
Figures and figure supplements

The antigenic switching network of *Plasmodium falciparum* and its implications for the immuno-epidemiology of malaria

Robert Noble, et al.

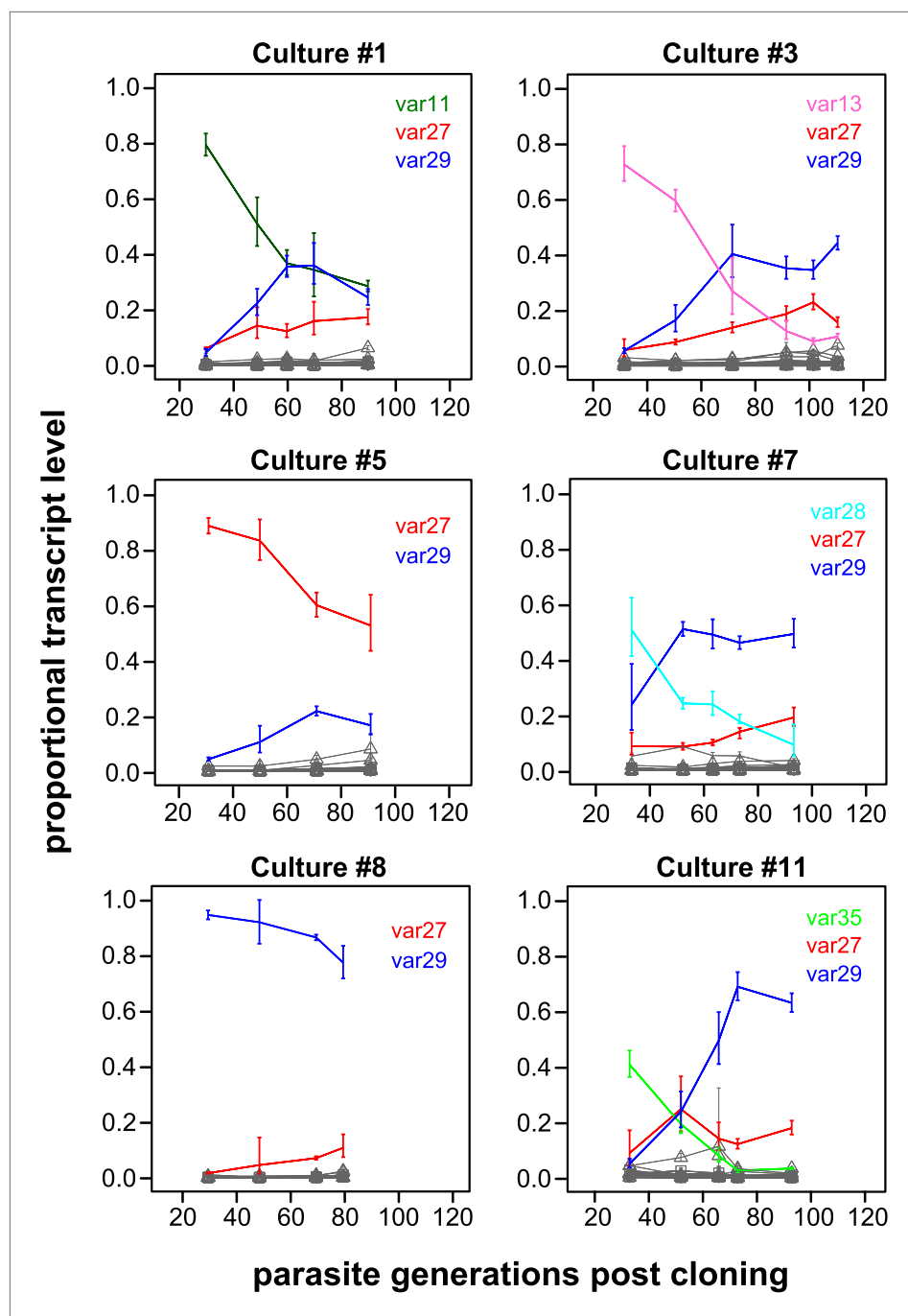


Figure 1. Proportional var transcript levels for six in vitro cultures. The parasite cultures initially expressed a variety of dominant 'starter genes', which switched off at notably different rates. Nevertheless, most cultures converged towards high level transcription of two centrally located genes var27 and var29 (red and blue lines, respectively), whereas most other gene transcripts (grey lines) remained at relatively low levels.

DOI: [10.7554/eLife.01074.004](https://doi.org/10.7554/eLife.01074.004)

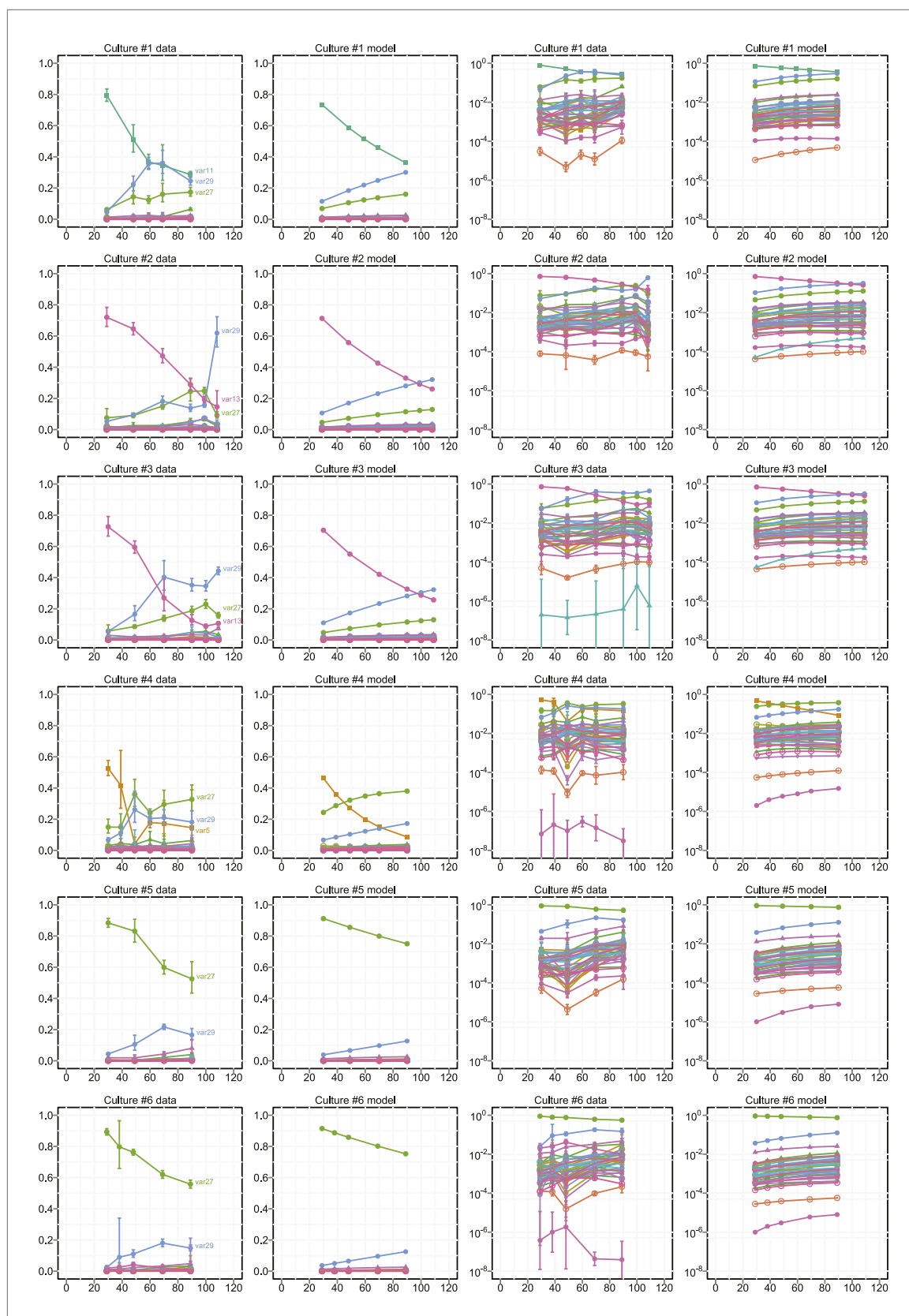


Figure 1—figure supplement 1. Proportional var transcript levels and model output for cultures 1–6. The relative transcript levels of activated var genes were measured by qRT-PCR at various time points during in vitro culture for 11 clones with different starter genes, plus the parent culture. Columns one
 Figure 1—figure supplement 1. Continued on next page

Figure 1—figure supplement 1. *Continued*

and two show the measured transcript levels at linear and logarithmic scale, respectively; columns three and four show the model fit. Error bars show approximate 95% confidence intervals estimated from qRT-PCR replicates.

DOI: [10.7554/eLife.01074.005](https://doi.org/10.7554/eLife.01074.005)

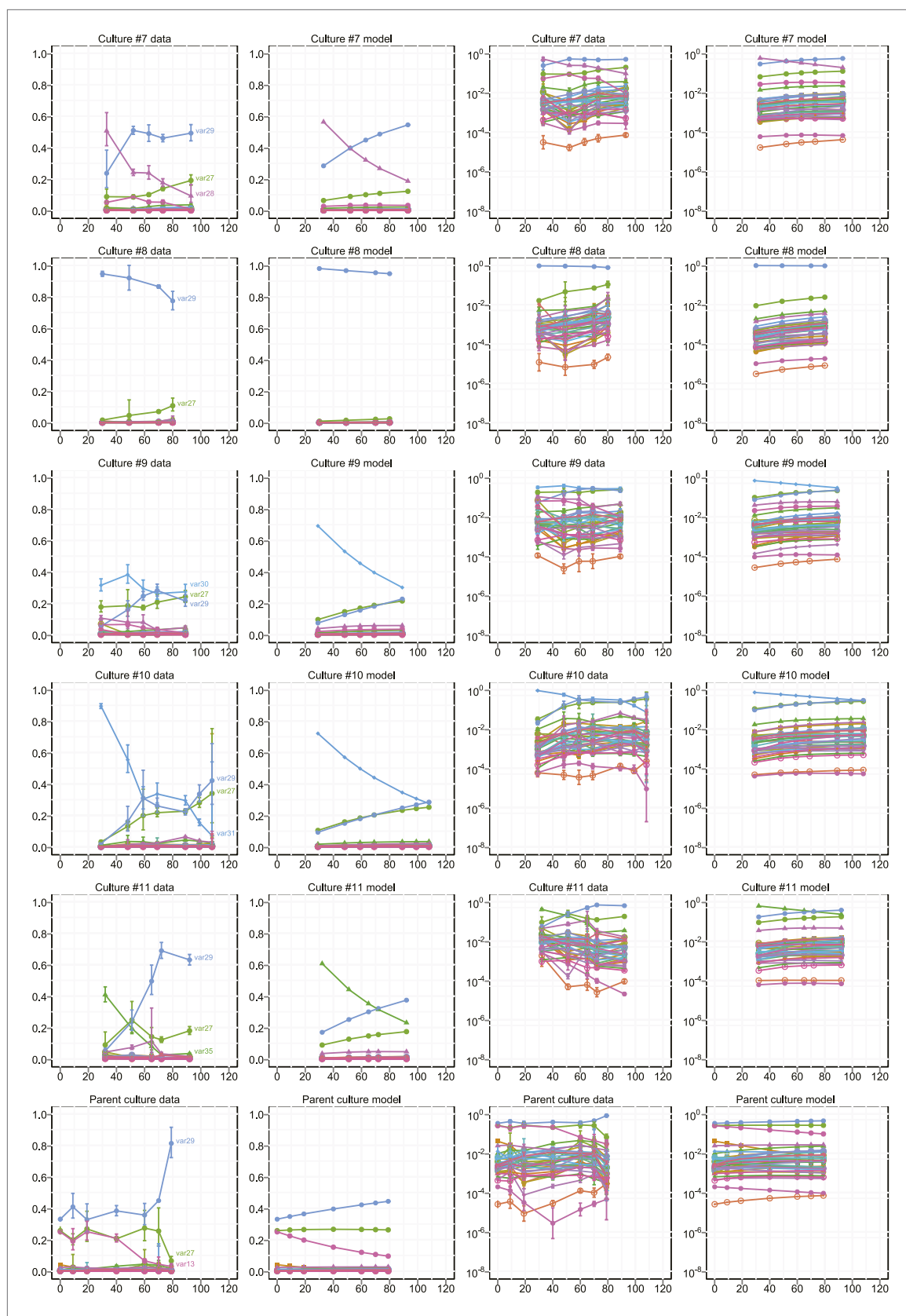


Figure 1—figure supplement 2. Proportional var transcript levels and model output for cultures 7–11. The relative transcript levels of activated var genes were measured by qRT-PCR at various time points during in vitro culture for 11 clones with different starter genes, plus the parent culture. *Figure 1—figure supplement 2. Continued on next page*

Figure 1—figure supplement 2. Continued

Columns one and two show the measured transcript levels at linear and logarithmic scale, respectively; columns three and four show the model fit. Error bars show approximate 95% confidence intervals estimated from qRT-PCR replicates.

DOI: [10.7554/eLife.01074.006](https://doi.org/10.7554/eLife.01074.006)

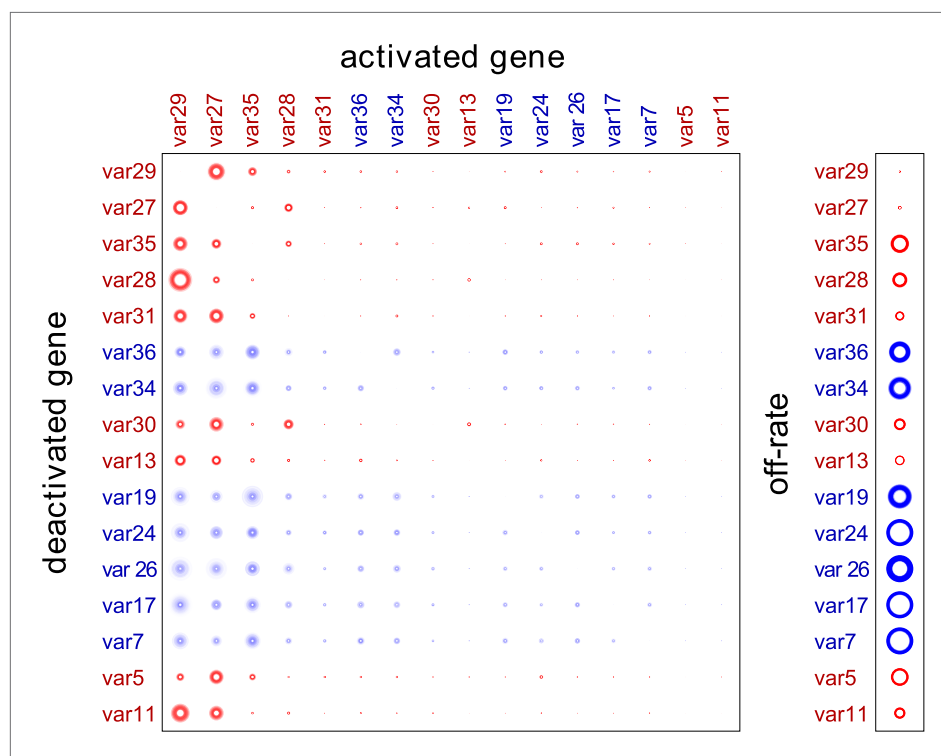


Figure 2. Estimation of switch parameters of highly transcribed genes. Parameter estimations for the 16 most transcribed var genes are represented as a switch matrix and an off-rate vector. The diameter of a circle in the i th row and j th column of the matrix is proportional to the switch bias $\beta_{ij} = \beta_j$ from gene i to gene j , and the diameter of a circle in the off-rate vector is proportional to the rate $\beta_{ij} = \beta_j$ at which gene i becomes silenced. The fuzziness indicates uncertainty in the estimate, such that the darkness of each concentric ring is proportional to the likelihood that the parameter is within the corresponding range (Noble and Recker, 2012). Starter gene parameters, which are more precisely estimated, are in red, and genes are arranged from left to right, and top to bottom, in order of average transcript level.

DOI: [10.7554/eLife.01074.007](https://doi.org/10.7554/eLife.01074.007)

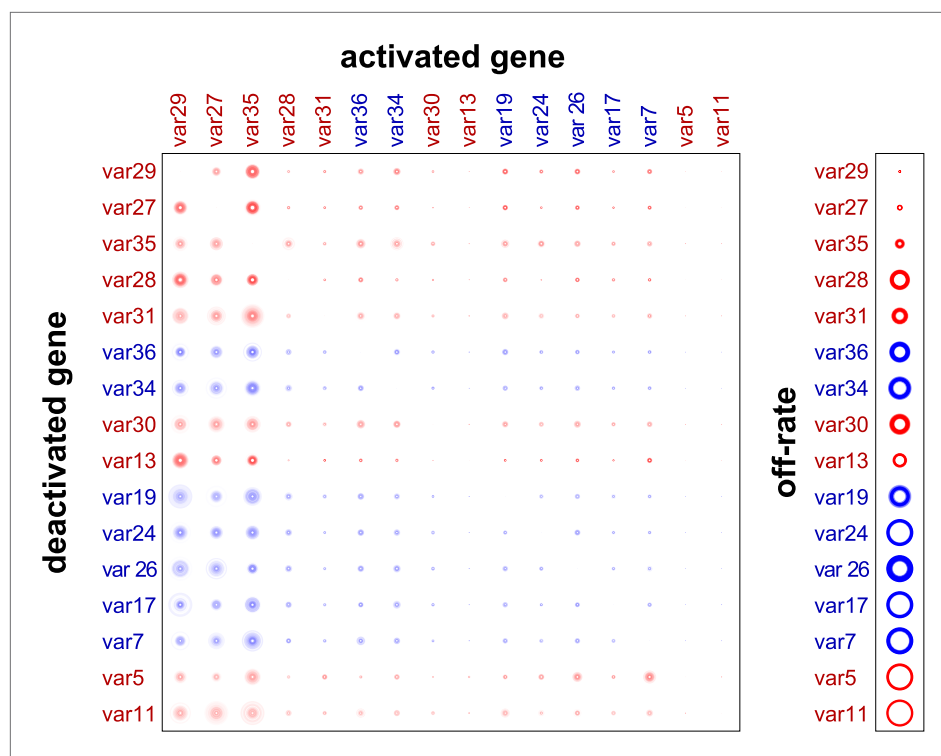


Figure 2—figure supplement 1. Validation of estimated switching parameters. Our method was able to estimate switching parameters not only for starter genes, which were initially transcribed in our cultures, but also for other *var* genes. To test the accuracy of the latter estimates, we used a cross-validation technique. For example, to generate an alternative set of parameter estimates for *var29*, we ran the MCMC algorithm without the data for culture #8, which began as a clone transcribing *var29*. The matrix shown here is a composite of estimates derived in this way. The similarity to **Figure 2** confirms that our method was reliable.

DOI: [10.7554/eLife.01074.009](https://doi.org/10.7554/eLife.01074.009)

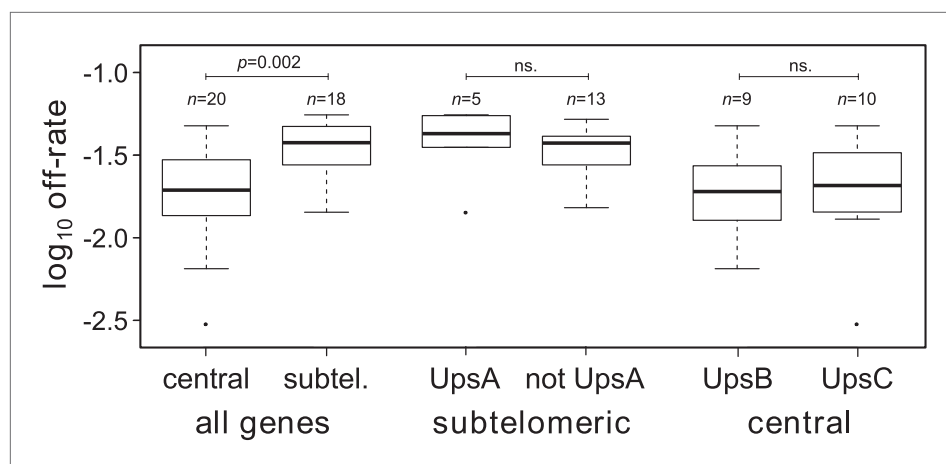


Figure 3. Off-rate estimates show strong dependency on chromosomal location. The estimated rates at which active var genes become silenced are significantly lower for centrally located var genes than for those in subtelomeric locations. There was no significant effect of upstream promoter type when differentially testing off-rates of genes in subtelomeric location (UpsA vs non-UpsA) or in central location (UpsB vs UpsC).
DOI: [10.7554/eLife.01074.011](https://doi.org/10.7554/eLife.01074.011)

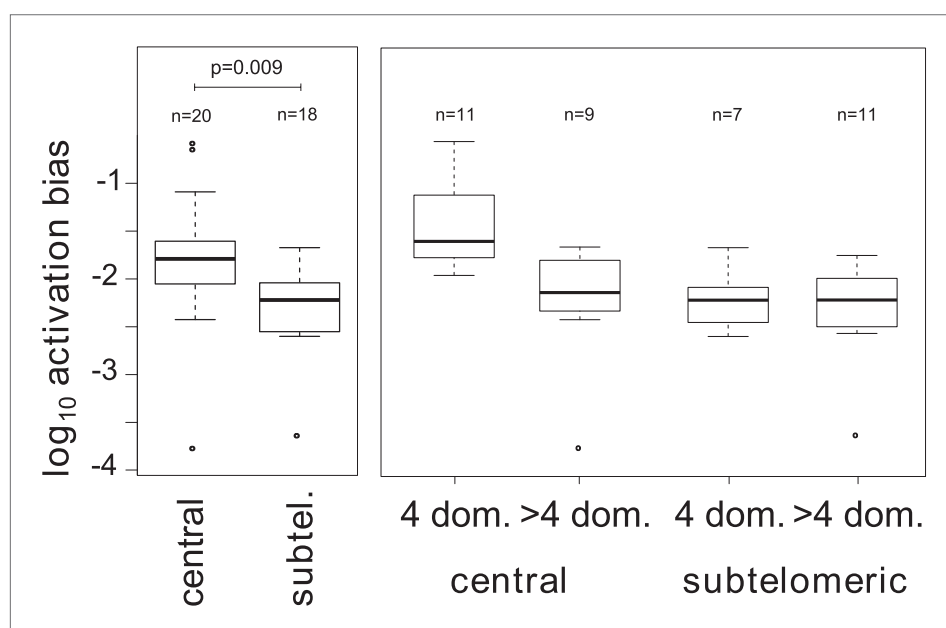


Figure 4. Associations between activation biases and genetic attributes. The mean activation bias is higher for centrally located var genes. Within the set of centrally located genes, those encoding only four binding domains have higher activation biases than longer genes. Together, short, central genes have activation biases ≈ 10 times higher than the rest of the repertoire.
DOI: [10.7554/eLife.01074.012](https://doi.org/10.7554/eLife.01074.012)

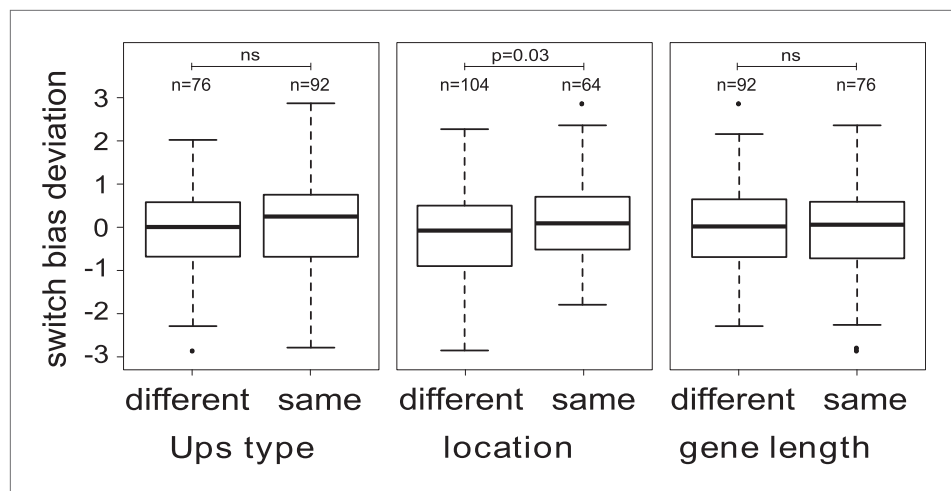


Figure 5. Variation in switch biases (deviation from the mean). Antigenic switching between *var* genes is more frequent between genes with matching chromosomal locations, for example from central to central or from subtelomeric to subtelomeric. No significant associations between switch bias and matching Ups type or gene length are found.

DOI: [10.7554/eLife.01074.013](https://doi.org/10.7554/eLife.01074.013)

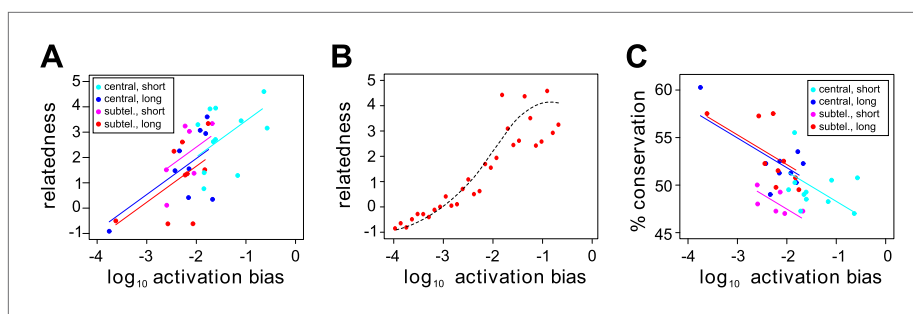


Figure 6. Effect of *var* gene activation on sequence evolution. **(A)** *Var* gene activation biases are significantly correlated with each gene's relatedness to the rest of the HB3 repertoire, as measured by the gene's centrality within a shared homology block network, independently of chromosomal location or gene length. Trend lines show the fit of a linear regression model. **(B)** Simulation of *var* gene evolution by gene conversion, whereby homology blocks are swapped among pairs of genes chosen at random according to their activation biases, shows a similar correlation between the genes' activation biases and their relatedness to the rest of the repertoire. The red points show the outcome of one simulation, and the dashed line is the smoothed average of 50 runs. **(C)** Activation biases are negatively correlated with the gene's average domain sequence conservation, as a measure of population-level diversity, independently of chromosomal location and gene length. Trend lines show the fit of a linear regression model.

DOI: [10.7554/eLife.01074.014](https://doi.org/10.7554/eLife.01074.014)

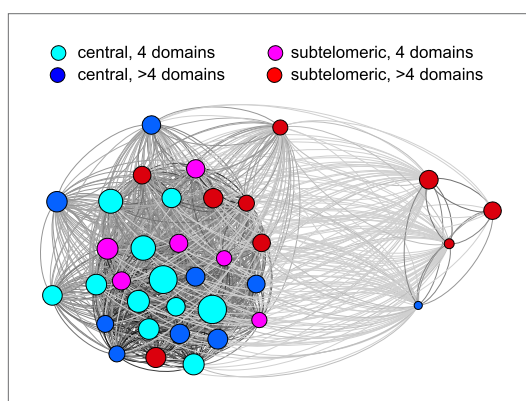


Figure 6—figure supplement 1. Relatedness network of the HB3 repertoire. Each node represents a *var* gene, colour coded according to the chromosomal location and the number of binding domains. Node size indicates estimated activation bias. The strength of a connection between two nodes is proportional to the number of shared homology blocks in the DBL1 and CIDR1 domains. Stronger connections are shown thicker and darker, and the layout is force-directed.

DOI: [10.7554/eLife.01074.015](https://doi.org/10.7554/eLife.01074.015)