
Figures and figure supplements

A coordinated transcriptional switching network mediates antigenic variation of human malaria parasites

Xu Zhang *et al.*

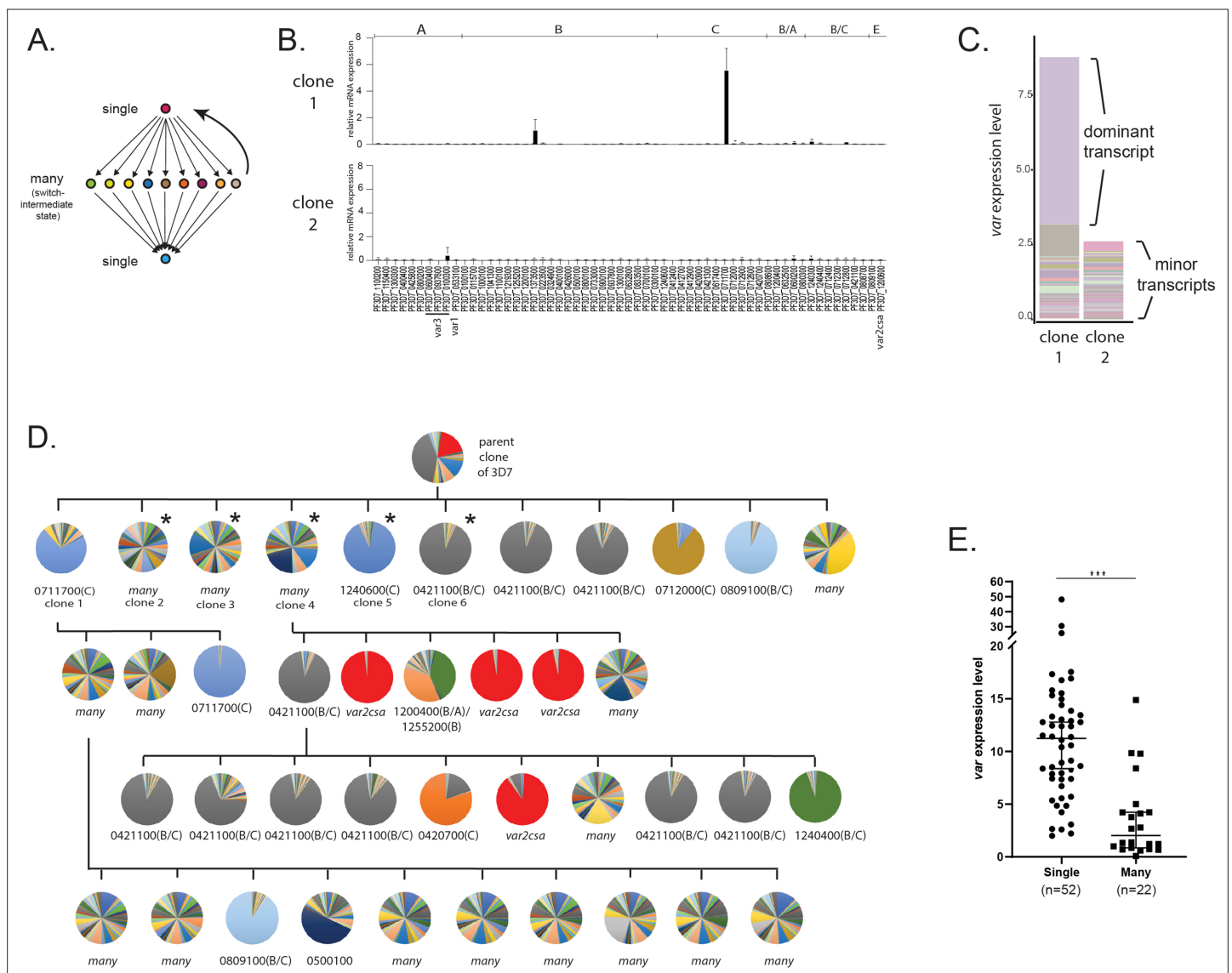


Figure 1. Detection of 'single' and 'many' *var* expression profiles in cultured parasite populations. **(A)** Schematic representation of the single-many-single (SMS) model for *var* gene transcriptional switching. Each circle represents expression of an individual *var* gene and arrows represent switches in expression. Switching events are hypothesized to transition from activation of a single gene (top) to a broad range of genes (middle) to a different single gene (bottom) or back to the original gene (reverse arrow). **(B)** *var* gene expression profiles for two clonal populations that display the single phenotype (clone 1, top) or the many phenotype (clone 2, bottom). *var* transcription levels were determined using a standardized quantitative real-time RT-PCR (qRT-PCR) protocol with the expression of each individual *var* gene displayed as relative copy number in the histogram. Error bars represent standard deviation of three biological replicates. **(C)** Total *var* expression for the two subclones shown in B, with transcripts from each individual *var* gene shown in a different color. For clone 1, transcripts from the dominant gene are marked, while both clones display similar levels of minor transcripts. **(D)** Clone tree of wildtype 3D7 parasites. Pie charts display the *var* expression profile for each subcloned population with each slice of the pie representing the expression level of a single *var* gene. Vertical and horizontal lines delineate sequential rounds of subcloning by limiting dilution. For parasite populations that display a dominantly expressed *var* gene, the annotation number is shown below the pie chart with the *var* type shown in parenthesis. The five subclones marked with an asterisk are further analysed in **Figure 2**. **(E)** Total *var* expression levels as determined by qRT-PCR for 74 subclones (see **Figure 1—figure supplement 1**). The median \pm 95% confidence interval is shown, and an unpaired t-test indicates a *** $p < 0.0001$.

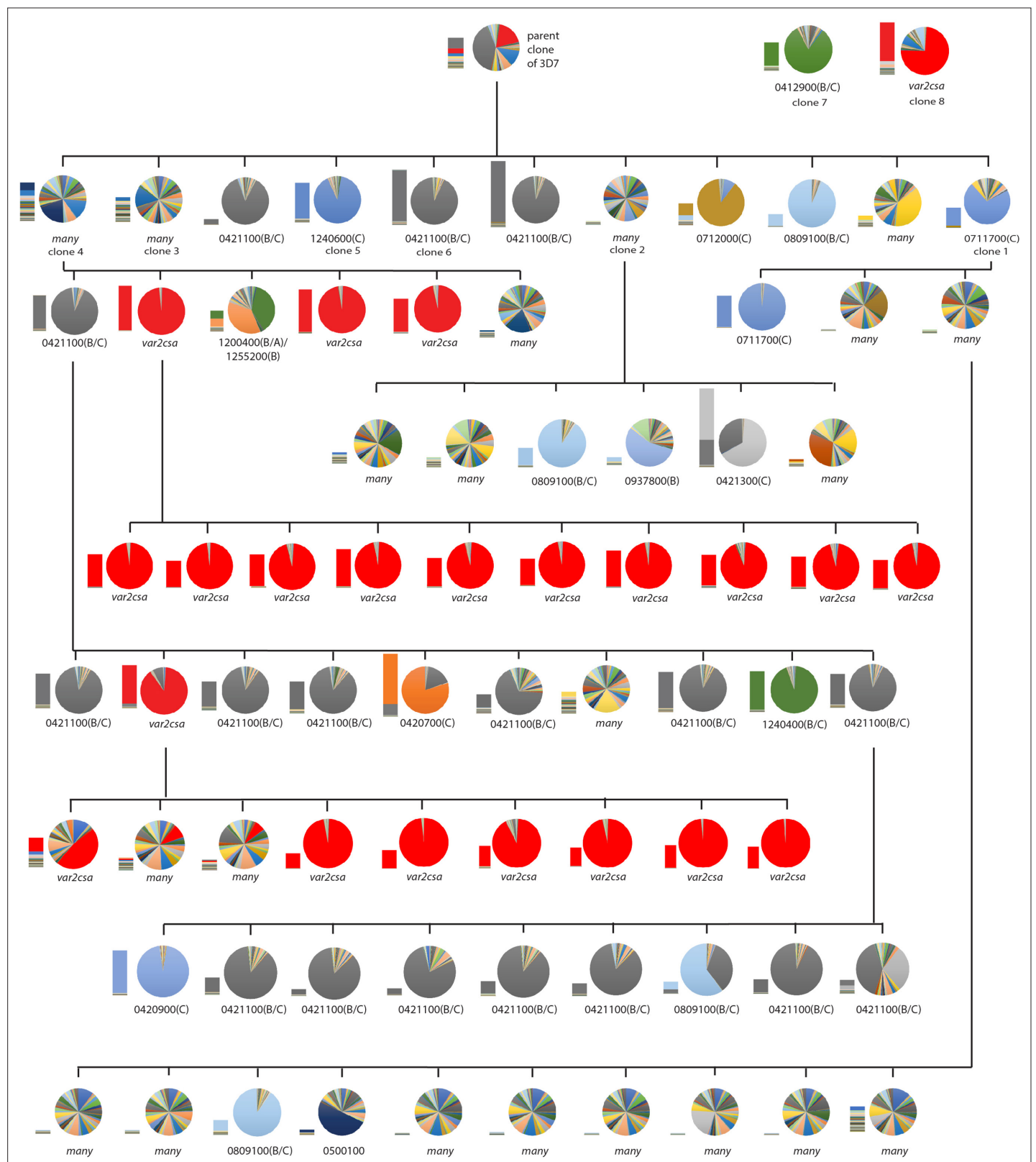


Figure 1—figure supplement 1. Clone tree of wildtype 3D7 parasites. Pie charts display the var expression profile for each subcloned population with each slice of the pie representing the expression level of a single var gene. Vertical and horizontal lines delineate sequential rounds of subcloning by limiting dilution. For parasite populations that display a dominantly expressed var gene, the annotation number is shown below the pie chart with the var type shown in parenthesis. A stacked bar graph is included next to each pie chart which displays the total var expression level, with transcripts from each individual var gene shown in a different color.

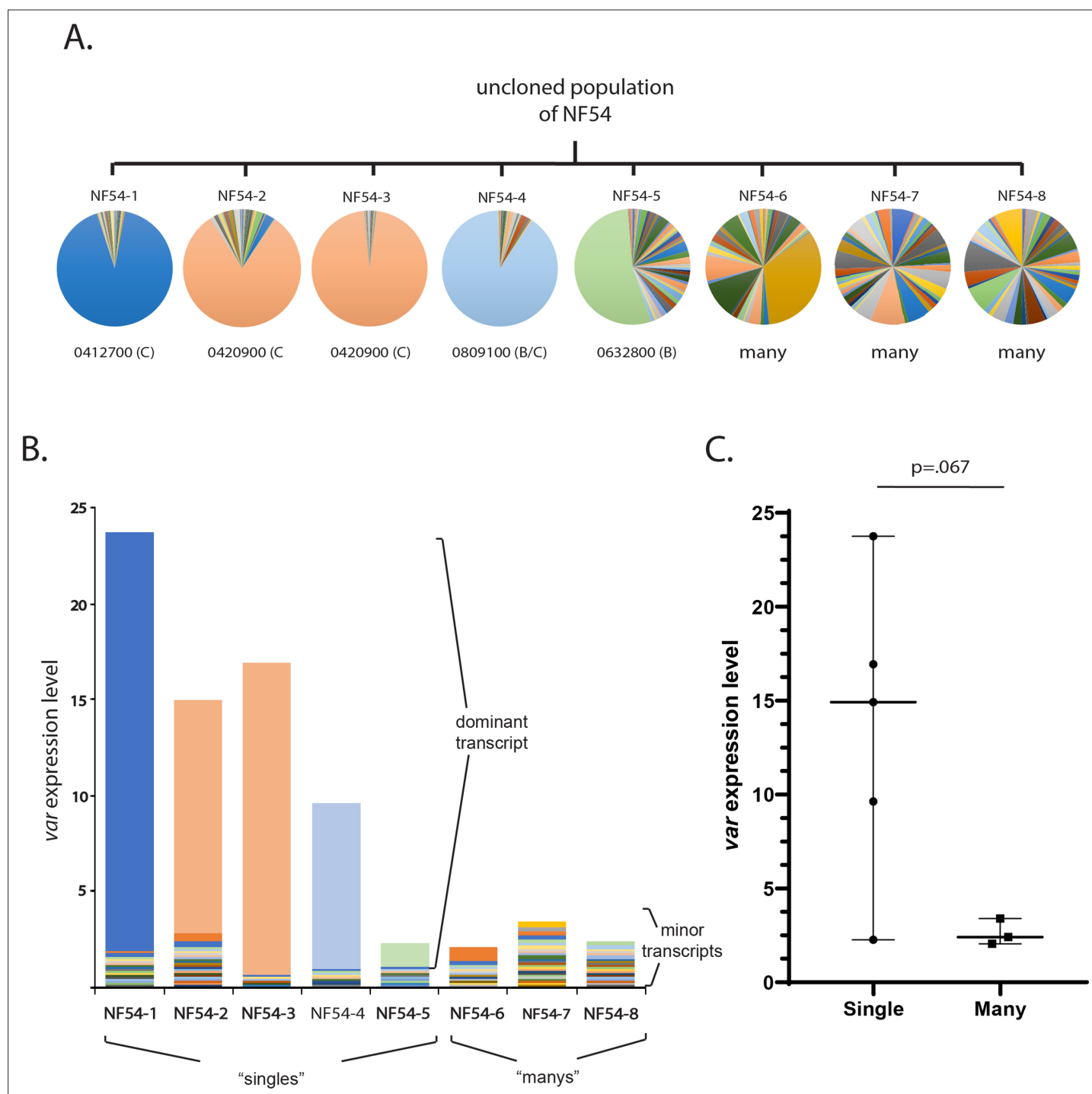


Figure 1—figure supplement 2. Examination of *var* gene expression profiles of recent clones of NF54. **(A)** Cloning by limiting dilution was performed using an uncloned, heterogenous population of the NF54 isolate. Eight individual clonal populations were generated and allowed to expand for 5 weeks, at which point they were synchronized and RNA isolated from ring stages. *var* gene expression profiles were generated and displayed as pie charts with the dominantly expressed *var* gene indicated below the chart for clones 1–4. **(B)** Total *var* expression levels for all eight clones shown in A, with transcripts for each *var* gene shown in a different color. **(C)** Total *var* expression levels as determined by quantitative real-time RT-PCR (qRT-PCR) for all eight NF54 subclones. The median \pm 95% confidence interval is shown, and an unpaired *t*-test indicates a *p* value of 0.0067.

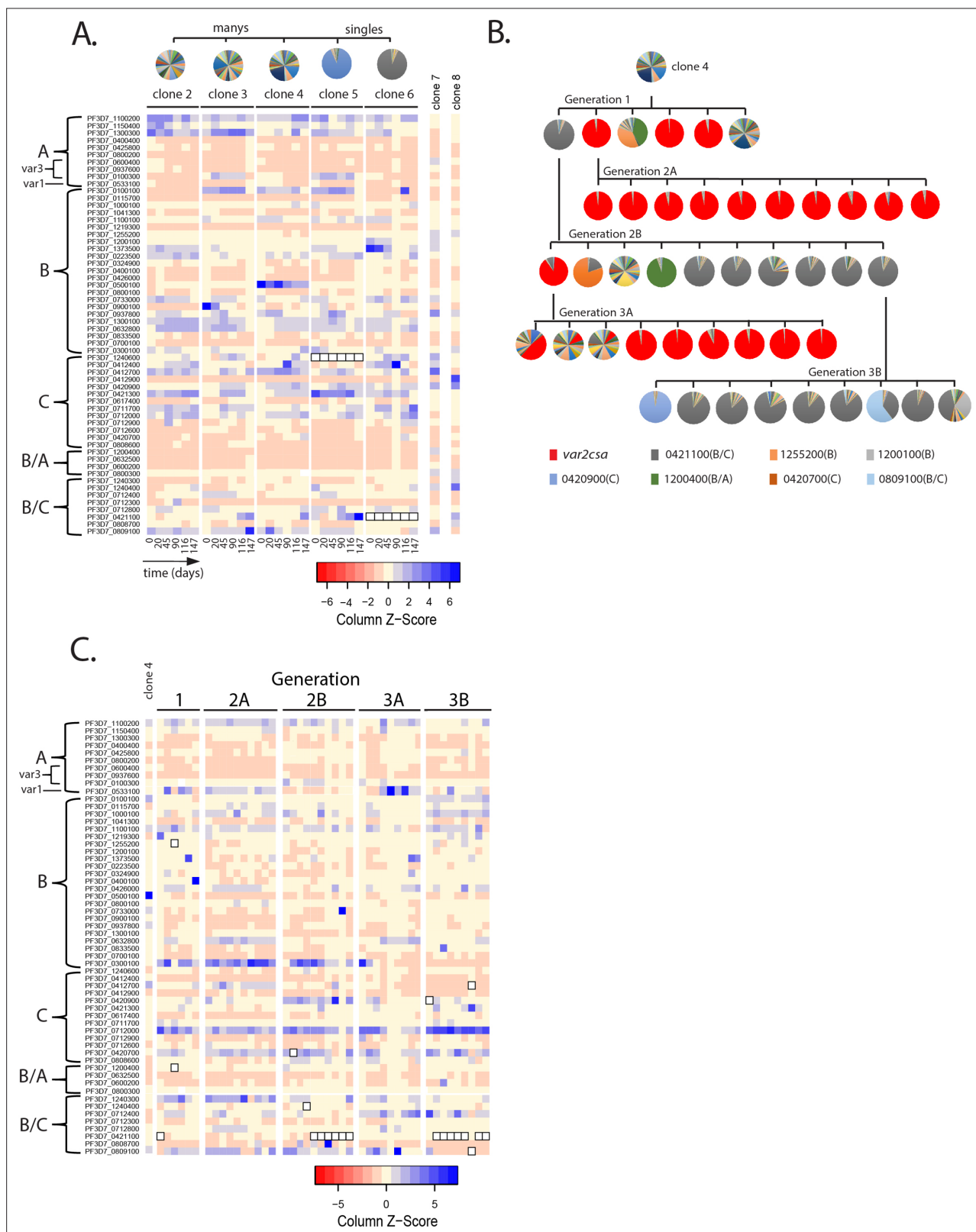


Figure 2. Detection of var gene minor transcript expression profiles in clonal populations overtime. **(A)** Heatmap of var minor transcripts for clones 2–6. The clones were all derived from the same parent population as shown in **Figure 1D**. Pie charts show the initial var gene expression profiles above the heatmap. The annotation numbers for each var gene are shown to the left of the heatmap and organized according to var type (left, var2csa is considered separately in **Figure 3**). Six time points are included for each clone. Two genetically identical parasite populations (clones 7 and 8) are shown for comparison. **Figure 2 continued on next page**

Figure 2 continued

and 8) also originally derived from 3D7 but grown separately for several years are shown for comparison. For clones 5 and 6, the dominant transcript (PF3D7_1240600 and PF3D7_0421100, respectively, marked by black boxes in the heatmap) was removed from the analysis to enable visualization of the minor transcripts. **(B)** Clone tree of wildtype 3D7. The tree is organized into three 'subclone generations' derived from initial clone 4 (top row). Pie charts display the *var* expression profile for each subclone. **(C)** Heatmap of *var* minor transcripts for the individual clones from each generation shown in B with the pattern of the parent population (clone 4) shown at the left for comparison. The annotation numbers for each *var* gene are shown to the left of the heatmap, and the gene order was organized according to *var* type. The order from left to right of each column in the heatmap corresponds to the order from left to right of the pie charts for each subclone generation shown in B. For parasites expressing the 'single' phenotype, the dominant transcript (marked by black boxes in the heatmap) was removed to enable visualization of the minor transcripts.

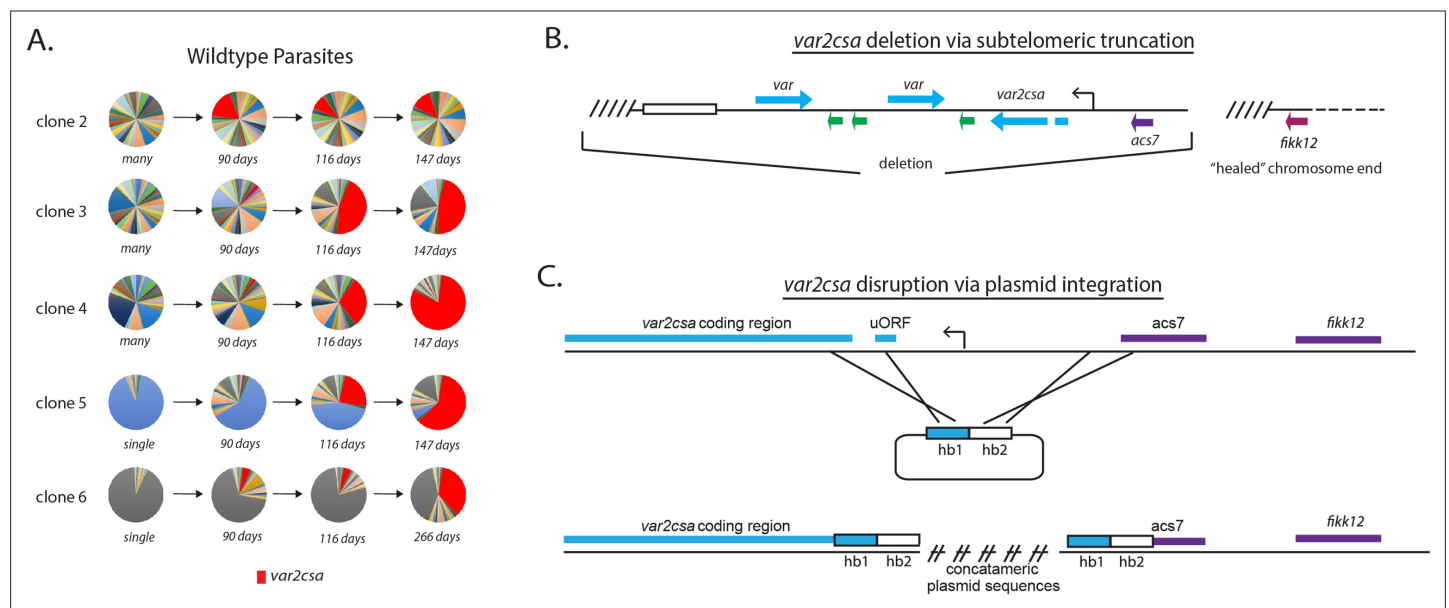


Figure 3. Convergence to *var2csa* expression in long-term cultures and targeted deletion of the *var2csa* locus. **(A)** Changes in *var* expression over time in five clonal parasite populations, three 'manys' (clones 2–4) and two 'singles' (clones 5 and 6), over several months of continuous culture. Expression of *var2csa* is shown in red. **(B)** Schematic diagram showing the truncation of the end of chromosome 12 by telomere healing. Telomere repeats are shown as slanted lines and the telomere-associated repeat elements (TAREs) are shown as an open box. The interior of the chromosome is displayed as a dashed line. A ~60 kb deletion, including three *var* genes (blue), three *rif* genes (green), and the *acs7* gene is shown. **(C)** Schematic diagram showing the *var2csa* locus and the plasmid containing homology blocks for targeted integration (hb1 and hb2). The crossed lines linking the plasmid to the chromosome signify sites of double cross over recombination leading to deletion of approximately 2.5 kb upstream of the *var2csa* gene, including the promoter.

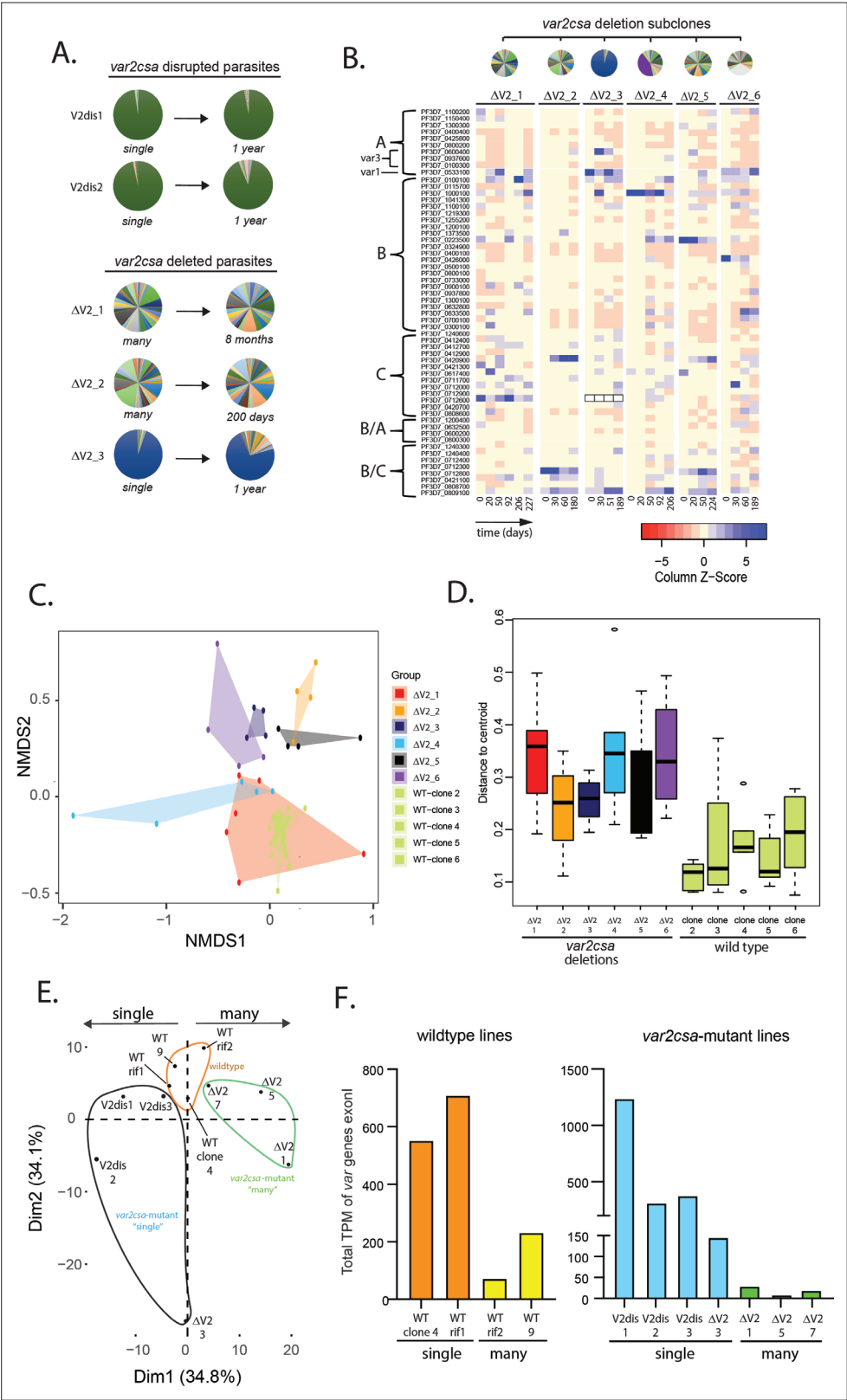


Figure 4. Alterations in *var* gene expression after deletion or disruption of *var2csa*. **(A)** Changes in *var* expression profile over time in five clonal parasite populations containing deletions or disruption of the *var2csa* locus, including two in which the locus was disrupted by plasmid integration (V2dis1 and V2dis2, top) and three in which the locus was deleted by chromosomal truncation (ΔV2_1, ΔV2_2, and ΔV2_3). V2dis1, V2dis2, and ΔV2_3 display

Figure 4 continued on next page

Figure 4 continued

a 'single' phenotype while $\Delta V2_1$ and $\Delta V2_2$ express many *var* genes. **(B)** Heatmap displaying the pattern of *var* minor transcripts over several time points for five *var2csa*-deleted lines. The clones were all derived from the same parent population, and pie charts showing the initial *var* gene expression profiles are shown above the heatmap. The annotation numbers for each *var* gene are shown to the left of the heatmap, organized according to *var* type. For $\Delta V2_3$, the dominant transcript (Pf3D7_0712600, marked by black boxes in the heatmap) was removed to enable visualization of the minor transcripts. **(C)** Non-metric multidimensional scaling (NMDS) plot based on Bray–Curtis dissimilarities displaying the expression profiles of the *var* minor transcripts for the six *var2csa*-deleted lines alongside five wildtype lines. **(D)** The variability within each cluster shown in C is inferred by its area on the plot (as defined by the mean distance from the group centroid for each time point) and displayed in a boxplot. **(E)** Transcriptomic analysis of wildtype and *var2csa*-mutated parasite lines. Principal component analysis (PCA) is shown based on the normalized expression level of 5721 genes $\log_{10}(\text{fragments per kilobase of transcript per million mapped fragments (FPKM)} + 1)$. RNAseq derived whole transcriptomes were obtained from 11 clonal parasite populations: two wildtype lines (WT 3.1 and WT 7) and two WT lines with a control plasmid integration into a *rif* gene (WT_rif1 and WT_rif2) are displayed in orange; four *var2csa*-mutated lines displaying a 'single' phenotype (V2dis1, V2dis2, V2dis3, and $\Delta V2_3$) are displayed in blue; and three *var2csa*-mutated lines displaying a "many" phenotype ($\Delta V2_1$, $\Delta V2_5$, and $\Delta V2_7$) are displayed in green. All analyses were performed using data obtained from synchronized populations ~16 hr after red cell invasion. **(F)** Comparisons of normalized *var* transcript counts from *var2csa*-mutated and wildtype lines displaying either the single (left) or many (right) phenotypes.

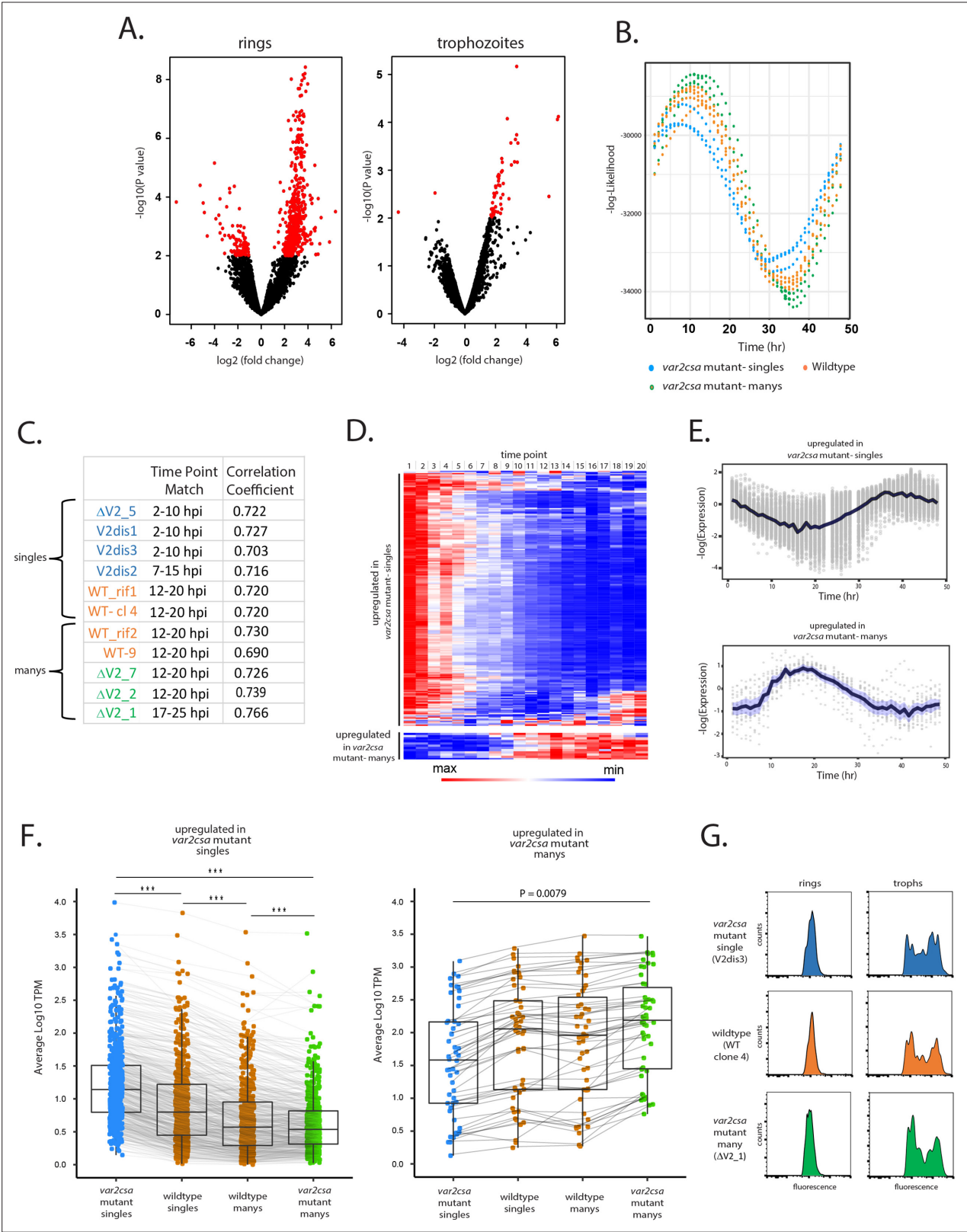


Figure 5. Parasites displaying ‘single’ vs. ‘many’ var expression profiles differ in replication cycle progression. **(A)** Volcano plots showing differentially expressed genes derived from whole transcriptome comparisons of *var2csa*-mutated ‘single’ and ‘many’ parasite lines. Differentially expressed genes are shown in red. For ring-stage parasites (left), 562 transcripts displayed higher expression levels in the single lines, while 49 were higher in the many lines. For trophozoite-stage parasites, only 51 genes were differentially expressed. **(B)** Estimation of cell cycle position using the method of

Figure 5 continued on next page

Figure 5 continued

Lemieux et al., 2009. Using the transcriptome profiles in **Supplementary file 8 (Bártfai et al., 2010)**, the likelihood (vertical axis) of each sample having been derived from a particular time point of the cycle (horizontal axis) is displayed. Eleven parasite lines are shown, with wildtype parasites in orange, *var2csa*-mutated, 'single' parasites in blue and *var2csa*-mutated, 'many' parasites in green type. **(C)** Table showing estimates of replication cycle progression for all 11 parasite lines. Estimates of the approximate time point within the 48-hr asexual cycle for each population were obtained by comparison with a modeled parasite population over a simulated infection time course using established datasets (**Poran et al., 2017**). The best time point match is shown for each population along with the correlation coefficient. Wildtype parasites are designated with orange type, *var2csa*-mutated, 'single' parasites with blue type and *var2csa*-mutated, 'many' parasites with green type. **(D)** Heatmap displaying changes in expression over the asexual replicative cycle for 611 differentially expressed genes according to the 48-hr asexual cycle of the *P. falciparum* HB3 transcriptome datasets defined by **Bozdech et al., 2003; Supplementary file 9**. Genes upregulated in *var2csa*-mutated 'single' lines are shown on top and those from *var2csa*-mutated 'many' lines are displayed on the bottom. **(E)** Average expression levels across the 48-hr asexual cycle of HB3 for 611 differentially expressed genes. Genes upregulated in *var2csa*-mutated, 'single' lines are shown on top and those upregulated in *var2csa*-mutated, 'many' lines are shown on the bottom. **(F)** Left panel, transcript expression levels for 559 differentially expressed genes that display higher expression in 'single' parasites. Average, normalized transcript levels were obtained from *var2csa*-mutated, 'single' parasites (blue), wildtype 'single' and 'many' parasites (orange, left and right, respectively), and *var2csa*-mutated, 'many' parasites (green). Box and whisker plots display the mean and standard deviation for each dataset. Right panel, analysis of transcript expression levels for 49 differentially expressed genes that display higher expression in 'many' parasites. Pairwise comparisons using t-tests with pooled standard deviation. p value adjustment method is from Bonferroni (significance codes: 0 '****', 0.001 '***', 0.01 '**'). **(G)** Monitoring of DNA content of infected cells by flow cytometry. Populations of *var2csa*-mutant 'singles' (top), wildtype (middle), and *var2csa*-mutant 'manys' (bottom) were tightly synchronized then assayed for DNA content by flow cytometry using Hoechst 33342 staining. Parasites were assayed at 16 and 36 hr after invasion.

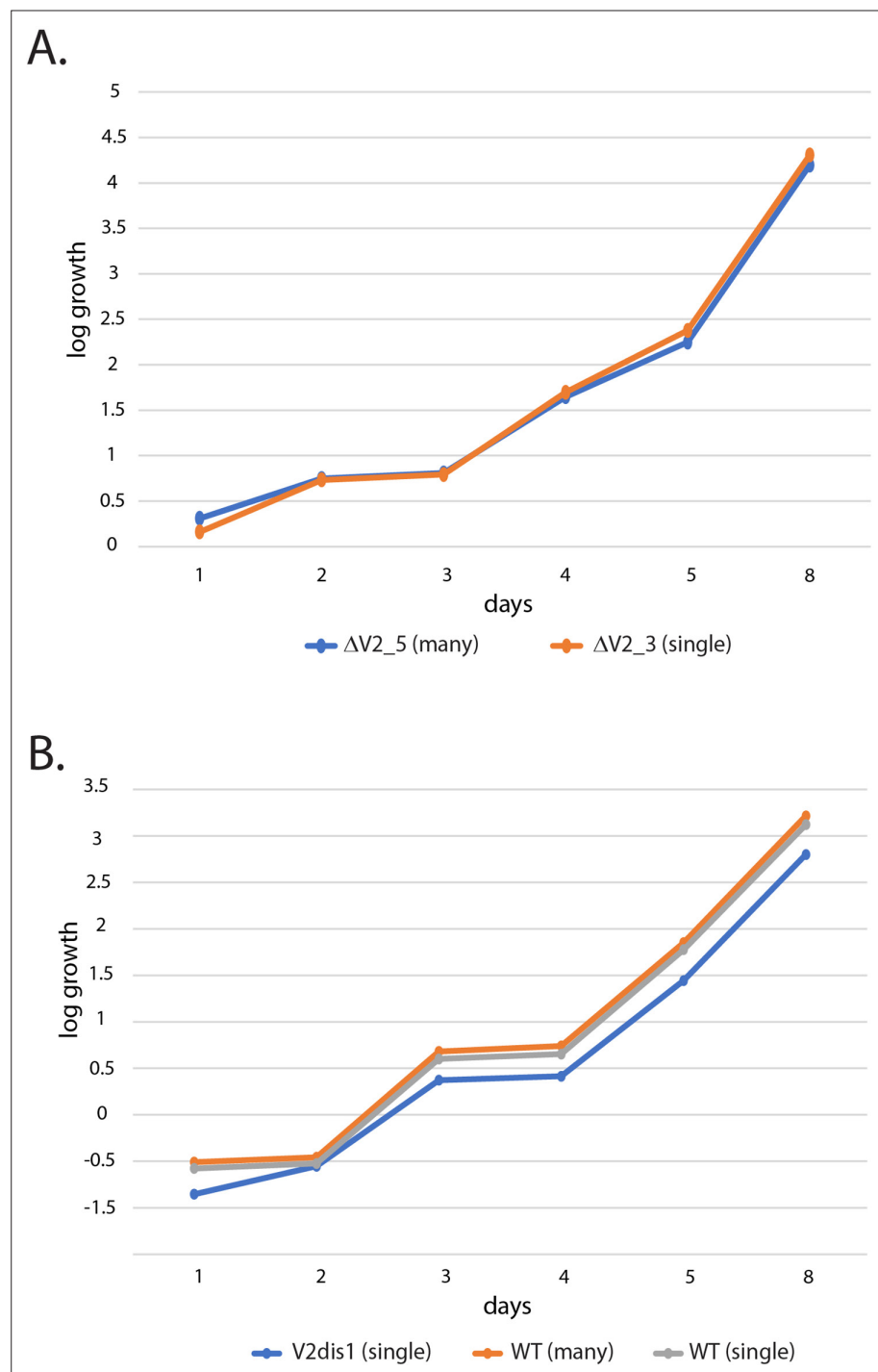


Figure 5—figure supplement 1. Growth assays of parasites displaying the ‘single’ and ‘many’ var expression phenotypes. Parasitemias were determined daily by flow cytometry and cultures were diluted by 10-fold whenever the parasitemia reached 0.5%. The daily parasitemia was multiplied by the exponential dilution factor and plotted as log growth over time. **(A)** Comparison of two clonal lines in which *var2csa* has been deleted. $\Delta V2_5$ (blue) displays the ‘many’ phenotype while $\Delta V2_3$ (orange) displays the ‘single’ expression profile. **(B)** Comparison of a *var2csa*-mutant line (V2dis1) displaying the ‘single’ phenotype with two wildtype lines, one displaying the many phenotype (orange) and one expressing a single dominant var gene (gray).