

CORRECTION

Open Access



# Correction: A manually curated annotation characterises genomic features of *P. falciparum* lncRNAs

Johanna Hoshizaki<sup>1</sup>, Sophie H. Adjalley<sup>1,2</sup>, Vandana Thathy<sup>3,4</sup>, Kim Judge<sup>1</sup>, Matthew Berriman<sup>1,5</sup>, Adam J. Reid<sup>1,6</sup> and Marcus C. S. Lee<sup>1\*</sup>

**Correction:** *BMC Genomics* 23, 780 (2022)  
<https://doi.org/10.1186/s12864-022-09017-2>

Following the publication of the original article [1], the authors detected a typo under the heading **Some lncRNAs contain structural RNA sequences**. The correction has been marked in **bold**.

### **Some lncRNAs contain structural RNA sequences**

Searches against the RNA families database (Rfam) revealed that 19 lncRNAs contained sequences associated with 22 **unique** described RNA families (Fig. 5A, Additional File 1: Supp. Table 4), including those encoding known structural RNAs such as the signal recognition particle RNA, the ribozyme ribonuclease P and several RNAs of unknown function (RUFs) [41].

Moreover, it was noted that the Figs. 1, 2, 3, 5 and 6 and their legends were not updated. The correct figures are included here. The corrections in the legends were marked in **bold**.

The original article [1] has been updated.

Published online: 06 April 2023

### Reference

1. Hoshizaki J, Adjalley SH, Thathy V, et al. A manually curated annotation characterises genomic features of *P. falciparum* lncRNAs. *BMC Genomics*. 2022;23:780. <https://doi.org/10.1186/s12864-022-09017-2>.

The original article can be found online at <https://doi.org/10.1186/s12864-022-09017-2>.

\*Correspondence:

Marcus C. S. Lee  
ml31@sanger.ac.uk

<sup>1</sup> Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK

<sup>2</sup> Micrographia Bio, London W12 0BZ, UK

<sup>3</sup> MRC Weatherall Institute of Molecular Medicine, University of Oxford, Oxford OX3 9DS, UK

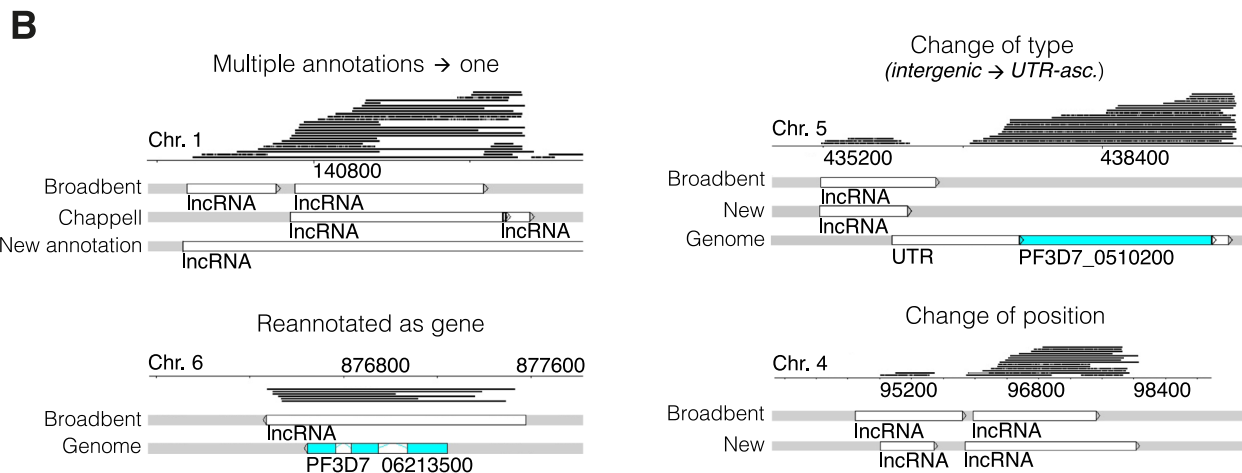
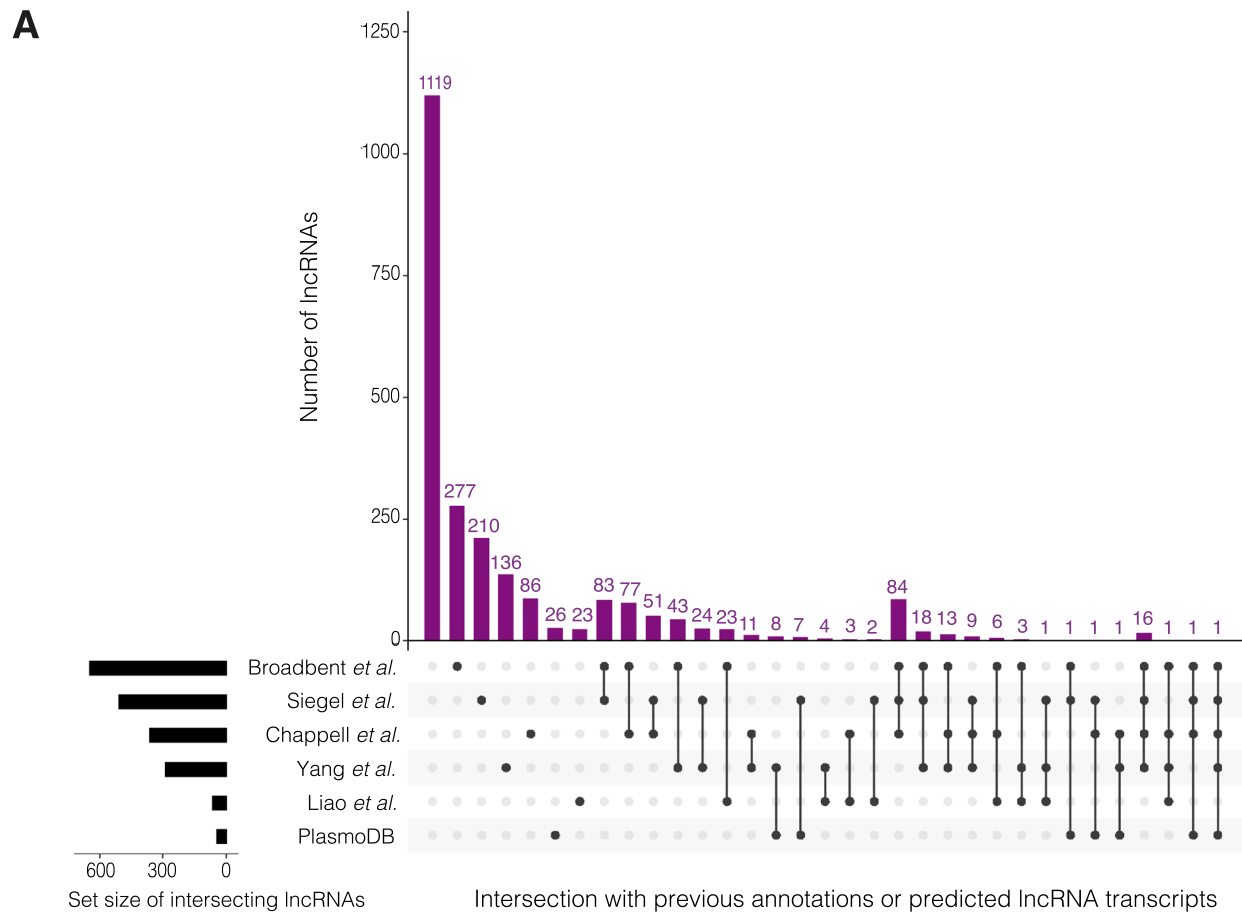
<sup>4</sup> Present address: Department of Microbiology and Immunology, Columbia University Medical Center, New York NY10032, USA

<sup>5</sup> Wellcome Centre for Integrative Parasitology, University of Glasgow, Glasgow G12 8TA, UK

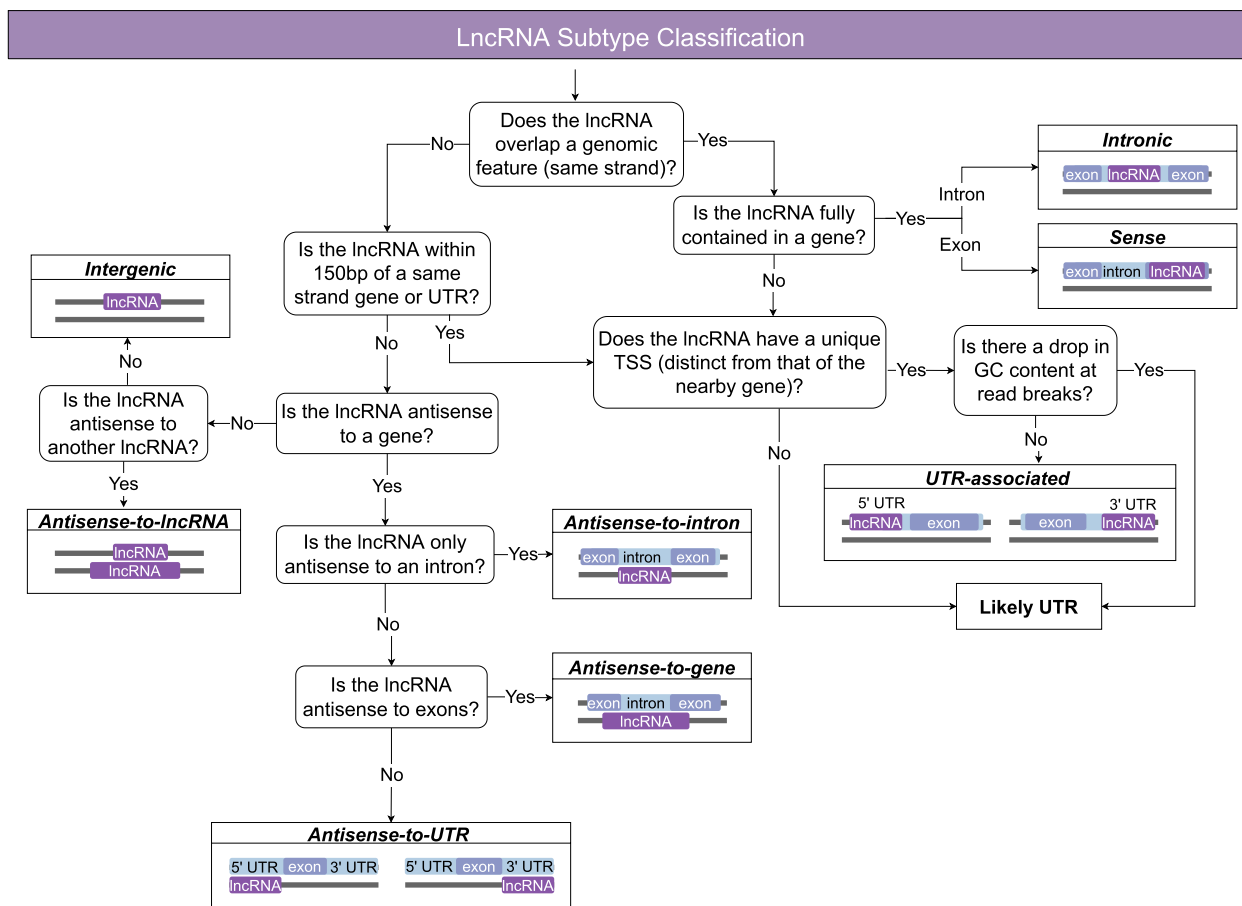
<sup>6</sup> Present address: Wellcome/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK



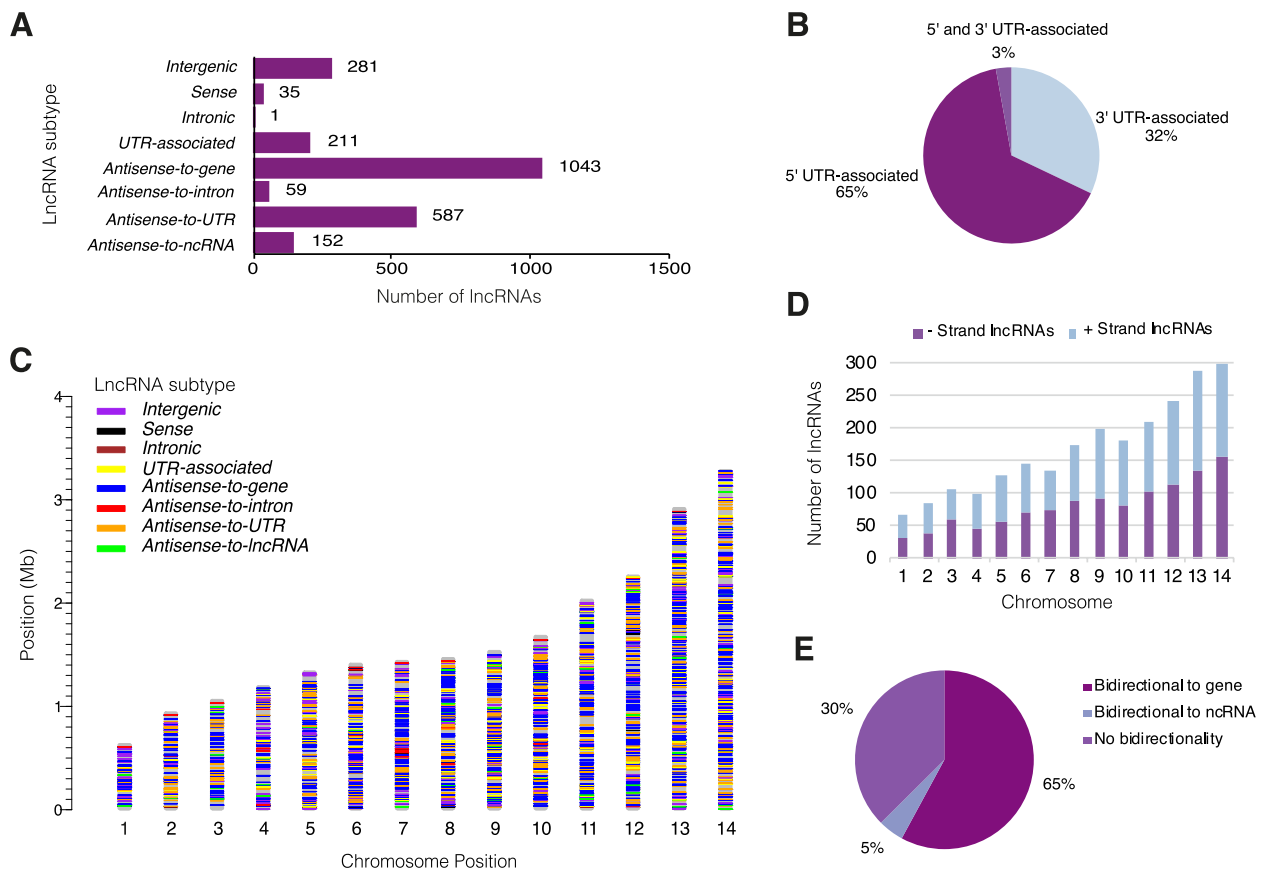
© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



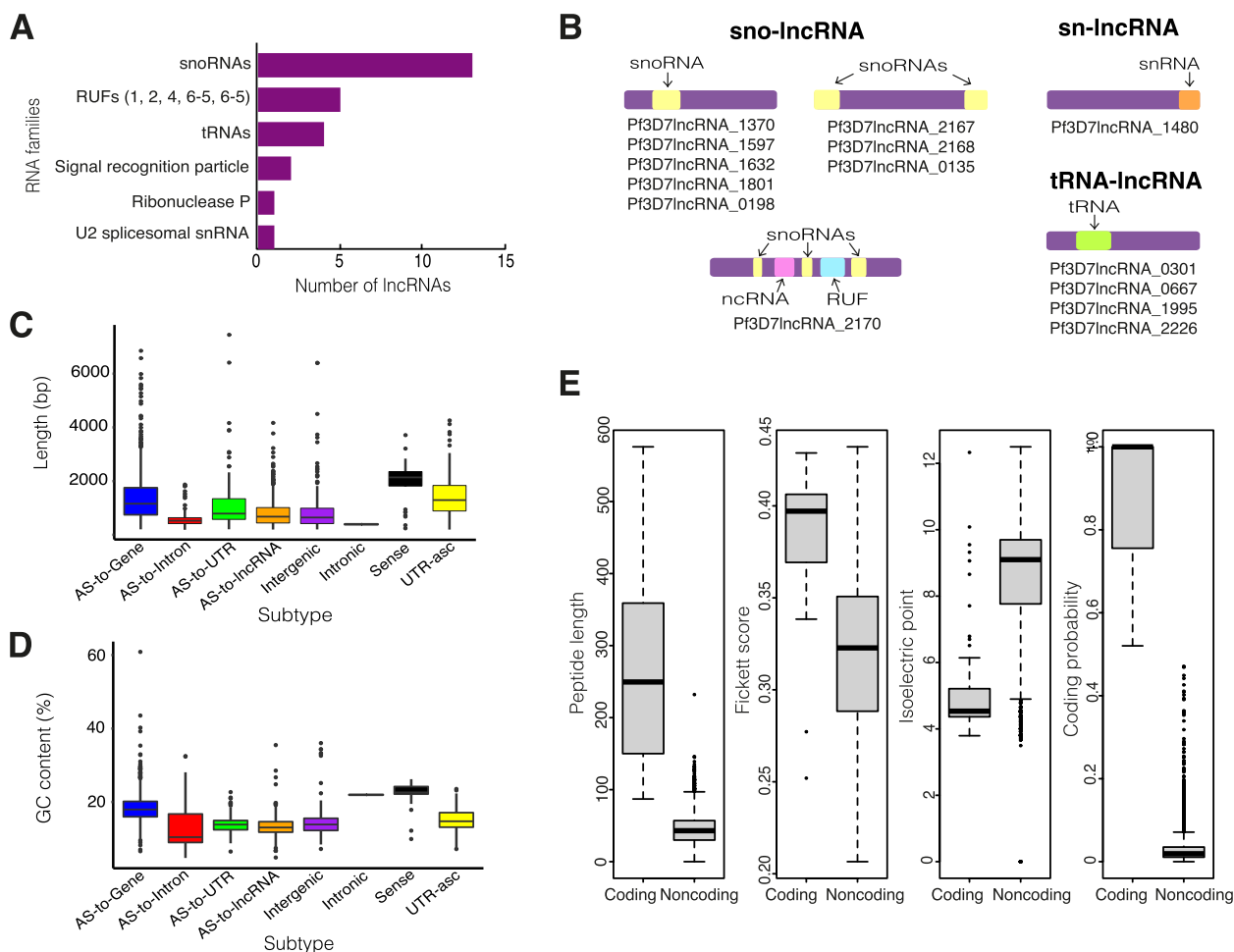
**Fig. 1** Verification of previous *P. falciparum* lncRNA annotations in the literature. Of the 2369 lncRNAs, 1119 were unique to this study and 1250 were previously annotated by Broadbent et al. or Liao et al., or predicted by Siegel et al., Chappell et al., Yang et al., and/or listed on PlasmoDB from various published studies [14, 18, 21, 31, 32, 37]. **a** An upset plot shows the number and membership of the previously annotated lncRNAs as well as the size of the set. Gene IDs from the Siegel et al. dataset were intersected with gene IDs of genes antisense to lncRNAs in this work [18]. **b** Snapshots demonstrate examples of changes to previous annotations



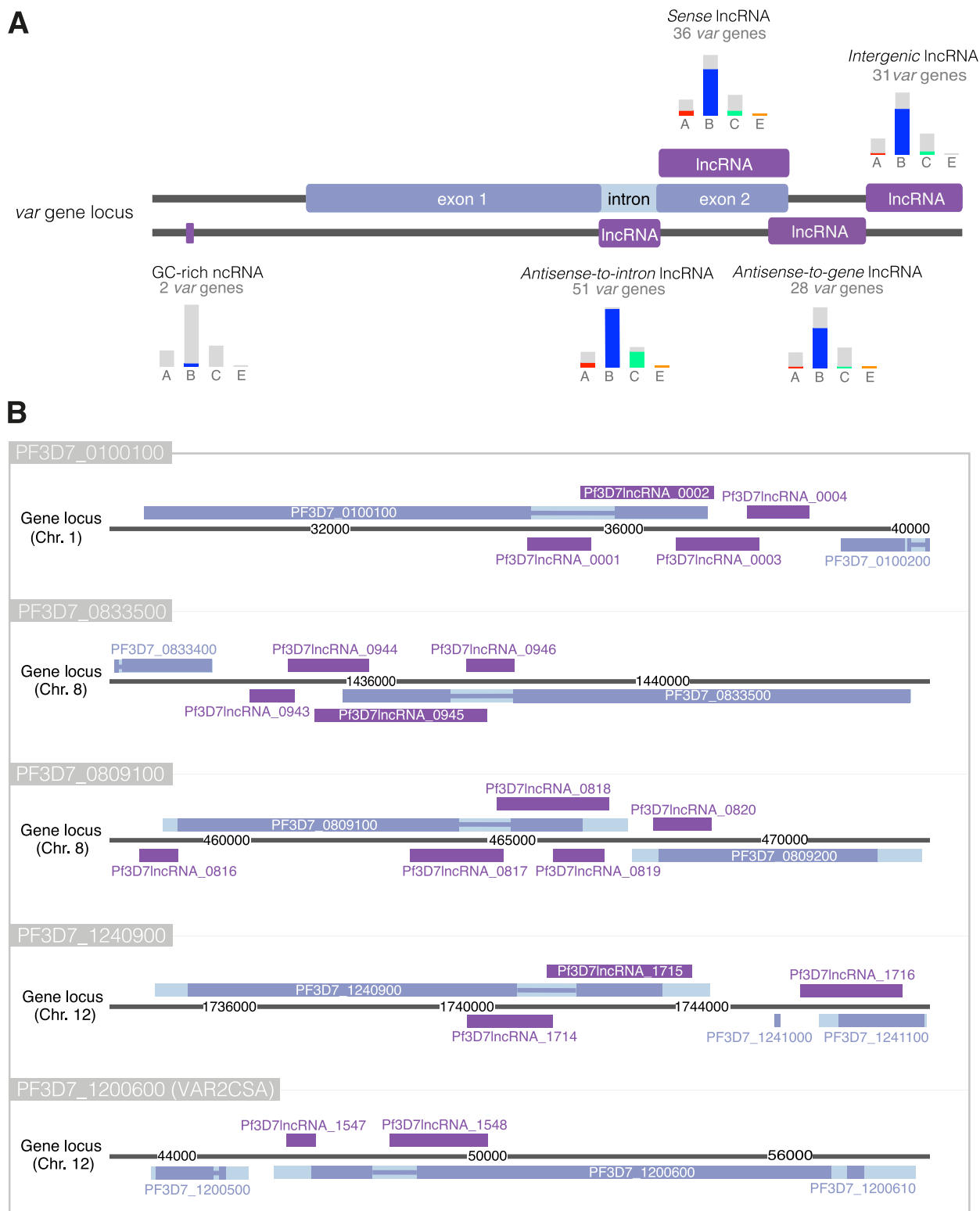
**Fig. 2** Schematic representation of the classification of lncRNA into genome context-based subtypes. Annotations were categorised by genomic context using a decision tree. lncRNAs that overlapped a gene on the same strand were classified as either intronic if contained within the intron or sense if contained within a single exon. No lncRNAs were annotated that spanned multiple exons in a gene. lncRNAs that overlapped a UTR and lncRNAs nearby genes (within 150bp of an annotated UTR or exon or read from the gene) were flagged as potential UTR-associated lncRNAs. To delineate UTR-associated lncRNAs from UTR transcripts (that could be fragmented due to drops in GC content or alternative start sites) careful examination of collative data was performed. This included an analysis of the level of overlap between reads from the putative lncRNA and gene/UTR, the presence of a unique transcriptional start site (distinct from the gene) and the lack of evidence of a drop in GC content. lncRNAs that were antisense (opposite strand) to genomic features were classified based on the type of antisense genomic feature: antisense-to-gene, antisense-to-intron, antisense-to-UTR and antisense-to-lncRNA. The antisense-to-intron lncRNAs were contained within the intron boundaries (with little to no overlap with the exon). The antisense-to-UTR lncRNAs only overlapped the UTR, not the exons and the level of overlap varied. Some lncRNAs could be classified as multiple subtypes if overlapping multiple features – the classification has a hierarchy starting with: intronic, sense, UTR-associated, antisense-to-intron, antisense-to-gene, antisense-to-UTR and antisense-to-lncRNA. lncRNAs not overlapping, antisense to, or nearby (150bp) any feature were classified as intergenic



**Fig. 3** Genomic features of *P. falciparum* lncRNAs. **a** The majority of lncRNAs were antisense-to-gene lncRNAs or antisense-to-UTR lncRNAs, followed by intergenic lncRNAs, UTR-associated lncRNAs and antisense-to-ncRNA lncRNAs. A minority were antisense-to-intron, intronic or sense lncRNAs. **b** Antisense-to-UTR lncRNAs were most commonly (65%) associated with a 5' UTR, followed by 32% associated with a 3' UTR and a small subset (3%) were nestled between 5' and 3' UTRs. **c** lncRNAs and their genome-context subtypes were distributed throughout the genome. **d** lncRNAs were equally distributed between positive and negative strands and their abundance in chromosomes was relative to chromosome size. **e** For lncRNAs with an associated transcriptional start site, 70% had evidence of bidirectionality (transcription in both directions in the same location and time point)



**Fig. 5** Sequence features of *P. falciparum* lncRNAs. **a** 26 RNA families (22 unique) were identified that aligned to annotated lncRNA sequences: 13 snoRNAs, 5 RUFs, 4 tRNAs, ribonuclease P, 2 signal recognition peptides and U6 spliceosomal snRNA using Rfam [39]. **b** Visual representation and IDs of lncRNAs containing snoRNAs, snRNAs and tRNAs. **c** lncRNA lengths ranged from 200 to 7452 bp and the distribution of lengths differed between subtypes (AS:antisense to). **d** lncRNAs were AT-rich with an average GC content of 15.97% and the distribution of GC content differed between subtypes (AS:antisense to). **e** 16 lncRNAs were identified as putatively coding by sequence feature-dependent coding potential analysis with CPC2 [40]. The distributions of the measures used to determine coding potential (peptide length, isoelectric point and Fickett score) are presented for putative coding and noncoding lncRNA annotations



**Fig. 6** lncRNAs present at var gene loci. **a** Schematic of lncRNAs found at var gene loci and the number of loci containing each subtype in this study. Bar charts denote the distribution of var gene subtypes for these loci by colour (A-red, B-blue, C-green and E-orange). The GC-rich (RUF6) family, antisense-to-intron and sense lncRNAs are well-established var-associated lncRNAs. Two additional lncRNAs; a downstream intergenic lncRNA and an antisense-to-gene lncRNA were observed in this study. **b** Examples of lncRNAs at five var gene loci