CJ, Vlachou D, Waters AP, Cowman AF, Crabb BS, de Koning-Ward TF. 2009. Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res* **37**: 3788-3798.

Baumgarten S, Bryant JM, Sinha A, Reyser T, Preiser PR, Dedon PC, Scherf A. 2019. Transcriptome-wide dynamics of extensive m(6)A mRNA methylation during Plasmodium falciparum blood-stage development. *Nat Microbiol* **4**: 2246-2259.

Beaudoin JD, Novoa EM, Vejnar CE, Yartseva V, Takacs CM, Kellis M, Giraldez AJ. 2018. Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat Struct Mol Biol* **25**: 677-686.

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

Brachman EE, Kmiec EB. 2004. DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. *J Cell Sci* **117**: 3867-3874.

Breaker RR. 2018. Riboswitches and Translation Control. *Cold Spring Harb Perspect Biol* **10**.

Broadbent KM, Broadbent JC, Ribacke U, Wirth D, Rinn JL, Sabeti PC. 2015. Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies

developmentally regulated long non-coding RNA and circular RNA. *BMC Genomics* **16**: 454.

Brower-Sinning R, Carter DM, Crevar CJ, Ghedin E, Ross TM, Benos PV. 2009. The role of RNA folding free energy in the evolution of the polymerase genes of the influenza A virus. *Genome Biol* **10**: R18.

Bunnik EM, Batugedara G, Saraf A, Prudhomme J, Florens L, Le Roch KG. 2016. The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. *Genome Biol* **17**: 147.

Bunnik EM, Chung DW, Hamilton M, Ponts N, Saraf A, Prudhomme J, Florens L, Le Roch KG. 2013. Polysome profiling reveals translational control of gene expression in the human malaria parasite Plasmodium falciparum. *Genome Biol* **14**: R128.

Caro F, Ahyong V, Betegon M, DeRisi JL. 2014. Genome-wide regulatory dynamics of translation in the Plasmodium falciparum asexual blood stages. *Elife* **3**.

Carthew RW, Sontheimer EJ. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**: 642-655.

Chakrabarti K, Pearson M, Grate L, Sterne-Weiler T, Deans J, Donohue JP, Ares M, Jr. 2007. Structural RNAs of known and unknown function identified in malaria parasites by comparative genomics and RNA analysis. *RNA* **13**: 1923-1939.

Chan S, Frasch A, Mandava CS, Ch'ng JH, Quintana MDP, Vesterlund M, Ghorbal M, Joannin N, Franzen O, Lopez-Rubio JJ et al. 2017. Regulation of PfEMP1-VAR2CSA translation by a Plasmodium translation-enhancing factor. *Nat Microbiol* **2**: 17068.

Chappell L, Ross P, Orchard L, Russell TJ, Otto TD, Berriman M, Rayner JC, Llinas M. 2020. Refining the transcriptome of the human malaria parasite Plasmodium falciparum using amplification-free RNA-seq. *BMC Genomics* **21**: 395.

Cheah MT, Wachter A, Sudarsan N, Breaker RR. 2007. Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* **447**: 497-500.

Chen Q, Schlichtherle M, Wahlgren M. 2000. Molecular aspects of severe malaria. *Clin Microbiol Rev* **13**: 439-450.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884-i890.

Cheviet T, Lefebvre-Tournier I, Wein S, Peyrottes S. 2019. Plasmodium Purine Metabolism and Its Inhibition by Nucleoside and Nucleotide Analogues. *J Med Chem* **62**: 8365-8391.

Clingman CC, Ryder SP. 2013. Metabolite sensing in eukaryotic mRNA biology. *Wiley Interdiscip Rev RNA* **4**: 387-396.

Collins CR, Hackett F, Atid J, Tan MSY, Blackman MJ. 2017. The Plasmodium falciparum pseudoprotease SERA5 regulates the kinetics and efficiency of malaria parasite egress from host erythrocytes. *PLoS Pathog* **13**: e1006453.

Corley M, Burns MC, Yeo GW. 2020. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol Cell* **78**: 9-29.

Coulson RM, Hall N, Ouzounis CA. 2004. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. *Genome Res* **14**: 1548-1554.

Cui L, Lindner S, Miao J. 2015. Translational regulation during stage transitions in malaria parasites. *Ann N Y Acad Sci* **1342**: 1-9.

Davila Lopez M, Samuelsson T. 2008. Early evolution of histone mRNA 3' end processing. *RNA* **14**: 1-10.

Ding Y, Kwok CK, Tang Y, Bevilacqua PC, Assmann SM. 2015. Genome-wide profiling of in vivo RNA structure at single-nucleotide resolution using structure-seq. *Nat Protoc* **10**: 1050-1066.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.

Dominguez D, Freese P, Alexis MS, Su A, Hochman M, Palden T, Bazile C, Lambert NJ, Van Nostrand EL, Pratt GA et al. 2018. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol Cell* **70**: 854-867 e859.

Duffy S, Avery VM. 2018. Routine In Vitro Culture of Plasmodium falciparum: Experimental Consequences? *Trends Parasitol* **34**: 564-575.

Epp C, Li F, Howitt CA, Chookajorn T, Deitsch KW. 2009. Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite Plasmodium falciparum. *RNA* **15**: 116-127.

Faghihi MA, Wahlestedt C. 2009. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol* **10**: 637-643.

Favaloro JM, Coppel RL, Corcoran LM, Foote SJ, Brown GV, Anders RF, Kemp DJ. 1986. Structure of the RESA gene of Plasmodium falciparum. *Nucleic Acids Res* **14**: 8265-8277.

Foley SW, Vandivier LE, Kuksa PP, Gregory BD. 2015. Transcriptome-wide measurement of plant RNA secondary structure. *Curr Opin Plant Biol* **27**: 36-43.

Gardiner DL, Holt DC, Thomas EA, Kemp DJ, Trenholme KR. 2000. Inhibition of Plasmodium falciparum clag9 gene function by antisense RNA. *Mol Biochem Parasitol* **110**: 33-41.

Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S et al. 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**: 498-511.

Goyal M, Alam A, Iqbal MS, Dey S, Bindu S, Pal C, Banerjee A, Chakrabarti S, Bandyopadhyay U. 2012. Identification and molecular characterization of an

Alba-family protein from human malaria parasite Plasmodium falciparum. *Nucleic Acids Res* **40**: 1174-1190.
Gunasekera AM, Patankar S, Schug J, Eisen G, Kissinger J, Roos D, Wirth DF. 2004. Widespread distribution of antisense transcripts in the Plasmodium falciparum genome. *Mol Biochem Parasitol* **136**: 35-42.
Halic M, Becker T, Pool MR, Spahn CM, Grassucci RA, Frank J, Beckmann R. 2004. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature* **427**: 808-814.
Henkin TM. 2008. Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* **22**: 3383-3390.
Hughes KR, Philip N, Starnes GL, Taylor S, Waters AP. 2010. From cradle to grave: RNA biology in malaria parasites. *Wiley Interdiscip Rev RNA* **1**: 287-303.
Jacobs E, Mills JD, Janitz M. 2012. The role of RNA structure in posttranscriptional regulation of gene expression. *J Genet Genomics* **39**: 535-543.
Jiang L, Mu J, Zhang Q, Ni T, Srinivasan P, Rayavara K, Yang W, Turner L, Lavstsen T, Theander TG et al. 2013. PfSETvs methylation of histone H3K36 represses virulence genes in Plasmodium falciparum. *Nature* **499**: 223-227.
Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. 2001. Protein-RNA interactions: a structural analysis. *Nucleic Acids Res* **29**: 943-954.
Kertesz M, Wan Y, Mazor E, Rinn JL, Nutter RC, Chang HY, Segal E. 2010. Genome-wide measurement of RNA secondary structure in yeast. *Nature* **467**: 103-107.
Kramer MC, Gregory BD. 2018. Does RNA secondary structure drive translation or vice versa? *Nat Struct Mol Biol* **25**: 641-643.
Kwok CK, Tang Y, Assmann SM, Bevilacqua PC. 2015. The RNA structurome: transcriptome-wide structure probing with next-generation sequencing. *Trends Biochem Sci* **40**: 221-232.
Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M, Yan SF, Williamson KC, Holder AA, Carucci DJ et al. 2004. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome Res* **14**: 2308-2318.
Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ et al. 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**: 1503-1508.
Leach MD, Cowen LE. 2013. Surviving the heat of the moment: a fungal pathogens perspective. *PLoS Pathog* **9**: e1003163.
Leppek K, Das R, Barna M. 2018. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**: 158-174.
Li F, Zheng Q, Ryvkin P, Dragomir I, Desai Y, Aiyer S, Valladares O, Yang J, Bambina S, Sabin LR et al. 2012. Global analysis of RNA secondary structure in two metazoans. *Cell Rep* **1**: 69-82.

Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923-930.

Lim YB, Thingna J, Kong F, Dao M, Cao J, Lim CT. 2020. Temperature-Induced Catch-Slip to Slip Bond Transit in Plasmodium falciparum-Infected Erythrocytes. *Biophys J* **118**: 105-116.

Liu L, Rice MC, Drury M, Cheng S, Gamper H, Kmiec EB. 2002. Strand bias in targeted gene repair is influenced by transcriptional activity. *Mol Cell Biol* **22**: 3852-3863.

Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. 2015. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**: 560-564.

Llinas M, Bozdech Z, Wong ED, Adai AT, DeRisi JL. 2006. Comparative whole genome transcriptome analysis of three Plasmodium falciparum strains. *Nucleic Acids Res* **34**: 1166-1173.

Llinas M, DeRisi JL. 2004. Pernicious plans revealed: Plasmodium falciparum genome wide expression analysis. *Curr Opin Microbiol* **7**: 382-387.

Lopez-Barragan MJ, Lemieux J, Quinones M, Williamson KC, Molina-Cruz A, Cui K, Barillas-Mury C, Zhao K, Su XZ. 2011. Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. *BMC Genomics* **12**: 587.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lu XM, Batugedara G, Lee M, Prudhomme J, Bunnik EM, Le Roch KG. 2017. Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum. *Nucleic Acids Res* **45**: 7825-7840.

Lu Z, Gong J, Zhang QC. 2018. PARIS: Psoralen Analysis of RNA Interactions and Structures with High Throughput and Resolution. *Methods Mol Biol* **1649**: 59-84.

Mair GR, Braks JA, Garver LS, Wiegant JC, Hall N, Dirks RW, Khan SM, Dimopoulos G, Janse CJ, Waters AP. 2006. Regulation of sexual development of Plasmodium by translational repression. *Science* **313**: 667-669.

Mair GR, Lasonder E, Garver LS, Franke-Fayard BM, Carret CK, Wiegant JC, Dirks RW, Dimopoulos G, Janse CJ, Waters AP. 2010. Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development. *PLoS Pathog* **6**: e1000767.

Mamillapalli A, Pattnaik P, Sharma M, Sharma SK, Tyagi PK, Joshi H, Chitnis CE. 2006. Sequence polymorphisms in the receptor-binding domain of Plasmodium falciparum EBA-175: implications for malaria vaccine development. *Mol Biochem Parasitol* **146**: 120-123.

Mao Y, Dong L, Liu XM, Guo J, Ma H, Shen B, Qian SB. 2019. m(6)A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun* **10**: 5332.

Mao Y, Liu H, Liu Y, Tao S. 2014. Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in Saccharomyces cerevisiae. *Nucleic Acids Res* **42**: 4813-4822.

Mauger DM, Cabral BJ, Presnyak V, Su SV, Reid DW, Goodman B, Link K, Khatwani N, Reynders J, Moore MJ et al. 2019. mRNA structure regulates protein expression through changes in functional half-life. *Proc Natl Acad Sci U S A* **116**: 24075-24083.

McGinnis JL, Liu Q, Lavender CA, Devaraj A, McClory SP, Fredrick K, Weeks KM. 2015. In-cell SHAPE reveals that free 30S ribosome subunits are in the inactive state. *Proc Natl Acad Sci U S A* **112**: 2425-2430.

Meyer S, Carlson PD, Lucks JB. 2017. Characterizing the Structure-Function Relationship of a Naturally Occurring RNA Thermometer. *Biochemistry* **56**: 6629-6638.

Militello KT, Patel V, Chessler AD, Fisher JK, Kasper JM, Gunasekera A, Wirth DF. 2005. RNA polymerase II synthesizes antisense RNA in Plasmodium falciparum. *RNA* **11**: 365-370.

Mourier T, Carret C, Kyes S, Christodoulou Z, Gardner PP, Jeffares DC, Pinches R, Barrell B, Berriman M, Griffiths-Jones S et al. 2008. Genome-wide discovery and verification of novel structured RNAs in Plasmodium falciparum. *Genome Res* **18**: 281-292.

Mustoe AM, Corley M, Laederach A, Weeks KM. 2018. Messenger RNA Structure Regulates Translation Initiation: A Mechanism Exploited from Bacteria to Humans. *Biochemistry* **57**: 3537-3539.

Narberhaus F. 2010. Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. *RNA Biol* **7**: 84-89.

Nikolova EN, Al-Hashimi HM. 2010. Thermodynamics of RNA melting, one base pair at a time. *RNA* **16**: 1687-1691.

Oakley MS, Gerald N, McCutchan TF, Aravind L, Kumar S. 2011. Clinical and molecular aspects of malaria fever. *Trends Parasitol* **27**: 442-449.

Otto TD, Wilinski D, Assefa S, Keane TM, Sarry LR, Bohme 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.

Painter HJ, Chung NC, Sebastian A, Albert I, Storey JD, Llinas M. 2018. Genome-wide real-time in vivo transcriptional dynamics during Plasmodium falciparum blood-stage development. *Nat Commun* **9**: 2656.

Pei X, Guo X, Coppel R, Bhattacharjee S, Haldar K, Gratzer W, Mohandas N, An X. 2007. The ring-infected erythrocyte surface antigen (RESA) of Plasmodium falciparum stabilizes spectrin tetramers and suppresses further invasion. *Blood* **110**: 1036-1042.

Piao M, Sun L, Zhang QC. 2017. RNA Regulations and Functions Decoded by Transcriptome-wide RNA Structure Probing. *Genomics Proteomics Bioinformatics* **15**: 267-278.

Qi F, Frishman D. 2017. Melting temperature highlights functionally important RNA structure and sequence elements in yeast mRNA coding regions. *Nucleic Acids Res* **45**: 6109-6118.

Rabani M, Kertesz M, Segal E. 2008. Computational prediction of RNA structural motifs involved in posttranscriptional regulatory processes. *Proc Natl Acad Sci U S A* **105**: 14885-14890.

Ralph SA, Bischoff E, Mattei D, Sismeiro O, Dillies MA, Guigon G, Coppee JY, David PH, Scherf A. 2005. Transcriptome analysis of antigenic variation in Plasmodium falciparum--var silencing is not dependent on antisense RNA. *Genome Biol* **6**: R93.

Reddy BP, Shrestha S, Hart KJ, Liang X, Kemirembe K, Cui L, Lindner SE. 2015. A bioinformatic survey of RNA-binding proteins in Plasmodium. *BMC Genomics* **16**: 890.

Rios KT, Lindner SE. 2019. Protein-RNA interactions important for Plasmodium transmission. *PLoS Pathog* **15**: e1008095.

Rodnina MV. 2016. The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Sci* **25**: 1390-1406.

Rogers MJ, Gutell RR, Damberger SH, Li J, McConkey GA, Waters AP, McCutchan TF. 1996. Structural features of the large subunit rRNA expressed in Plasmodium falciparum sporozoites that distinguish it from the asexually expressed subunit rRNA. *RNA* **2**: 134-145.

Roundtree IA, Evans ME, Pan T, He C. 2017. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **169**: 1187-1200.

Rouskin S, Zubradt M, Washietl S, Kellis M, Weissman JS. 2014. Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* **505**: 701-705.

Sanchez de Groot N, Armaos A, Grana-Montes R, Alriquet M, Calloni G, Vabulas RM, Tartaglia GG. 2019. RNA structure drives interaction with proteins. *Nat Commun* **10**: 3246.

Sharp PA. 2009. The centrality of RNA. *Cell* **136**: 577-580.

Siegel TN, Hon CC, Zhang Q, Lopez-Rubio JJ, Scheidig-Benatar C, Martins RM, Sismeiro O, Coppee JY, Scherf A. 2014. Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in Plasmodium falciparum. *BMC Genomics* **15**: 150.

Sims JS, Militello KT, Sims PA, Patel VP, Kasper JM, Wirth DF. 2009. Patterns of gene-specific and total transcriptional activity during the Plasmodium falciparum intraerythrocytic developmental cycle. *Eukaryot Cell* **8**: 327-338.

Siwo GH, Tan A, Button-Simons KA, Samarakoon U, Checkley LA, Pinapati RS, Ferdig MT. 2015. Predicting functional and regulatory divergence of a drug resistance transporter gene in the human malaria parasite. *BMC Genomics* **16**: 115.

Smola MJ, Christy TW, Inoue K, Nicholson CO, Friedersdorf M, Keene JD, Lee DM, Calabrese JM, Weeks KM. 2016. SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. *Proc Natl Acad Sci U S A* **113**: 10322-10327.

Smola MJ, Weeks KM. 2018. In-cell RNA structure probing with SHAPE-MaP. *Nat Protoc* **13**: 1181-1195.

Sorber K, Dimon MT, DeRisi JL. 2011. RNA-Seq analysis of splicing in Plasmodium falciparum uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. *Nucleic Acids Res* **39**: 3820-3835.

Spitale RC, Crisalli P, Flynn RA, Torre EA, Kool ET, Chang HY. 2013. RNA SHAPE analysis in living cells. *Nat Chem Biol* **9**: 18-20.

Strick TR, Savery NJ. 2017. Understanding bias in DNA repair. *Proc Natl Acad Sci U S A* **114**: 2791-2793.

Sudarsan N, Barrick JE, Breaker RR. 2003. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**: 644-647.

Toenhake CG, Fraschka SA, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bartfai R. 2018. Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. *Cell Host Microbe* **23**: 557-569 e559.

Trager W, Jensen JB. 1976. Human malaria parasites in continuous culture. *Science* **193**: 673-675.

Val F, Costa FT, King L, Brito-Sousa JD, Bassat Q, Monteiro WM, Siqueira AM, Luzzatto L, Lacerda MV. 2019. Tafenoquine for the prophylaxis, treatment and elimination of malaria: eagerness must meet prudence. *Future Microbiol* **14**: 1261-1279.

Vembar SS, Droll D, Scherf A. 2016. Translational regulation in blood stages of the malaria parasite Plasmodium spp.: systems-wide studies pave the way. *Wiley Interdiscip Rev RNA* **7**: 772-792.

Vembar SS, Scherf A, Siegel TN. 2014. Noncoding RNAs as emerging regulators of Plasmodium falciparum virulence gene expression. *Curr Opin Microbiol* **20**: 153-161.

Wan Y, Kertesz M, Spitale RC, Segal E, Chang HY. 2011. Understanding the transcriptome through RNA structure. *Nat Rev Genet* **12**: 641-655.

Wan Y, Qu K, Ouyang Z, Kertesz M, Li J, Tibshirani R, Makino DL, Nutter RC, Segal E, Chang HY. 2012. Genome-wide measurement of RNA folding energies. *Mol Cell* **48**: 169-181.

Wan Y, Qu K, Zhang QC, Flynn RA, Manor O, Ouyang Z, Zhang J, Spitale RC, Snyder MP, Segal E et al. 2014. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature* **505**: 706-709.

Warner KD, Hajdin CE, Weeks KM. 2018. Principles for targeting RNA with drug-like small molecules. *Nat Rev Drug Discov* **17**: 547-558.

Xu ZZ, Mathews DH. 2016. Experiment-Assisted Secondary Structure Prediction with RNAstructure. *Methods Mol Biol* **1490**: 163-176.

Zhang J, Tan D, DeRose EF, Perera L, Dominski Z, Marzluff WF, Tong L, Hall TM. 2014a. Molecular mechanisms for the regulation of histone mRNA stem-loop-binding protein by phosphorylation. *Proc Natl Acad Sci U S A* **111**: E2937-2946.

Zhang Q, Siegel TN, Martins RM, Wang F, Cao J, Gao Q, Cheng X, Jiang L, Hon CC, Scheidig-Benatar C et al. 2014b. Exonuclease-mediated degradation of nascent RNA silences genes linked to severe malaria. *Nature* **513**: 431-435.

Zhang R, Chandramohanadas R, Lim CT, Dao M. 2018. Febrile Temperature Elevates the Expression of Phosphatidylserine on Plasmodium falciparum (FCR3CSA) Infected Red Blood Cell Surface Leading to Increased Cytoadhesion. *Sci Rep* **8**: 15022.

Zheng Q, Ryvkin P, Li F, Dragomir I, Valladares O, Yang J, Cao K, Wang LS, Gregory BD. 2010. Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. *PLoS Genet* **6**: e1001141.

Figure Legend

Figure-1: RNA sequencing, analysis pipeline and read distribution profile. A) Total RNA samples collected from three different developmental stages (Ring, Trophozoite and Schizont) of *P. falciparum* in human RBCs were subjected to nuclease-based structure mapping using RNase One, followed by purification of base-paired RNAs and strand-specific RNA sequencing. Differentially expressed regions of the *P. falciparum* genome that show significant accumulation of duplex RNA reads were further analyzed by immunoprecipitation using anti-dsRNA for detection of antisense RNAs. Also, notable structural RNAs that were identified by dsRNA-seq were further validated by mapping to secondary structures following RNA-SHAPE analysis. B) Distribution of paired-end duplex RNA-seq reads that align to different annotated categories of coding and noncoding transcripts in *P.falciparum* genome (PlasmoDB). Only reads mapping uniquely to the genome are considered and reads aligning to known rRNAs and tRNAs are removed from this analysis.

Figure-2: (A) Distribution of uniquely mapped reads across known coding regions on all 14 chromosomes by developmental state; *red:* Ring; *blue:* Trophozoite; *green:* Schizont (B) Distribution of uniquely mapped reads across various annotated coding and non-coding regions of *P.falciparum* genome. The relative percentage of reads from forward stranded reads across various gene categories are shown on top and reads from reverse stranded reads are shown on the bottom. The bars denote average percentage of reads across biological replicates and the error bars show standard deviation.

Figure-3: Mapping and validation of duplex RNA-seq data with known RNA secondary structures of noncoding RNAs: U3 RNA (A-C) and SRP RNA (D-F).(A) Two Cy5 labeled denaturing gels for U3 snoRNA domains: Stem I', Hinge, Box B, Stem II', Stem II', Box C and Stem III (Corresponding RNA secondary structure data in panel C). (B) IGV screenshot of the U3 RNA peaks representing duplex RNA-seq read enrichments in three different P. falciparum RBC stages. (C) Secondary structure of the Pf U3 snoRNA based on in-gel SHAPE modification of mix stage RNA using NAI (shown in panel A). Unpaired RNA nucleotides are shown in red and orange color based on their reactivities, where high reactivity is signified by red color and low reactivity by orange with the SHAPE reagent. Base-paired RNA regions identified from duplex RNA-seq in B are indicated by green, yellow and magenta color skirting (also shown by matching colored bars at the bottom of the peaks/reads identified in duplex RNA-seq). (D) Schematic representation of the P. falciparum SRP RNA secondary structure model. Mapped reads from duplex RNA-seq are shown in skirted colored lines along the structure model. Corresponding peaks of the mapped read regions in P. falciparum genome are shown by three color matching boxed regions in F. Due to shorter insert sizes in 3 replicates of the data (**Suppl. Table 2**), the peaks are seen more distinctly and at a higher resolution in these replicates and thus, we selected to display the coverage in these replicates. The remaining three replicates also show the same structures but at a lower resolution. (E) Northern blot analysis of the total RNA from five different time point collections during asexual RBC cycle. (F) Black arrow showing a peak missing from the ring stage samples (also shown by an arrow in structure

model in D), indicating the ability of the duplex RNA-seq to detect stage specific RNA structural rearrangements during *P. falciparum* development.

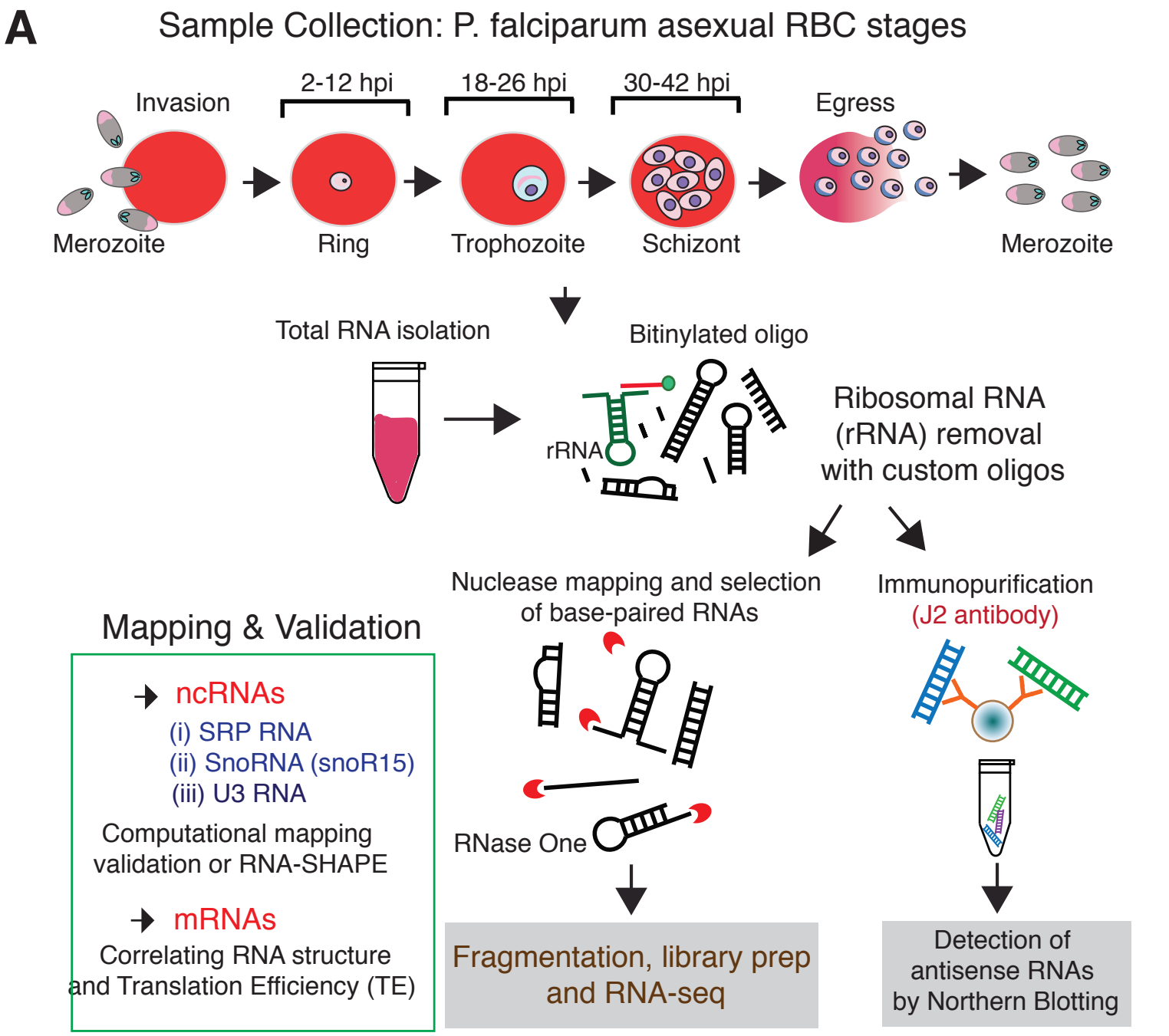
Figure-4: (A) Gene expression of duplex RNA-seq as compared to *P. falciparum* total RNA-seq data (Otto et al. 2010). Representative scatter plot showing correlation between normalized gene counts from total RNA-seq and duplex RNA-seq datasets for Ring stage. The other stages are shown in **Suppl. Fig. 5A and 5B**. A pseudocount of 1.0 is added to the normalized counts for both duplex and total RNA-seq data prior to plotting. The axes are in logarithmic scale. The color and size of the points show the ratio of the duplex RNA-seq normalized counts by the total RNA-seq normalized counts. Larger red points represent genes with higher expression in duplex RNA-seq compared to total RNA-seq as shown in the legend. (B) Heatmap of the 100 more expressed differentially expressed genes across the three developmental stages. The normalized gene counts from DESeq2 are log transformed in the heatmap. The highlighted genes on the right are referenced across the text. (C) Screen shot of duplex RNA enrichments for all six samples for Histone H2A visualized by IGV browser. The gene contains two sets of peaks in the duplex RNA-seq data, highlighted by dashed boxes. The second replicates show a higher resolution of peaks due to shorter insert sizes (as discussed before) (D) A highly represented sequence-structure motif identified from duplex RNA reads of *P. falciparum* mRNAs.

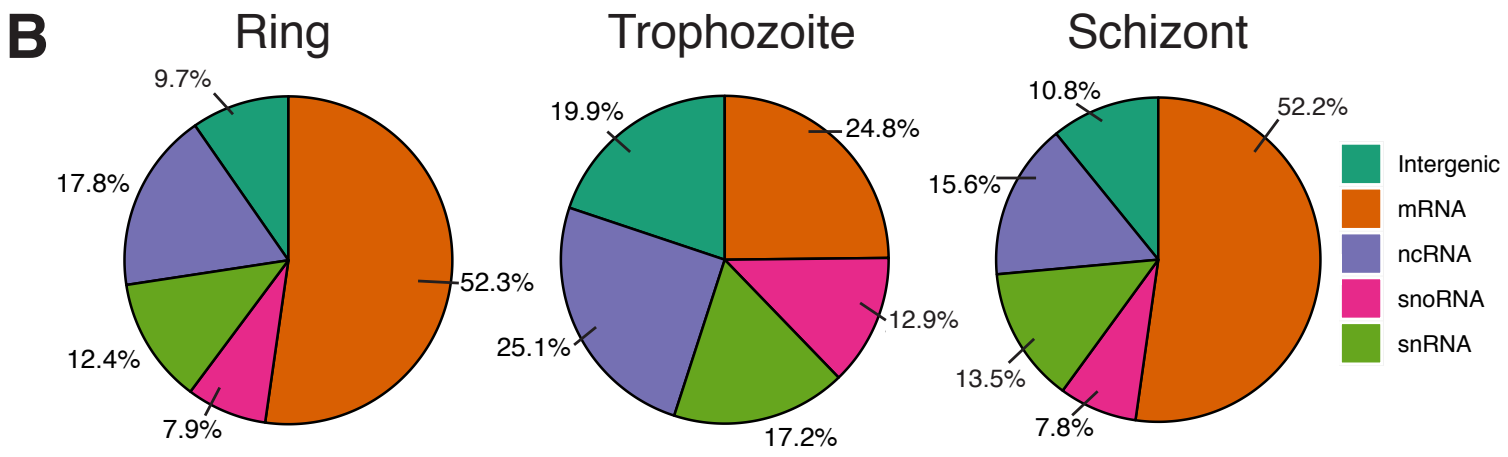
Figure-5: Distribution of duplex RNA-seq data across the top 20% expressed (A) 5' UTRs, (B) 3'UTRs and (C) splice junctions. The y-axis shows the normalized expression ratio, which is normalized against the highest duplex RNA-seq expression in the region being investigated. This ratio represents the relative distribution of enrichment of base-paired RNAs (See Methods). For the UTRs, all UTRs shorter than 500bp (the median UTR length) are binned into 50bp bins and all UTRs longer than 500bp are binned into 500bp bins, shown by the top and bottom panels in (A) and (B). We focus on 100bp of the coding regions downstream of the start codon and upstream of the stop codon for the 5'UTR and 3'UTR respectively. (C) Coding regions of 50bp flanking the splice junction is plotted. The values for the replicates for each stage are averaged for plotting.

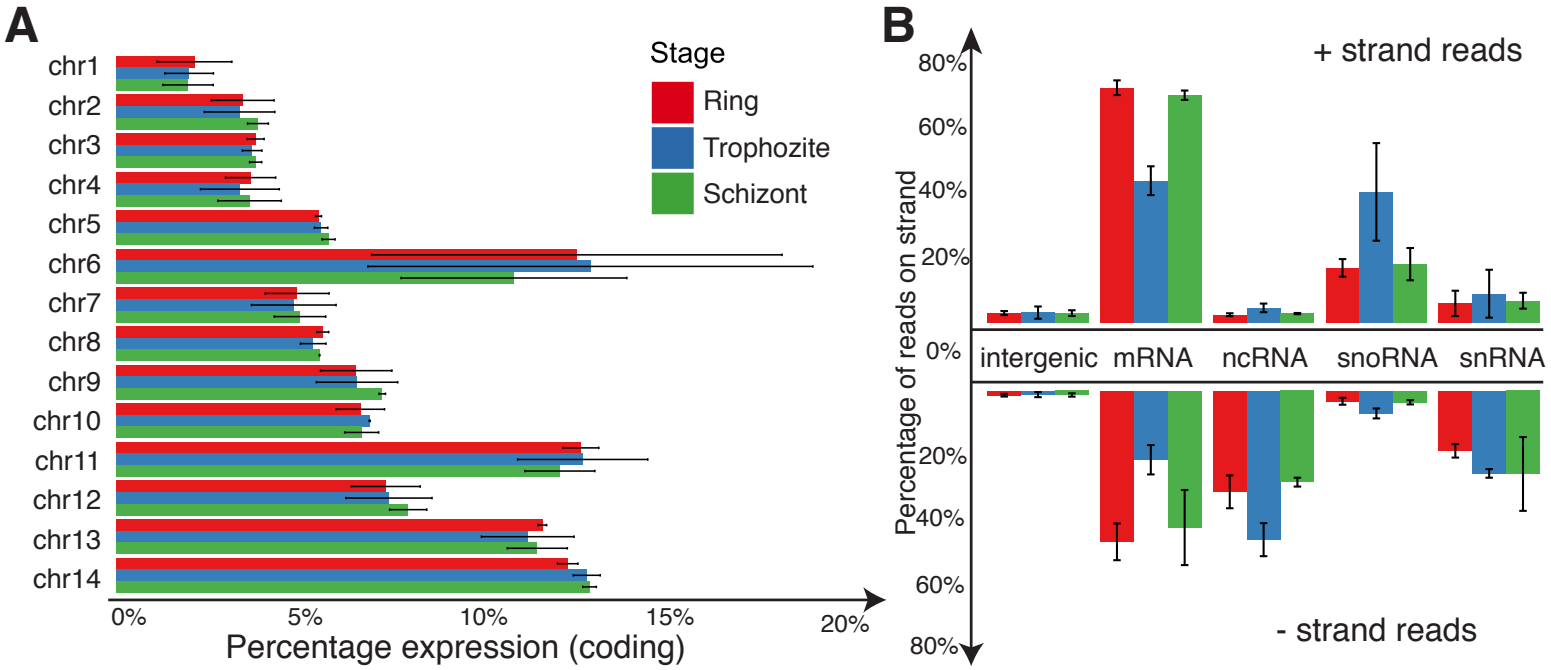
Figure-6: Antisense expression in duplex RNA-seq (A) Antisense duplex RNA-seq counts for the top 50 antisense genes plotted against mRNA-seq counts for Ring Stage. The other stages are shown in Suppl. Fig 4C-D. The size and color of the points indicate the ratio of the duplex RNA-seq antisense to sense counts. Larger red points indicate genes that have more antisense signal than sense signal. Few selected genes with very high and very low ratios are highlighted. A pseudocount of one is added to the normalized counts and axes are in logarithmic scale. (B) A schematic indicating the three types of antisense peaks seen in the duplex RNA-seq data in the 3' UTRs of the genes (i) antisense peak without any overlapping genes (ii) antisense peak with overlapping UTR of another protein coding gene (iii) antisense peak from overlapping non-coding gene(s). (C) An example gene with high antisense peaks originating from non-coding snoRNA and ncRNA (snoR31 and RUF4) across all three stages is shown. (D) Antisense transcripts detected by northern blotting in the J2 antibody immunoprecipitated samples from

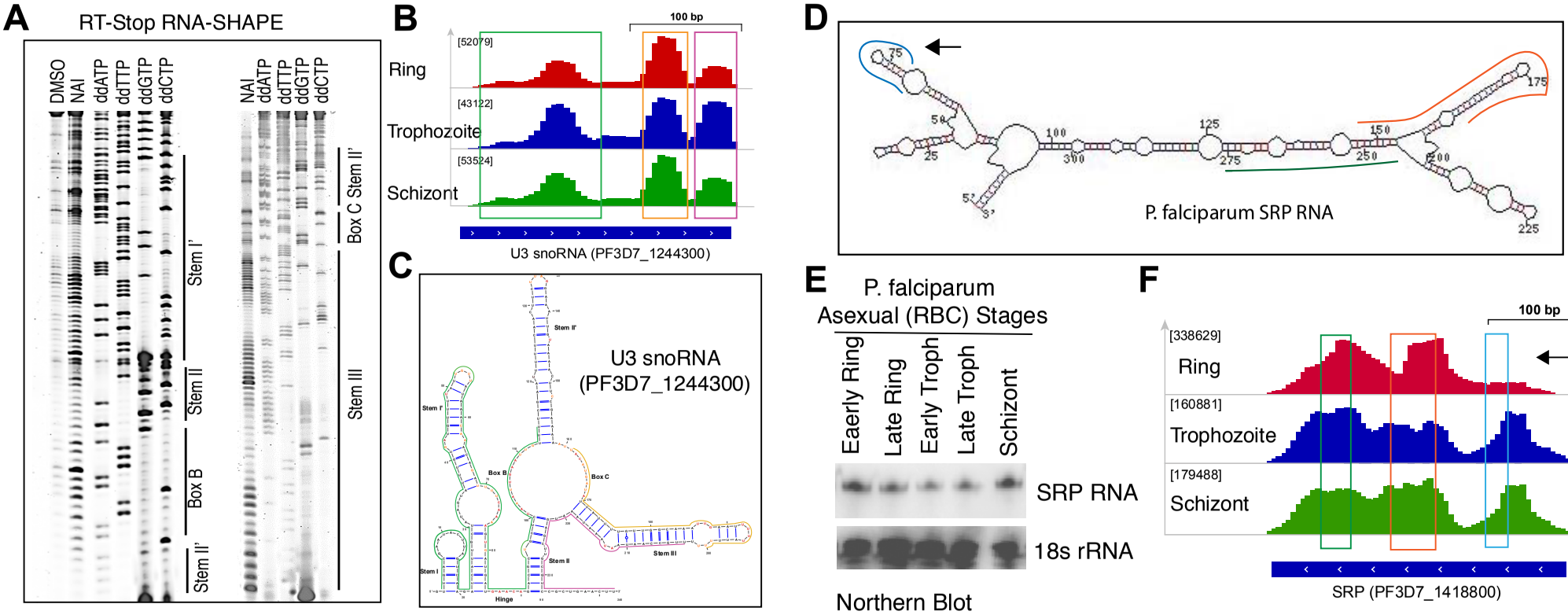
mix (asynchronous) stages *P. falciparum* total RNA. Left blot (in all panels) shows sense gene expression and the right blot shows antisense expression from immunoprecipitated samples.

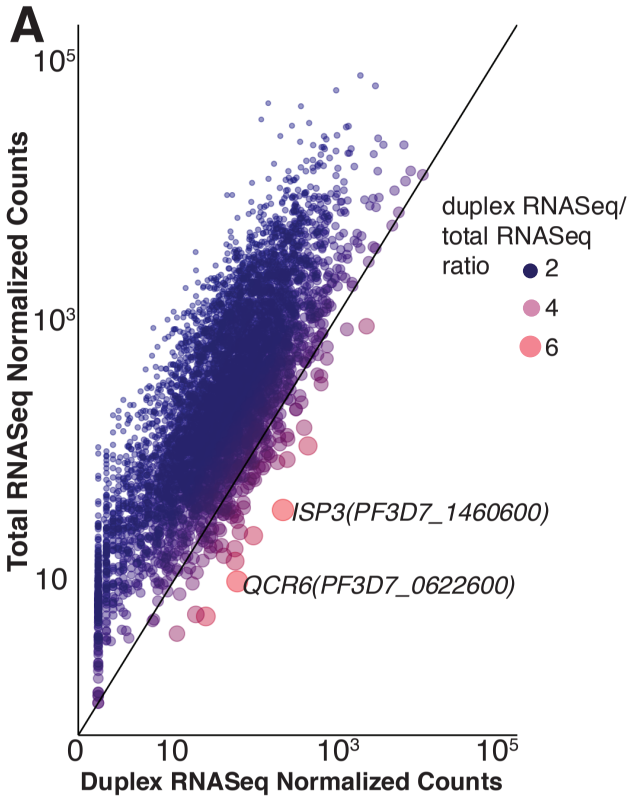
Figure-7: Translation efficiency correlation with duplex RNA-seq structures. Scatter plot of normalized duplex RNA-seq counts plotted against ribosome footprint RPKM data (translation efficiency) from Caro et. al. 2014 for the (A) Ring stage (B) Trophozoite and (C) Schizont stages. The top 50% protein coding genes with structures are plotted. The Pearson correlation coefficient for the three stages are 0.49, 0.51 and 0.53 respectively ($p < 0.001$).

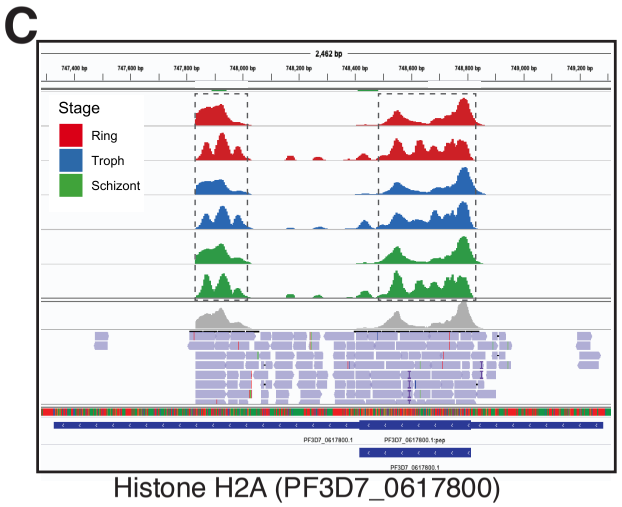

A
Sample Collection: P. falciparum asexual RBC stages
Invasion
2-12 hpi
18-26 hpi
30-42 hpi
Egress
Merozoite
Ring
Trophozoite
Schizont
Merozoite
Total RNA isolation
Bitinylated oligo
rRNA
Ribosomal RNA (rRNA) removal with custom oligos
Nuclease mapping and selection of base-paired RNAs
Immunopurification
(J2 antibody)
Mapping & Validation
ncRNAs
(i) SRP RNA
(ii) SnoRNA (snoR15)
(iii) U3 RNA
Computational mapping validation or RNA-SHAPE
mRNAs
Correlating RNA structure and Translation Efficiency (TE)
RNase One
Fragmentation, library prep and RNA-seq
Detection of antisense RNAs by Northern Blotting


B
Ring
Trophozoite
Schizont
9.7%
17.8%
52.3%
12.4%
7.9%
19.9%
24.8%
12.9%
25.1%
17.2%
10.8%
52.2%
15.6%
13.5%
7.8%
Intergenic
mRNA
ncRNA
snoRNA
snRNA

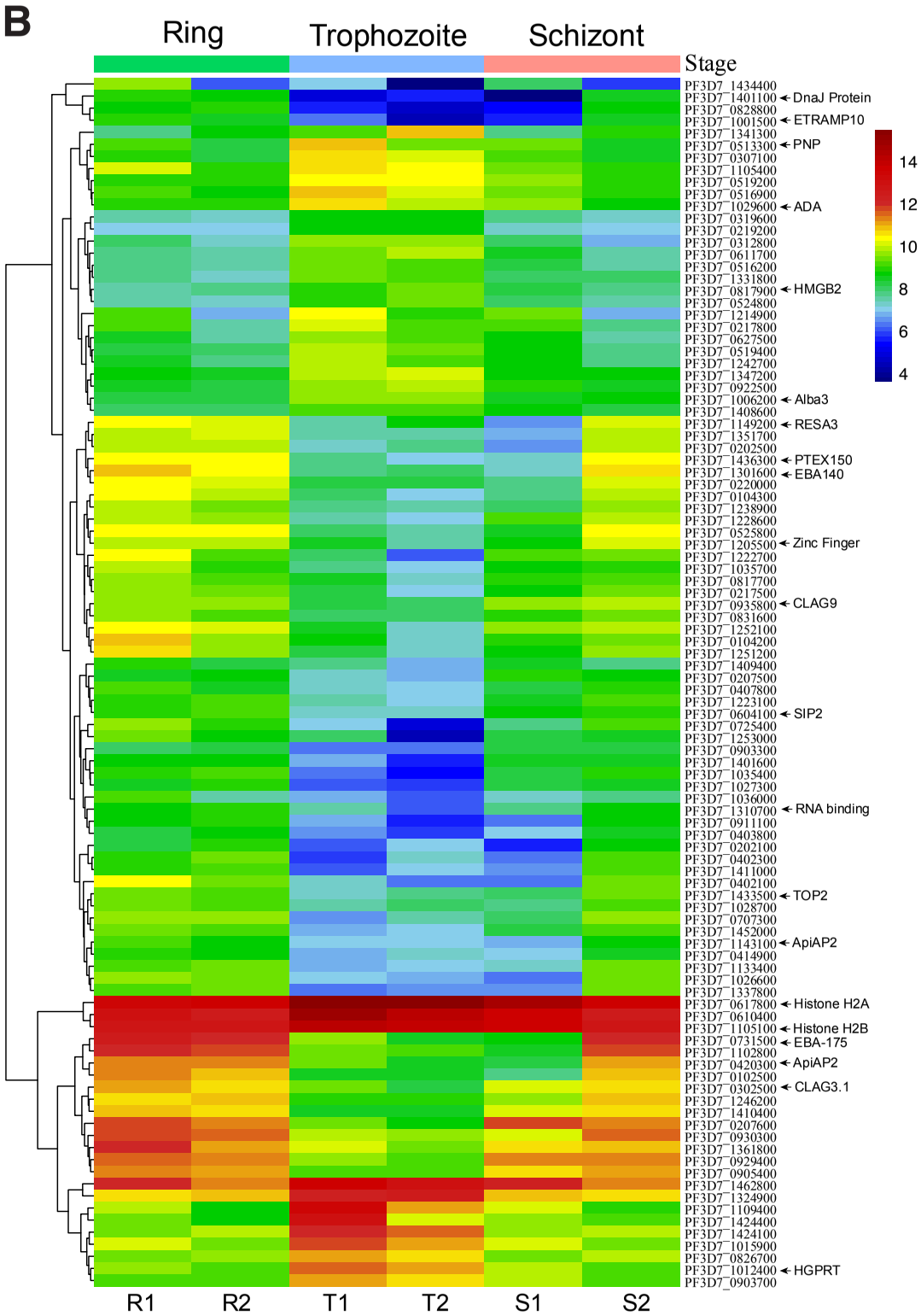

A
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
Stage
Ring
Trophozite
Schizont
0%
5%
10%
15%
20%
Percentage expression (coding)
B
+ strand reads
- strand reads
Percentage of reads on strand
80%
60%
40%
20%
0%
intergenic
mRNA
ncRNA
snoRNA
snRNA

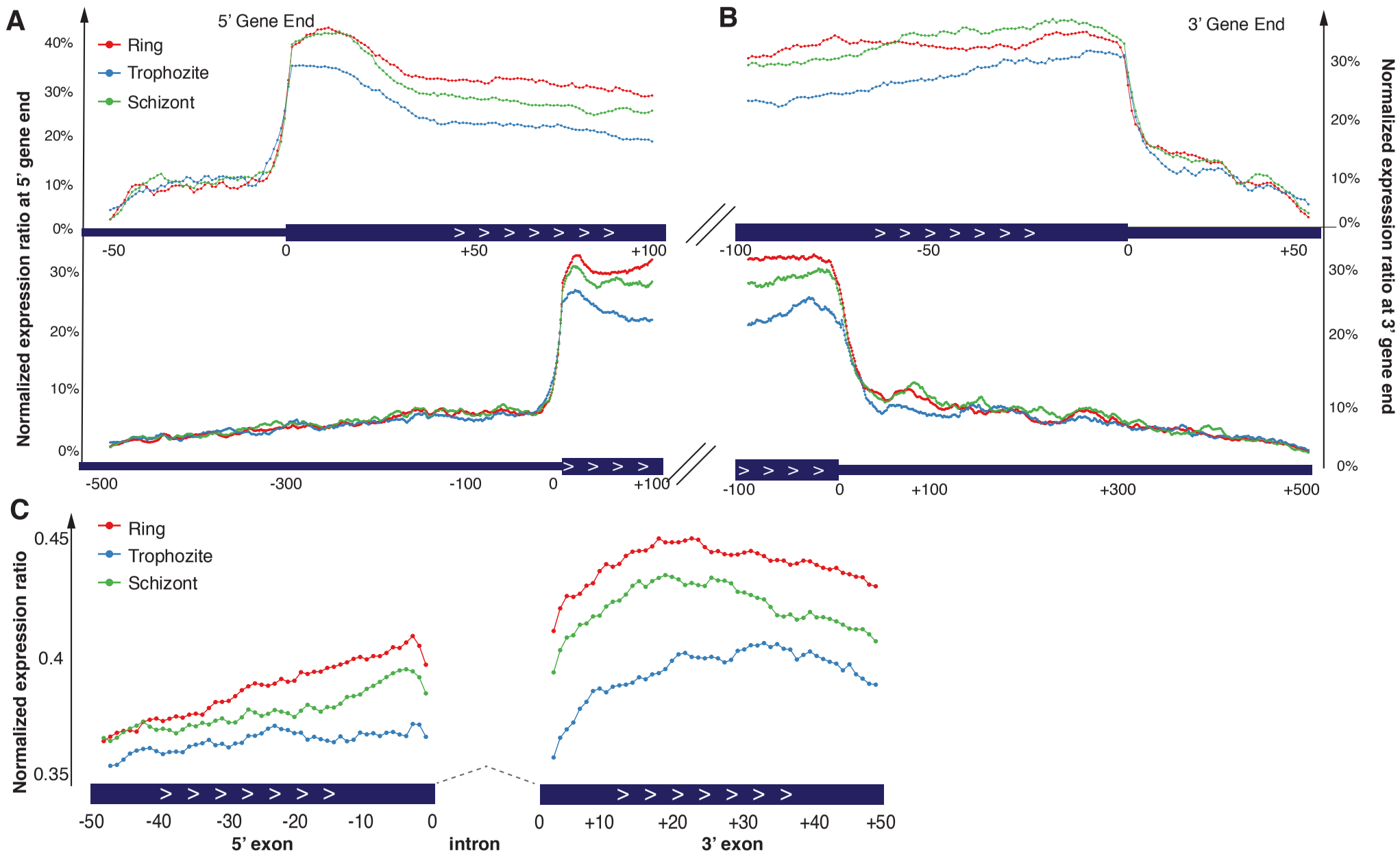

A
RT-Stop RNA-SHAPE
DMSO
NAI
ddATP
ddTTP
ddGTP
ddCTP
NAI
ddATP
ddTTP
ddGTP
ddCTP
Stem I'
Stem II
Box B
Stem II'
Box C Stem II'
Stem III
B
100 bp
[52079]
Ring
[43122]
Trophozoite
[53524]
Schizont
U3 snoRNA (PF3D7_1244300)
C
U3 snoRNA
(PF3D7_1244300)
Stem II'
Stem I'
Box B
Box C
Stem I
Stem II
Stem III
Hinge
D
P. falciparum SRP RNA
E
P. falciparum
Asexual (RBC) Stages
Eaerly Ring
Late Ring
Early Troph
Late Troph
Schizont
SRP RNA
18s rRNA
Northern Blot
F
100 bp
[338629]
Ring
[160881]
Trophozoite
[179488]
Schizont
SRP (PF3D7_1418800)

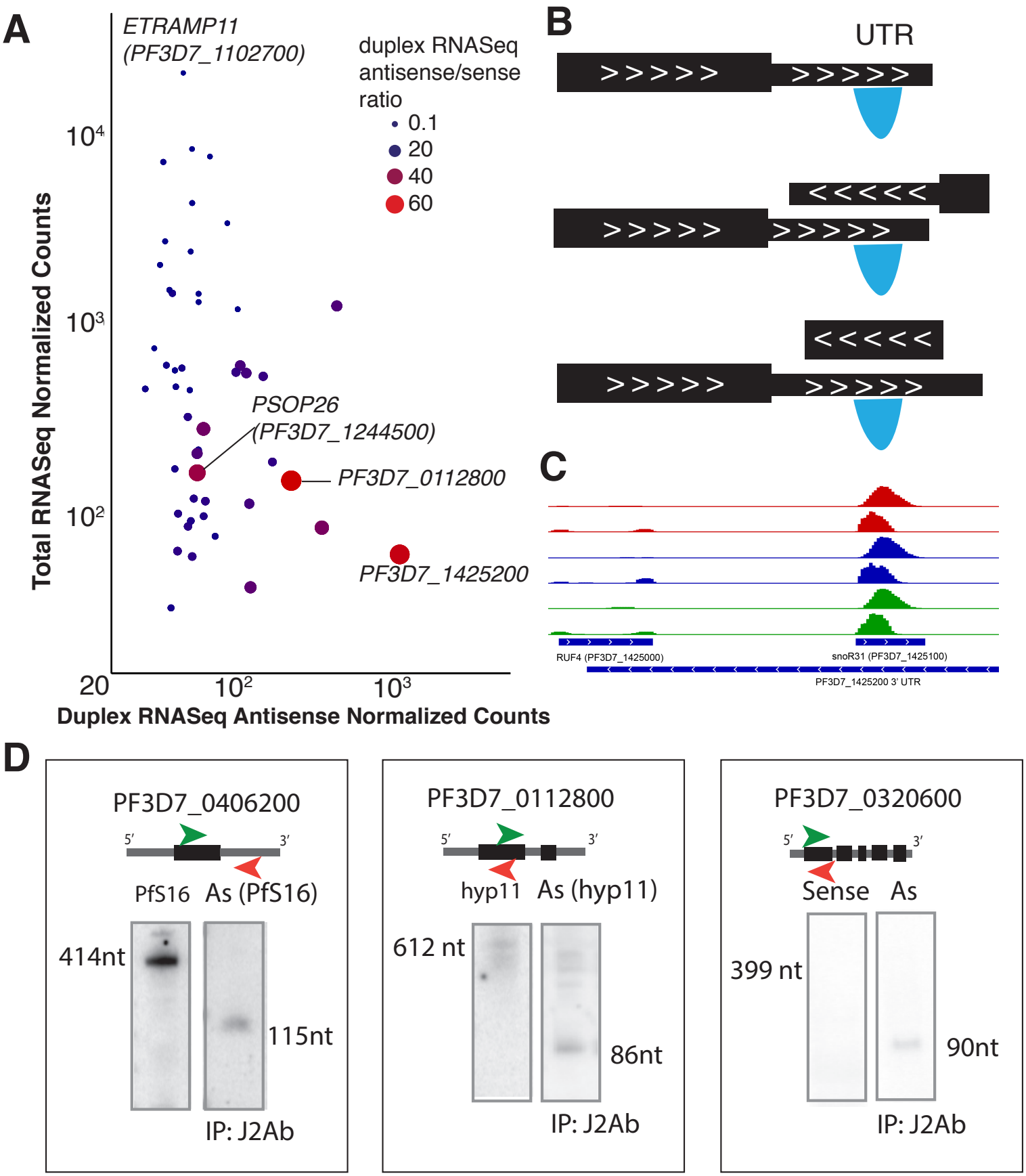

A
Total RNASeq Normalized Counts
Duplex RNASeq Normalized Counts
duplex RNASeq/
total RNASeq
ratio
2
4
6
ISP3(PF3D7_1460600)
QCR6(PF3D7_0622600)

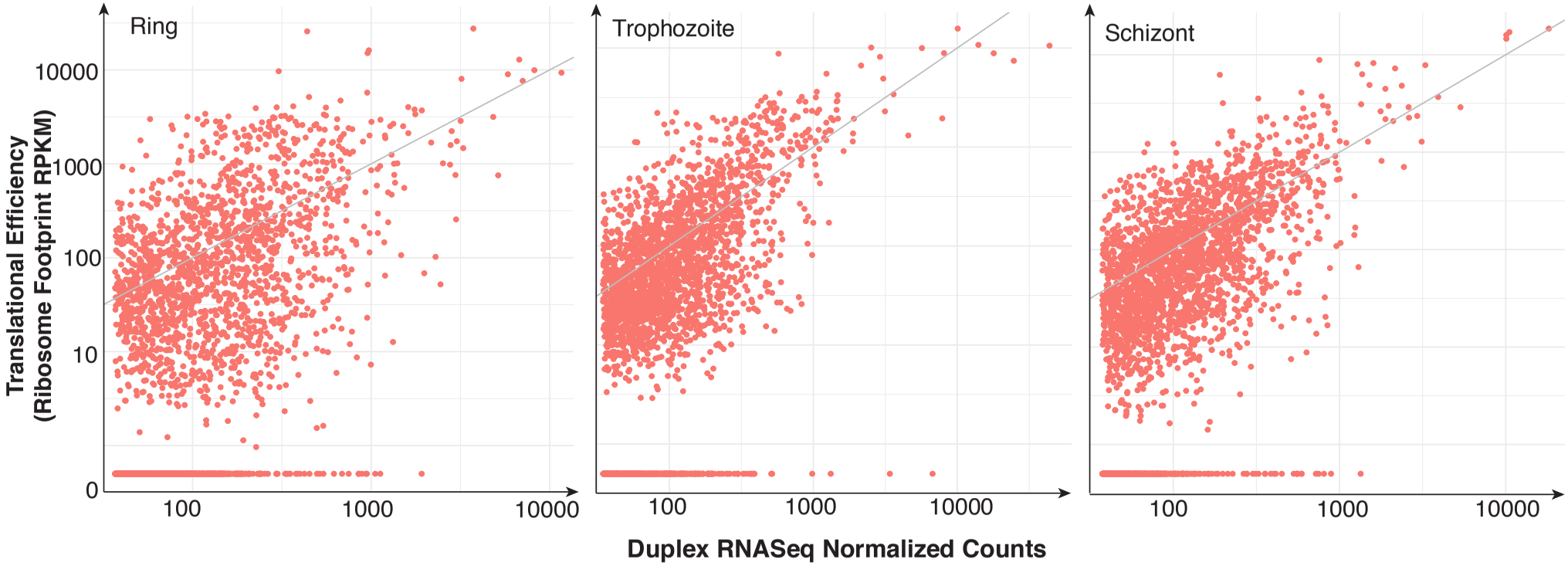

C
Stage
Ring
Troph
Schizont
Histone H2A (PF3D7_0617800)

D
bits


B
Ring
Trophozoite
Schizont
Stage
PF3D7_1434400
PF3D7_1401100 DnaJ Protein
PF3D7_0828800
PF3D7_1001500 ETRAMP10
PF3D7_1341300
PF3D7_0513300 PNP
PF3D7_0307100
PF3D7_1105400
PF3D7_0519200
PF3D7_0516900
PF3D7_1029600 ADA
PF3D7_0319600
PF3D7_0219200
PF3D7_0312800
PF3D7_0611700
PF3D7_0516200
PF3D7_1331800
PF3D7_0817900 HMGB2
PF3D7_0524800
PF3D7_1214900
PF3D7_0217800
PF3D7_0627500
PF3D7_0519400
PF3D7_1242700
PF3D7_1347200
PF3D7_0922500
PF3D7_1006200 Alba3
PF3D7_1408600
PF3D7_1149200 RESA3
PF3D7_1351700
PF3D7_0202500
PF3D7_1436300 PTEX150
PF3D7_1301600 EBA140
PF3D7_0220000
PF3D7_0104300
PF3D7_1238900
PF3D7_1228600
PF3D7_0525800
PF3D7_1205500 Zinc Finger
PF3D7_1222700
PF3D7_1035700
PF3D7_0817700
PF3D7_0217500
PF3D7_0935800 CLAG9
PF3D7_0831600
PF3D7_1252100
PF3D7_0104200
PF3D7_1251200
PF3D7_1409400
PF3D7_0207500
PF3D7_0407800
PF3D7_1223100
PF3D7_0604100 SIP2
PF3D7_0725400
PF3D7_1253000
PF3D7_0903300
PF3D7_1401600
PF3D7_1035400
PF3D7_1027300
PF3D7_1036000
PF3D7_1310700 RNA binding
PF3D7_0911100
PF3D7_0403800
PF3D7_0202100
PF3D7_0402300
PF3D7_1411000
PF3D7_0402100
PF3D7_1433500 TOP2
PF3D7_1028700
PF3D7_0707300
PF3D7_1452000
PF3D7_1143100 ApiAP2
PF3D7_0414900
PF3D7_1133400
PF3D7_1026600
PF3D7_1337800
PF3D7_0617800 Histone H2A
PF3D7_0610400
PF3D7_1105100 Histone H2B
PF3D7_0731500 EBA-175
PF3D7_1102800
PF3D7_0420300 ApiAP2
PF3D7_0102500
PF3D7_0302500 CLAG3.1
PF3D7_1246200
PF3D7_1410400
PF3D7_0207600
PF3D7_0930300
PF3D7_1361800
PF3D7_0929400
PF3D7_0905400
PF3D7_1462800
PF3D7_1324900
PF3D7_1109400
PF3D7_1424400
PF3D7_1424100
PF3D7_1015900
PF3D7_0826700
PF3D7_1012400 HGPRT
PF3D7_0903700
R1
R2
T1
T2
S1
S2
14
12
10
8
6
4


A
5' Gene End
Ring
Trophozite
Schizont
Normalized expression ratio at 5' gene end
40%
30%
20%
10%
0%
-50
0
+50
+100
-500
-300
-100
0
+100
B
3' Gene End
Normalized expression ratio at 3' gene end
-100
-50
0
+50
-100
0
+100
+300
+500
C
Normalized expression ratio
0.45
0.4
0.35
-50
-40
-30
-20
-10
0
5' exon
intron
0
+10
+20
+30
+40
+50
3' exon


A
ETRAMP11
(PF3D7_1102700)
duplex RNASeq
antisense/sense
ratio
0.1
20
40
60
PSOP26
(PF3D7_1244500)
PF3D7_0112800
PF3D7_1425200
Total RNASeq Normalized Counts
Duplex RNASeq Antisense Normalized Counts
B
UTR
C
RUF4 (PF3D7_1425000)
snoR31 (PF3D7_1425100)
PF3D7_1425200 3' UTR
D
PF3D7_0406200
PfS16
As (PfS16)
414nt
115nt
IP: J2Ab
PF3D7_0112800
hyp11
As (hyp11)
612 nt
86nt
IP: J2Ab
PF3D7_0320600
Sense
As
399 nt
90nt
IP: J2Ab


Ring
Trophozoite
Schizont
Translational Efficiency
(Ribosome Footprint RPKM)
10000
1000
100
10
0
100
1000
10000
Duplex RNASeq Normalized Counts

Supplementary Figures:

Supplementary Figure 1: Correlation scatter plots between the biological replicates of the three time point collections of *P. falciparum* after normalization of gene counts by sequencing depth (REC) (see Materials and Methods). The x-axis shows the normalized gene counts for replicates sequenced on HiSeq 2500 while the y-axis shows the normalized gene counts for replicates sequenced on NextSeq500. The axes are plotted in log10 scale and a pseudo count of one is added to all gene counts. The plots show all coding and non-coding genes (5386 genes). Both tRNA and rRNA reads were discarded prior to analysis. The genes are colored by type as indicated: *orange*: mRNA; *blue*: ncRNA; *pink*: snoRNA; *green*: snRNA. Pearson correlation R^2: 0.87, 0.63, 0.75 respectively.

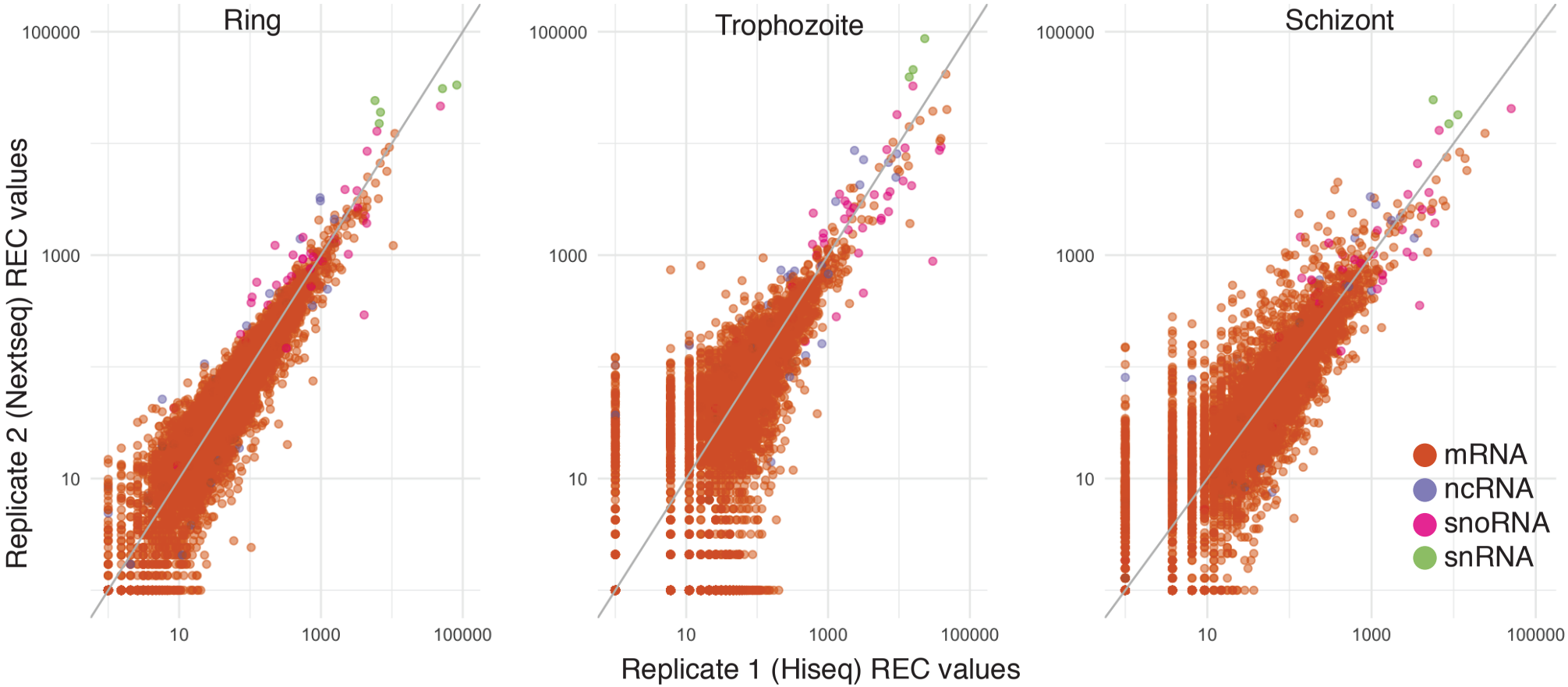
Supplementary Figure 2: Distribution of uniquely mapped reads mapping to annotated non-coding transcripts in the *P. falciparum* genome. Reads mapping to tRNA and rRNA genes were not included in this analysis. The x-axis shows the 14 chromosomes and the y-axis on the left shows the relative percentage of non-coding reads across each of the chromosomes. The error bars represent the standard deviation across the two replicates of the same stage. The y-axis on the right represents the number of non-coding genes annotated on each chromosome (Chappell et al. 2020). This is shown by the black dots and line on the plot. Colors represent the stage *red*: Ring; *blue*: Trophozoite; *green*: Schizont

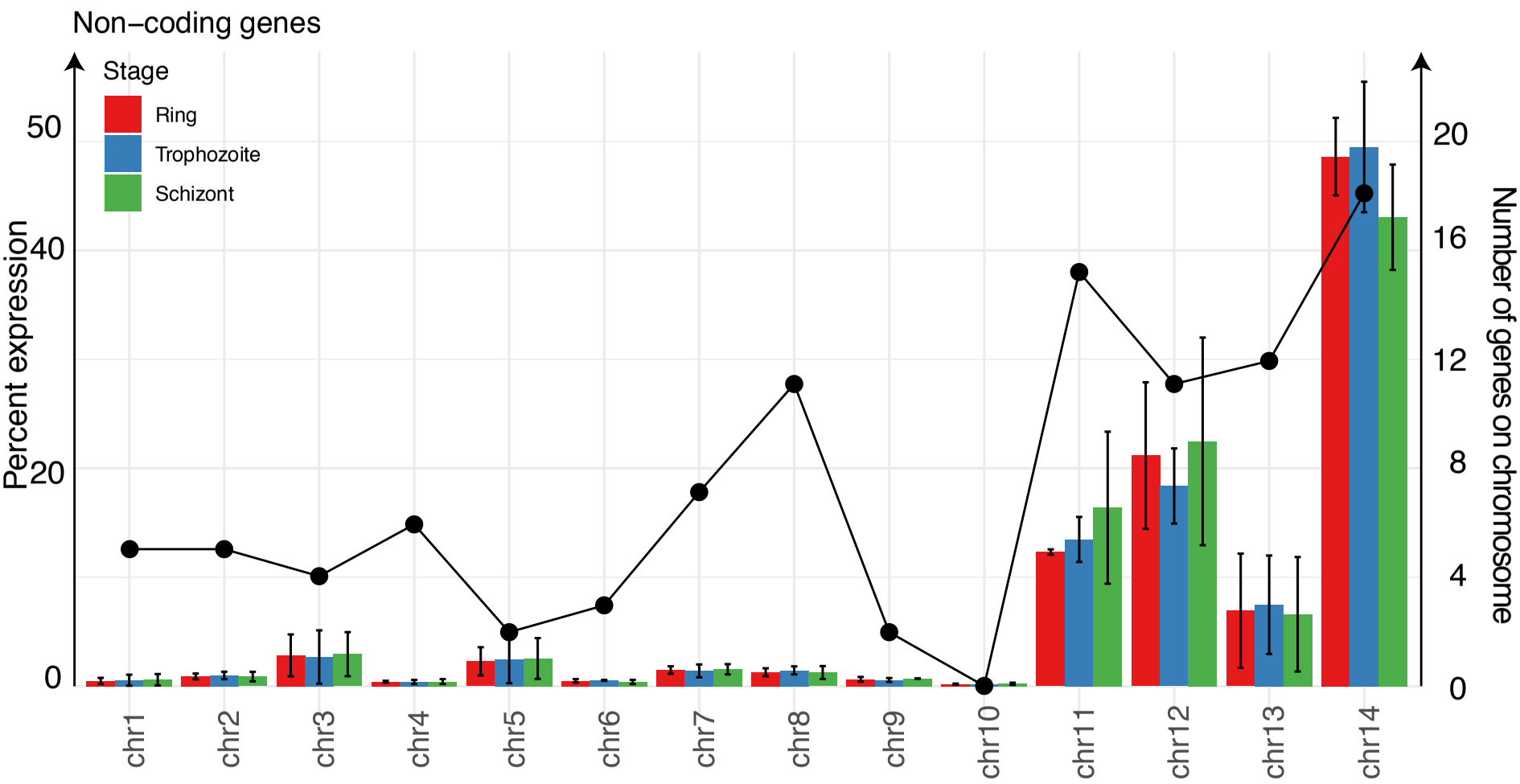
Supplementary Figure 3: Duplex RNA-seq data for snoR15. (A) Secondary structure of the PfsnoR15 based on in-gel SHAPE modification of mix stage RNA using NAI (B) IGV screenshot of the snoR15 peaks representing duplex RNA-seq read enrichments in three different P. falciparum RBC stages.

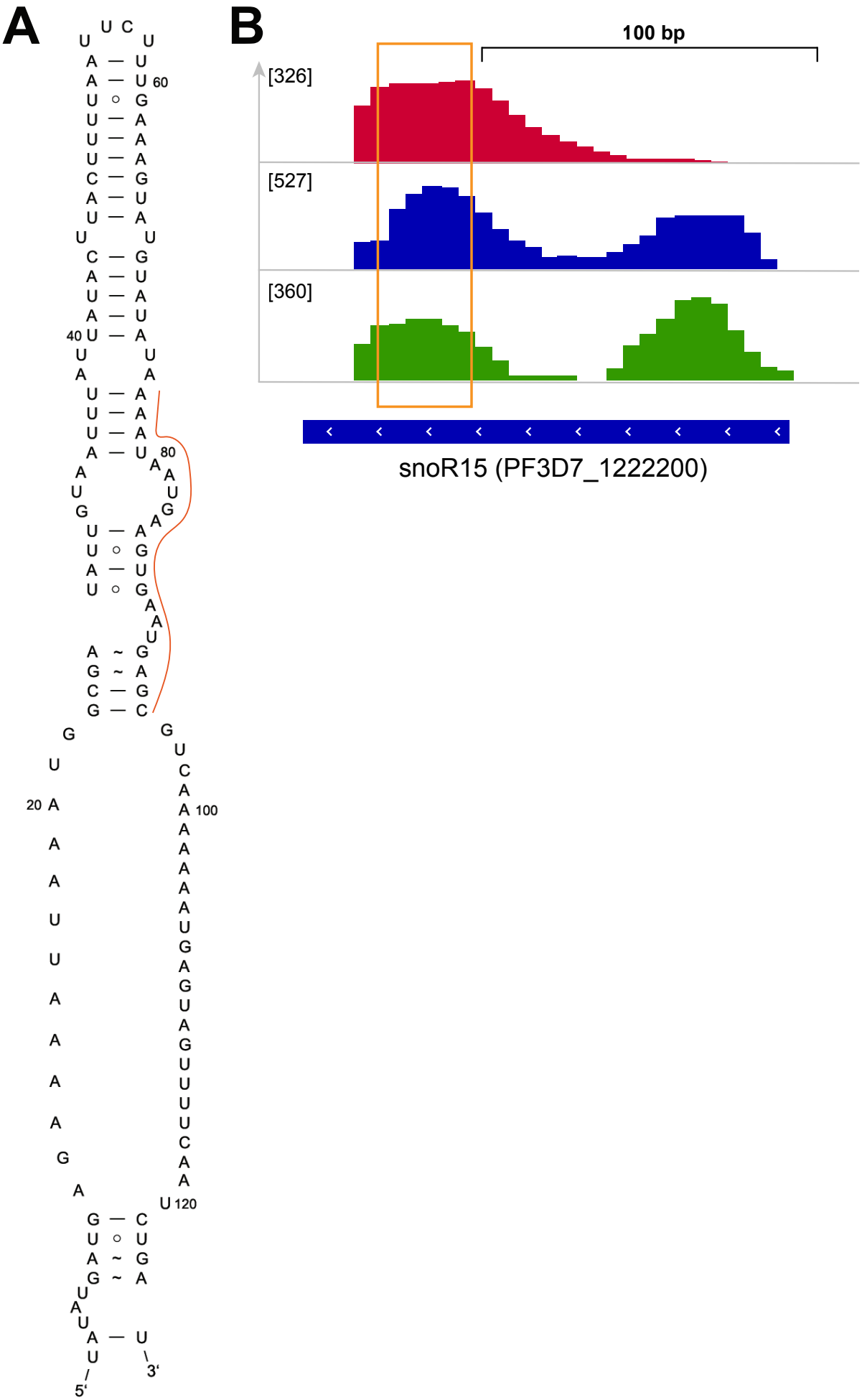
Supplementary Figure 4: Box plots illustrating the REC values (described in 'Materials and methods') of "expressed" coding and non-coding genes (3829 genes) (See Materials and Methods) across the three time point collections of *P. falciparum*. The asterisks indicate a significant difference (two-sided Wilcoxon test, $p < 0.001$).

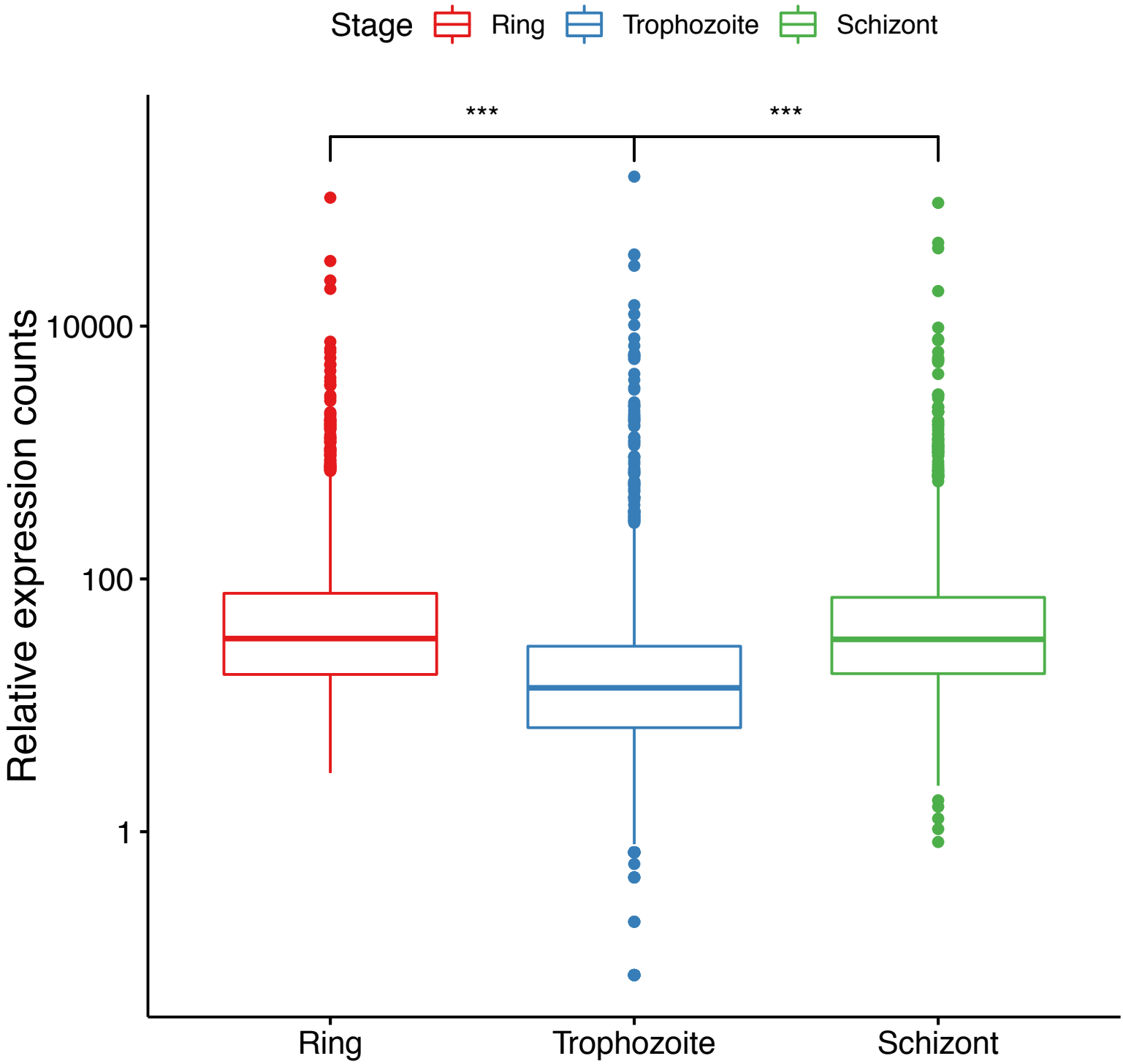
Supplementary Figure 5: (A-B) Scatter plots for Trophozoite (A) and (B) Schizont showing correlation between normalized gene counts from mRNA-seq and Duplex RNA-seq datasets for Ring stage. A pseudo count of 1.0 is added to the normalized counts for both duplex and mRNA-seq data. The axes are in logarithmic scale. The color and size of the points show the ratio of the duplex RNA-seq normalized counts by the mRNA-seq normalized counts. Larger red points represent genes with higher expression in duplex RNA-seq compared to total mRNA-seq as shown in the legend. (C-D) Antisense duplex RNA-seq counts for the top 50 antisense genes plotted against mRNA-seq counts for (C) Trophozoite and (D) Schizont stages. The size and color of the points indicate the ratio of the duplex RNA-seq antisense to sense counts. Larger red points indicate genes that have more antisense signal than sense signal. Few selected genes with very high and very low ratios are highlighted. A pseudo count of one is added to the normalized counts and axes are in logarithmic scale.

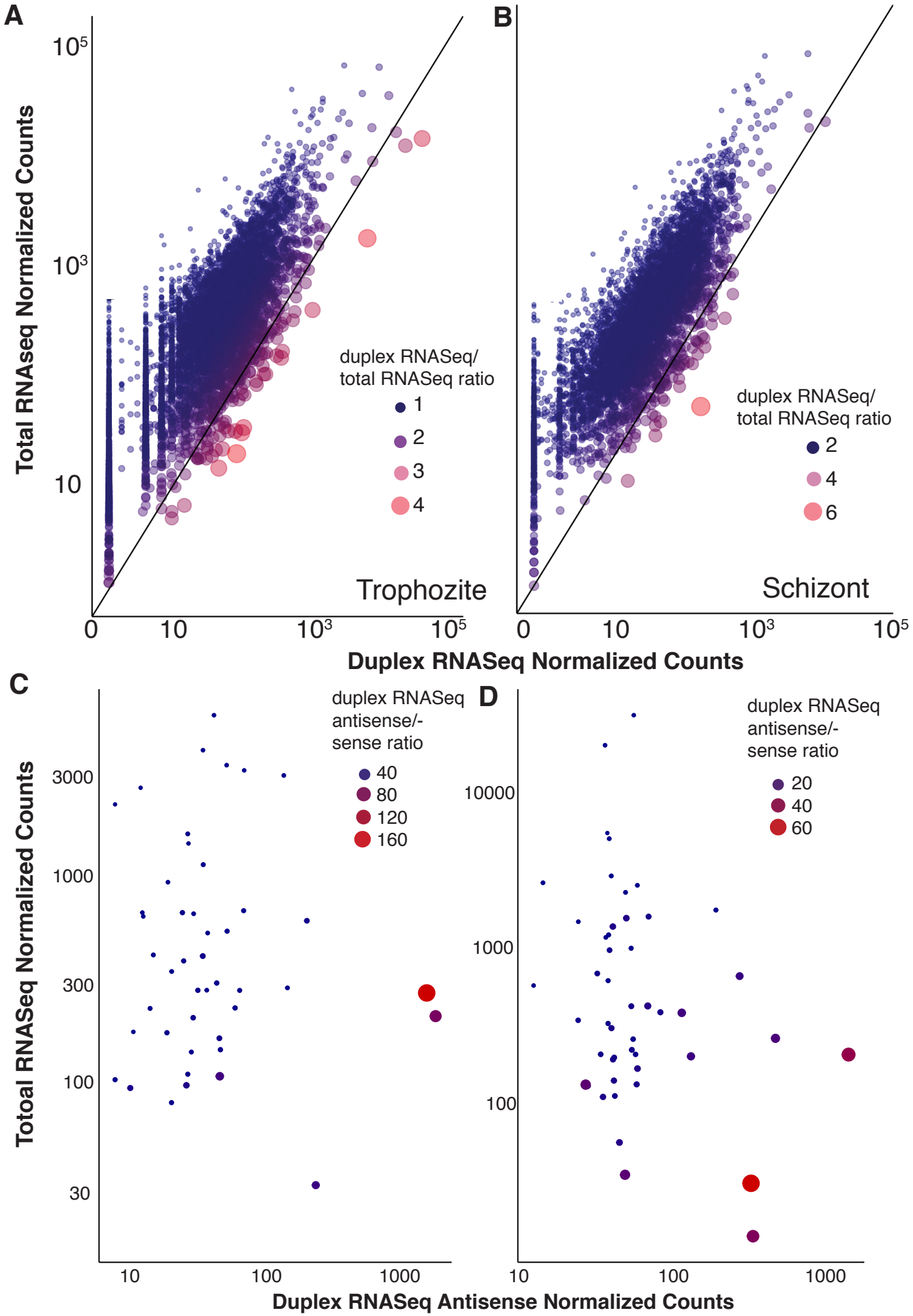
Supplementary Figure 6: Box plots illustrating the lengths of the 20% most expressed values of “expressed” coding and non-coding genes (3829 genes) (See Materials and Methods) across the three time point collections of *P. falciparum*. The asterisks indicate a significant difference (two-sided Wilcoxon test, $p < 0.001$).


Ring
Trophozoite
Schizont
100000
1000
10
Replicate 2 (Nextseq) REC values
Replicate 1 (Hiseq) REC values
mRNA
ncRNA
snoRNA
snRNA


Non−coding genes
Stage
Ring
Trophozoite
Schizont
Percent expression
0
20
40
50
Number of genes on chromosome
0
4
8
12
16
20
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14


A
B
100 bp
[326]
[527]
[360]
snoR15 (PF3D7_1222200)
5'
3'


Stage
Ring
Trophozoite
Schizont
***
***
Relative expression counts
10000
100
1
Ring
Trophozoite
Schizont


A
B
C
D
Total RNAseq Normalized Counts
Duplex RNASeq Normalized Counts
duplex RNASeq/
total RNASeq ratio
1
2
3
4
Trophozite
duplex RNASeq/
total RNASeq ratio
2
4
6
Schizont
Totoal RNASeq Normalized Counts
Duplex RNASeq Antisense Normalized Counts
duplex RNASeq
antisense/-
sense ratio
40
80
120
160
duplex RNASeq
antisense/-
sense ratio
20
40
60

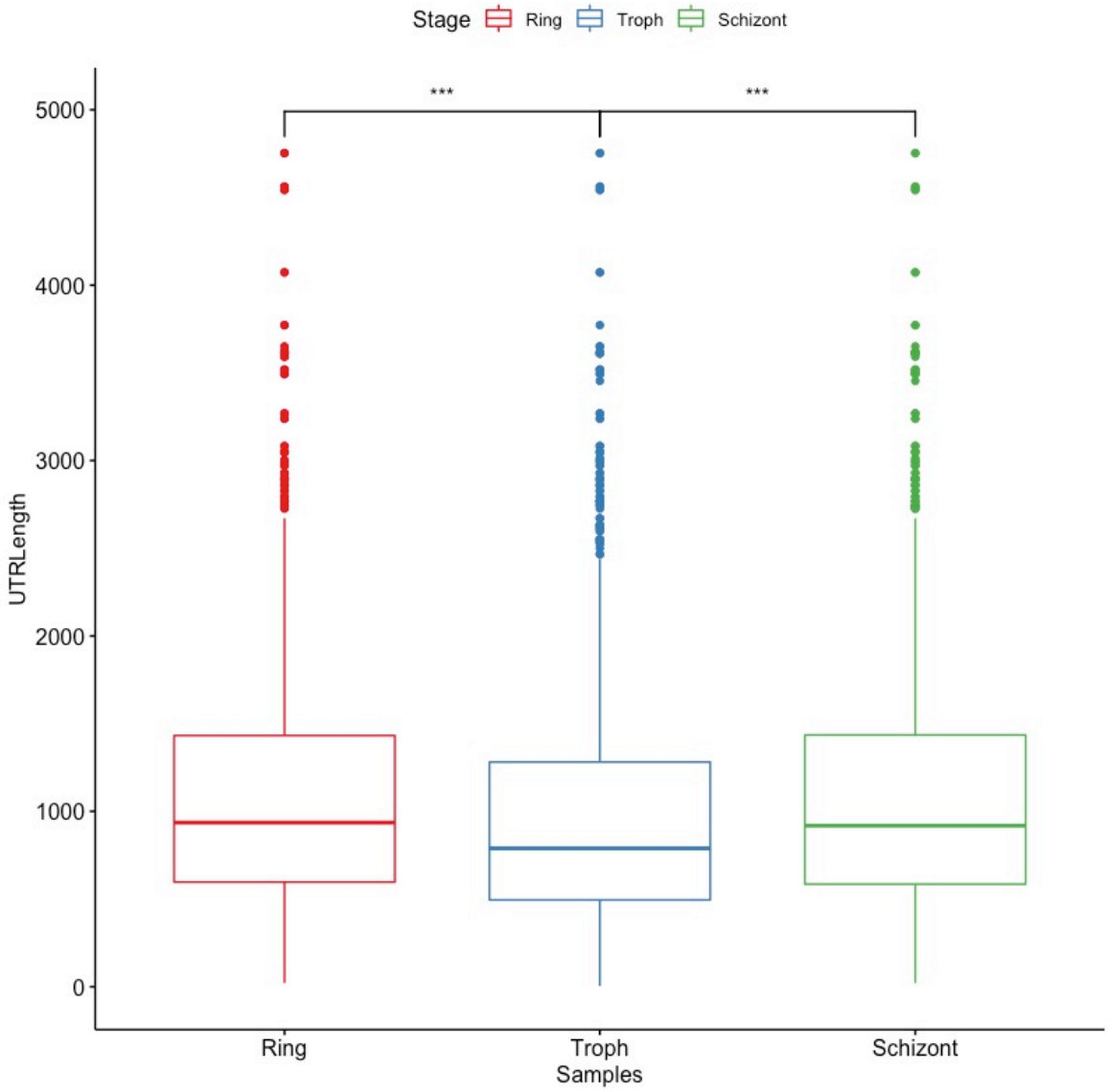

Stage
Ring
Troph
Schizont
***
***
5000
4000
3000
2000
1000
0
UTRLength
Ring
Troph
Schizont
Samples

Table S1: HPLC purified 3' Biotinylated oligos (3Bio) used for the depletion of *Plasmodium falciparum* ribosomal RNA (rRNA) species.

***Plasmodium falciparum* rRNA**	**Biotinylated oligos**	**Melting temperature (Tm)**
5S rRNA	TGACTCGTTCATACTACAGTGGCCA/3Bio/	59.4°C
5.8S rRNA	CCAAGACATCCATCGTTAAGACTGTGC/3Bio/	59.4°C
18S rRNA	TGTACTTTCACTTGCATGGCTTAATCTGTG/3Bio/	59.1°C
28S rRNA	CCTTCTTCACTCGCCGTTACTAAAGG/3Bio/	59.0°C

Table S2. Summary of sequence reads obtained from duplex RNA seq libraries. The %total bases Q>=30 and GC content are calculated post-trimming adapter sequences.

Library	Read length	Total reads	%Total Bases Q≥30	GC Content (%)	Median Insert Size
Duplex RNA-seq-1					
Ring	125X125	67,970,256	75.89	42.25	45
Trophozoite	125X125	17,691,860	75.03	42.82	49
Schizont	105X105	10,201,818	81.45	43.43	45
Duplex RNA-seq-2					
Ring	75X75	178,562,120	88.76	51.49	30
Trophozoite	75X75	437,153,310	88.32	50.41	33
Schizont	75X75	224,498,038	89.64	51.27	30

Table S3 Normalized gene counts of 73 highly base-paired segments of known, conserved structural RNAs

Gene	GeneType	Ring Replicate	Ring Replicate	Trophozoite R	Trophozoite R	Schizont Rep	Schizont Rep	Average
PF3D7_1418800	ncRNA	234564.23	136722.93	828810.50	455948.14	221959.70	133108.54	335185.67
PF3D7_1213100	snRNA	81252.92	33711.32	208307.64	101793.57	135634.23	34747.31	99241.17
PF3D7_1137000	snRNA	51065.79	31306.55	193819.90	114183.66	122142.25	31934.64	90742.13
PF3D7_1244300	snoRNA	47740.37	21830.53	172635.24	79335.42	50379.06	20534.10	65409.12
PF3D7_1341100	snRNA	5704.78	24458.35	22773.52	87886.71	5646.71	24640.64	28518.45
PF3D7_1424900	snoRNA	4352.35	1950.79	107670.57	12136.38	5930.84	1940.23	22330.19
PF3D7_1446000	snRNA	6824.10	19245.38	13755.18	39836.49	11309.98	18072.34	18173.91
PF3D7_0308400	snRNA	6545.85	15246.26	15591.52	46321.65	8766.61	15011.14	17913.84
PF3D7_0510900	snoRNA	6076.49	12998.20	15522.22	33053.49	6675.65	13119.25	14574.21
PF3D7_1109000	snoRNA	3264.52	2664.40	38622.42	9447.00	4154.35	2568.97	10120.28
PF3D7_1341400	snoRNA	4159.67	2271.02	36701.94	8782.93	5376.38	2472.00	9960.66
PF3D7_1424500	snoRNA	4460.50	8771.37	9369.76	18822.31	3666.09	6799.49	8648.25
PF3D7_1424600	snoRNA	4531.38	443.59	37083.07	1124.07	4482.61	528.64	8032.22
PF3D7_1304800	snoRNA	2153.07	3911.97	12007.94	9228.59	2783.36	3504.53	5598.25
PF3D7_0707420	snoRNA	3183.15	3815.51	6647.43	8918.62	5015.01	3643.30	5203.84
PF3D7_0218800	ncRNA	2273.83	2212.92	9171.77	8164.09	2248.20	2190.76	4376.93
PF3D7_1109600	snoRNA	707.72	1049.62	14854.02	4239.24	2711.64	1064.38	4104.44
PF3D7_0826800	snoRNA	1602.34	1306.74	11280.34	4674.97	2375.10	1688.85	3821.39
PF3D7_0918500	ncRNA	1494.18	1308.89	7003.81	6846.98	3315.75	1422.12	3565.29
PF3D7_1328700	ncRNA	958.15	3306.30	2326.35	8760.87	968.24	3338.75	3276.44
PF3D7_0104900	ncRNA	966.55	3097.59	3123.25	7239.69	1128.24	2843.38	3066.45
PF3D7_1241000	ncRNA	743.42	357.17	8919.34	5033.48	1001.35	479.44	2755.70
PF3D7_1109300	snoRNA	383.78	696.04	8949.04	4201.73	571.02	663.71	2577.55
PF3D7_1109500	snoRNA	2401.40	1028.83	5513.95	2074.94	3216.45	973.10	2534.78
PF3D7_1424800	ncRNA	1522.53	2120.41	2766.88	4310.94	1754.43	2050.85	2421.01
PF3D7_1130310	snoRNA	1551.93	1515.80	4553.71	3878.52	1199.96	1815.40	2419.22
PF3D7_1421150	snoRNA	1040.57	889.69	6800.87	3419.63	1362.71	675.94	2364.90
PF3D7_0601800	snoRNA	925.07	542.92	7365.14	3745.05	617.91	570.72	2294.47
PF3D7_1423900	snoRNA	717.69	532.52	5538.70	2168.71	1177.89	498.21	1772.29
PF3D7_1425100	snoRNA	764.42	973.96	2252.11	2667.31	1062.03	1021.44	1456.88
PF3D7_0218900	ncRNA	501.39	1413.96	1262.17	3057.81	626.19	1426.38	1381.32
PF3D7_0804100	snoRNA	550.74	1460.94	1707.64	3092.01	444.12	1006.94	1377.07
PF3D7_1351500	snoRNA	398.48	1014.48	1430.46	3563.03	300.68	1286.19	1332.22
PF3D7_1424000	snoRNA	547.59	935.23	3029.21	1771.59	623.43	913.10	1303.36
PF3D7_1130300	snoRNA	242.03	495.23	4316.13	1405.36	311.71	570.15	1223.44
PF3D7_1424300	snoRNA	695.64	759.87	1895.73	1705.40	1370.99	592.62	1170.04

PF3D7_1331900	snoRNA	546.01	934.15	1682.89	2138.92	744.80	845.71	1148.75
PF3D7_1402600	snoRNA	378.01	652.65	2029.37	2436.76	444.12	740.49	1113.57
PF3D7_1424200	snoRNA	121.80	577.71	1851.18	2863.66	184.82	601.15	1033.39
PF3D7_1109800	snoRNA	221.03	1236.81	608.81	2405.88	135.17	1453.97	1010.28
PF3D7_0219000	snoRNA	230.48	291.54	3103.45	462.20	366.88	271.85	787.74
PF3D7_1478900	ncRNA	1218.03	500.61	994.89	686.13	526.88	522.38	741.49
PF3D7_0709050	snoRNA	232.06	545.07	871.15	1288.43	463.43	541.43	656.93
PF3D7_0307000	snoRNA	309.23	440.72	841.45	1440.66	391.71	323.89	624.61
PF3D7_0309400	snoRNA	101.33	378.68	846.40	1586.27	140.69	623.90	612.88
PF3D7_1309120	ncRNA	322.36	376.17	796.90	808.58	491.02	528.64	553.94
PF3D7_0803900	snoRNA	104.48	425.30	593.96	1266.37	201.37	424.84	502.72
PF3D7_1210800	snoRNA	275.11	285.09	678.11	1004.93	231.72	366.55	473.58
PF3D7_1424700	snoRNA	320.26	147.38	1286.92	283.50	424.81	137.92	433.47
PF3D7_0804200	ncRNA	255.68	469.41	331.63	729.15	226.20	529.77	423.64
PF3D7_0626840	ncRNA	185.33	457.22	292.03	678.41	231.72	600.30	407.50
PF3D7_1425000	ncRNA	87.15	234.88	212.84	742.39	132.41	247.68	276.23
PF3D7_1222200	snoRNA	85.58	169.62	306.88	523.98	80.00	90.43	209.41
PF3D7_0809000	ncRNA	28.88	59.17	262.33	642.01	22.07	116.87	188.55
PF3D7_1336500	snoRNA	71.40	196.87	311.83	275.78	74.48	184.55	185.82
PF3D7_0712700	ncRNA	68.78	17.93	806.80	162.16	27.59	15.64	183.15
PF3D7_1330900	snoRNA	86.63	88.22	470.22	169.88	60.69	28.15	150.63
PF3D7_0829900	ncRNA	219.45	153.84	123.74	38.61	66.20	151.28	125.52
PF3D7_0531500	ncRNA	21.53	106.50	138.59	359.61	27.59	96.68	125.08
PF3D7_0808500	ncRNA	51.45	21.16	475.17	126.86	30.34	28.44	122.24
PF3D7_0712100	ncRNA	48.83	53.43	287.08	233.86	66.20	31.00	120.07
PF3D7_0829800	ncRNA	154.35	170.69	153.44	13.24	77.24	112.61	113.60
PF3D7_0105000	ncRNA	69.30	62.40	202.94	110.31	74.48	51.47	95.15
PF3D7_0711800	ncRNA	58.80	22.23	306.88	109.21	60.69	6.54	94.06
PF3D7_0830100	ncRNA	43.58	59.89	24.75	99.28	85.51	72.23	64.21
PF3D7_0421400	ncRNA	14.70	17.21	59.40	258.13	13.79	6.26	61.58
PF3D7_1370500	ncRNA	26.78	8.25	84.14	147.82	24.83	34.12	54.32
PF3D7_0108900	ncRNA	44.63	73.51	19.80	112.52	5.52	66.54	53.75
PF3D7_1109700	snoRNA	7.35	41.96	49.50	168.78	11.03	40.10	53.12
PF3D7_1421700	ncRNA	4.73	50.92	9.90	156.64	0.00	79.62	50.30
PF3D7_1363900	ncRNA	24.68	28.69	44.55	110.31	60.69	27.58	49.42
PF3D7_0731000	ncRNA	31.50	77.46	14.85	43.02	5.52	76.21	41.43
PF3D7_0809300	ncRNA	26.78	16.14	24.75	0.00	68.96	28.72	27.56