Studying cis-regulatory heterogeneity in single-cells at allelic resolution

Veronika Petrova[1,2], Muqing Niu[1,2], Thomas Vierbuchen[3,4], Emily S Wong[1,2*]

[1] Victor Chang Cardiac Research Institute, Sydney 2010, Australia
[2] School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney 2033, Australia
[3] Developmental Biology Program, Sloan Kettering Institute for Cancer Research, New York, NY 10065, USA
[4] Center for Stem Cell Biology, Sloan Kettering Institute for Cancer Research, New York, NY 10065, USA

* Correspondence: e.wong@victorchang.edu.au

Abstract

The use of F1 hybrids, the offspring of inbred parental strains, is a powerful strategy for detecting cis- and trans-regulatory contributions to transcription. At single-cell resolution, this system enables unique insights into a diversity of transcriptional phenomena. However, the detection of allelic imbalance and variance in single-cell expression data is limited by extreme sparsity and overdispersion of the count data. Here, we present ASPEN, a statistical framework for robustly modelling cis-regulatory variation in single-cell RNA-sequencing data from F1 hybrids. ASPEN integrates a sensitive allelic quantification pipeline with an adaptive shrinkage approach, facilitating accurate inference of both mean allelic imbalance and variance without over-regularizing stably expressed genes. Through extensive simulation, we show that ASPEN reliably identifies cell state-specific allelic imbalances, even under sparse data conditions. Applying ASPEN to F1 hybrid mouse brain organoids and antigen-activated T cells, we identified 2,997 genes exhibiting significant allelic imbalance. In T cells, ASPEN revealed dynamic changes in allelic variance in 59 genes, including key transcriptional regulators, kinases, and components of immune signaling pathways. ASPEN also detected random monoallelic expression events and incomplete X-inactivation in neuronal progenitors. Together, these findings underscore ASPEN's capacity to elucidate the regulatory architecture of gene expression at allelic resolution.

Background

F1 hybrids have served as a foundational strategy for studying the genetic architecture of gene regulation. By combining evolutionarily divergent haplotypes within a shared nuclear environment, F1 hybrids disentangle the contributions of cis- and trans-regulatory variation to gene regulation. F1 hybrids have been used to reveal fundamental principles of transcriptional regulation and epigenetic processes, including parent-of-origin effects [1–3], X-chromosome inactivation[4,5], cis regulatory evolution [6,7], and the molecular basis of hybrid vigor [9,2]. This experimental system has been applied across taxa — from model mammals (e.g., mice [8,9]) and insects (e.g., *Drosophila* [10–12]) to plants (e.g., *Arabidopsis* [13–15]) and fungi [16–18].

Examination of cis-regulatory mechanisms in F1 hybrids can be combined with single cell genomic techniques, enabling inference of transcriptional burst kinetics in differentiated cell types [19,20], the detection of incomplete X inactivation [21,22], and characterization of cis-regulatory changes during lineage differentiation [29]. However, the full potential of these approaches has not yet been realized because statistical frameworks capable of robustly modeling allelic imbalance and allelic variance in single-cell data remain underdeveloped. Tools such as Airpart [23], scDALI [24], and SCALE [19] have advanced the field by addressing distinct biological questions using varied statistical paradigms, yet challenges persist—particularly in the context of data sparsity and the need to detect subtle regulatory shifts across distinct cell states/types.

To address these limitations, we introduce ASPEN (Allele-Specific Parameter Estimation in scRNA-seq), a flexible statistical method for modeling allelic imbalance and variance in single-cell transcriptomic data. ASPEN couples stringent allele-specific quantification with a moderated beta-binomial model that incorporates adaptive shrinkage to stabilize dispersion estimates without masking biological signal. This framework enables the sensitive and accurate detection of allele-specific expression patterns across diverse cellular contexts using scRNA-seq.

Validated against simulated and empirical datasets, ASPEN outperforms the state-of-the-art approach in sensitivity by 30% with robust false discovery control. Applying ASPEN to mouse brain organoids and T cells,

we uncover dynamic cis-regulatory changes during cellular differentiation, stochastic monoallelic expression, and signatures of incomplete dosage compensation in X-linked genes. ASPEN thus offers a robust and generalizable framework for uncovering transcriptional regulatory dynamics, with significant implications for understanding mechanisms of cis-regulatory control of gene expression in diverse cell types and tissues.

Results and Discussion

Precise regulatory control, not technical noise, underlies low allelic variation

Single-cell allele-specific expression analysis (scASE) is particularly susceptible to technical artifacts arising from sparse read coverage, which can obscure true biological variation [26]. To mitigate these issues, we implemented a stringent read-mapping pipeline that integrates both single nucleotide polymorphisms (SNPs) and small insertions/deletions (indels) to align reads against custom F1 hybrid genome assemblies. Only uniquely mappable reads with unambiguous allelic assignment were retained for downstream analysis. This strategy enabled the resolution of 20–38% of total sequencing reads across four widely used mouse hybrid crosses (**Methods**).

To improve statistical power while accounting for data sparsity, we used a moderated beta-binomial modeling framework with empirical Bayes shrinkage of dispersion parameters (**Fig. 1**). Within this framework, allelic variance is parameterized as dispersion—quantifying the deviation around the beta-binomial mean allele ratio. Shrinkage was applied across the transcriptome, stabilizing dispersion estimates by borrowing information across genes while still accommodating gene-specific departures from the global trend. This approach reduces susceptibility to technical noise and facilitates the inference of meaningful allelic variation [27,28].

We hypothesized that genes exhibiting consistently low allelic dispersion—while meeting minimal read coverage thresholds—may represent loci under stringent cis-regulatory control, with stable transcriptional output across individual cells. Importantly, in such cases, overshrinkage could artificially inflate dispersion estimates and mask biologically relevant allelic stability. To evaluate this, we compared allelic dispersion profiles between single-cell and bulk RNA-seq datasets generated from activated CD8+ T cells (day 7 post-infection) in *Mus musculus* × *Mus spretus* F1 hybrids [23,29]. Genes with low dispersion in single-cell data ($\theta < 0.005$) maintained highly stable allelic ratios and exhibited robust expression in bulk RNA-seq (Wilcoxon $Z = -8.23$, $p < 2.2 \times 10^{-16}$; **Fig. 2**). Functional enrichment analysis revealed that these genes are significantly associated with essential housekeeping processes (FDR < 0.1; **Supp. Fig. 1A**), including oxidative phosphorylation and ribosome biogenesis—suggesting these pathways may be subject to strong evolutionary constraint and tightly coordinated expression.

Further supporting this notion, genes critical for CD8+ T cell clonal expansion and effector differentiation, such as *Il2ra* and *Cd8a*, exhibited significantly lower allelic variance relative to nonessential genes matched for expression level (Wilcoxon $Z = -5.54$, $p = 3.04 \times 10^{-8}$; **Supp. Fig. 1B**) [30]. These findings support that reduced allelic variance is not simply a consequence of technical sparsity, but rather reflects precise regulatory control, particularly in genes central to cell identity and function.

ASPEN accurately detects allelic imbalance in sparse single-cell data

Building upon our observation that low-dispersion genes reflect bona fide allelic stability rather than technical noise, we implemented an adaptive statistical strategy for testing allelic imbalance. ASPEN evaluates allelic count data by maximum likelihood estimation (MLE) under two competing models. The alternative model estimates the mean allelic ratio directly from the data. The null model assumes allelic counts are drawn from a beta-binomial distribution with a fixed mean corresponding to allelic balance (e.g. 0.5 for autosomes), incorporating gene-specific dispersion estimates that are subject to empirical Bayes shrinkage. Based on our above observations, we exempt from shrinkage those genes with low dispersion ($\theta < 0.005$) to preserve biological signal (**Fig. 3A, Methods**).

To validate this framework, we performed extensive simulations across a broad range of gene expression levels, allelic ratios, and dispersion parameters. ASPEN consistently demonstrated high sensitivity for detecting allelic imbalance while maintaining stringent control over false positives (**Fig. 3B; Supp. Table 1**). For low-dispersion genes, exemption from shrinkage improved sensitivity by up to 22%. In contrast, genes with high allelic dispersion benefited from shrinkage, which enhanced detection accuracy by up to 30% through stabilized dispersion estimates. We benchmarked ASPEN against scDALI [25], a method that models ASE via generalized linear mixed models, and found that ASPEN detected up to 30% more imbalanced genes across

simulated datasets, with no detectable increase in false positives among allelically balanced controls (**Fig. 3C; Supp. Fig. 2A**).

We further assessed ASPEN's performance on empirical data using positive control gene sets. In brain organoids derived from male *C57BL/6J* × *Mus Spretus (SPRET/Ei)* F1 hybrids, ASPEN reliably identified 100% of expressed X-linked genes (43/43) as monoallelic, as expected due to hemizygosity. Additionally, all known imprinted loci present in the dataset (10/10), including *Meg3* and *Peg3*, were correctly recovered (**Fig. 3D; Supp. Table 2**), confirming ASPEN's capacity to resolve allele-specific expression with high precision.

To compare single-cell results with bulk-level expression trends, we aggregated scRNA-seq data into pseudobulk profiles across three major brain organoid subpopulations. ASPEN's single-cell allelic imbalance calls showed strong concordance with pseudobulk-derived results, with 74–82% overlap in brain organoids and 77–93% agreement in T cells across different activation states (**Fig. 3E; Supp. Fig. 2B**). The slightly lower concordance in organoids may reflect a higher degree of cell-to-cell heterogeneity in allelic expression, consistent with prior observations of increased random monoallelic expression during embryonic stem cell differentiation [39] (**Supp. Fig. 2C**).

Finally, we applied ASPEN to a time course of CD8+ T cell responses during acute viral infection in *C57BL/6J* × *Mus Spretus (SPRET/Ei)* F1 hybrid [29]. ASPEN identified significant allelic imbalance in 491 out of 809 expressed genes (FDR < 0.05, **Supp. Table 3**), including canonical markers of T-cell activation (*Cd69*), effector differentiation (*Gzmb*, *Ifng*), and memory formation (*Tcf7*, *Il7r*) (**Supp. Fig. 2D**).

Allelic variance reveals transcriptional plasticity across cell states

ASPEN incorporates a dedicated statistical framework to test for allelic over- and under- dispersion, enabling the identification of loci where allelic variance deviates from expression-level expectations (**Fig. 4A**; **Methods**). Conceptually, reduced allelic variance may reflect selective pressure for transcriptional precision, whereas increased variance can signal regulatory plasticity, stochastic gene expression, or context-dependent modulation of cis-regulatory inputs [31].

To detect such deviations, ASPEN first models the expected allelic dispersion as a function of gene expression, capturing the transcriptome-wide trend. It then evaluates whether individual genes significantly depart from this trend using an empirical permutation-based procedure applied to stabilized (shrunk) dispersion estimates (**Methods**). In female *C57BL/6J* × *Mus Castaneus (CAST/Ei)* F1 hybrid brain organoids, we identified 232 out of 1,261 expressed genes whose allelic variance significantly deviated from the expectation (**Fig. 4B**). Of these, 35% (83/232) were significantly under dispersed, while 65% (149/232) displayed significant over dispersion.

Genes with reduced allelic variance were enriched for core cellular processes, including RNA splicing, mRNA metabolism, and cell cycle regulation (FDR < 0.5; **Fig. 4C**), consistent with prior observations. These findings suggest that allelic stability is selectively maintained in genes essential for cell integrity. By contrast, over dispersed genes were enriched for neurodevelopmental pathways, including glial differentiation (*Sox4, Olig1, Ptn, Nfib*), axonogenesis (*Dcx, Pac3, Map1b, Nefm*), neurogenesis (*Vim, Nrp1, Sox11, Syt4*), and CNS development (*Tubb2a, Cenpf, Fabp7, Neurog2*) (FDR < 0.05; **Fig. 4C**).

To capture dynamic regulatory shifts during differentiation, ASPEN implements two likelihood ratio tests: one to detect temporal changes in allelic mean (*dyn-mean*) and another to test for temporal shifts in allelic variance (*dyn-var*) (**Fig. 4A, D**). In both cases, a null model assumes a shared set of beta-binomial parameters across timepoints, whereas the alternative model allows group-specific parameters. For the *dyn-mean* test, dispersion is held constant to isolate mean shifts, while in *dyn-var*, the mean is fixed to assess dynamic changes in variance (**Fig. 4E**; **Methods**).

We validated these models using simulated single-cell data along a pseudo-temporal axis defined by principal component analysis (**Fig. 4F**). Cells were grouped into early, intermediate, and late stages, and simulations were constructed to model (1) dynamic changes in allelic mean, (2) dynamic changes in allelic variance, and (3) a null model with stable parameters (**Supp. Fig. 3A**). ASPEN demonstrated high sensitivity (98%) and specificity (92%) with a low false positive rate (<12%), confirming its capacity to distinguish distinct modes of regulatory variation during lineage differentiation (**Fig. 4G**).

Applying ASPEN to early neurogenesis in female *C57BL/6J × Mus Castaneus (CAST/Ei)* F1 hybrids revealed 33 genes with significant changes in mean allelic expression and 193 with dynamic variance, independent of total expression changes (**Supp. Fig. 3B; Supp. Table 4**). These included regulatory kinases (*Csnk1a1*), chromatin and transcriptional modulators (*Pbx1, Hmgb1, Fos*), and transcriptional co-regulators (*Taf1d, Ccna, Trim28r1, Eid1*). Notably, early progenitor populations exhibited the lowest variance ($\theta = 0.07 \pm 0.01$), which marginally increased in intermediate progenitors ($\theta = 0.09 \pm 0.01$) and differentiated cortical neurons ($\theta = 0.10 \pm 0.02$). This pattern is consistent with increasing heterogeneity in cell states during lineage commitment [32].

Four genes with dynamic variance—*Ankrd11, Csde1, Nrxn1,* and *Zbtb20*—are high-confidence autism risk genes (SFARI database) [33]. Additionally, four genes exhibiting variable allelic expression (*Eif3h, Lars, Oaz1, Selenok*) have been implicated in anatomical brain phenotypes based on mouse knockout models [34] (**Supp. Fig. 3C**). It may be possible that the temporal modulation of variance could serve as a regulatory mechanism that fine-tunes gene expression programs during development and may be perturbed in disease states.

Monoallelic expression reveals regulatory complexity

We next leveraged ASPEN's estimation of beta-binomial shape parameters to identify genes exhibiting monoallelic expression (**Fig. 5A**). Among 1,385 genes expressed in neuronal cells from female *C57BL/6J × Mus Castaneus (CAST/Ei)* F1 hybrids, 72 demonstrated monoallelic patterns, defined by extreme skew in the beta-binomial shape parameters, wherein either α or β was less than 1 (**Fig. 5B; Supp. Table 5**). In this framework, α and β represent the inferred support for expression from each allele; values below 1 indicate peaked distributions suggestive of allelic skew or monoallelic expression. The approach successfully recovered several canonical imprinted genes, including *Meg3*, *Peg3*, *Kcnq1ot1*, *Rian*, and *Snrpn* (**Fig. 5B; Supp. Fig. 4A**).

To further resolve random monoallelic expression (RME)—a phenomenon in which one allele is stochastically expressed in individual cells [35–38], we applied more stringent criteria. Specifically, we required both α and β to fall below 1, with a difference of less than 0.5 between them ($|\alpha - \beta| < 0.5$), reflecting symmetrical yet peaked distributions consistent with cell-to-cell allelic randomness. We reasoned that such a bimodal distribution of allelic expression is characteristic of RME where each cell shows monoallelic expression, whereas genes with low transcriptional burst frequencies are likely to show a broader unimodal expression centered around an allelic ratio of approximately 0.5, reflecting transcription from both alleles. These candidates were further filtered using ASPEN's allelic variance test (FDR < 0.05), yielding 39 high-confidence RME genes (**Fig. 5C**). Of these, 27 were X-linked, consistent with random X-inactivation dynamics in female cells, and five (*Bex2*, *Ndufb11*, *Pcsk1n*, *Sh3bgrl*, *Uba1*) displayed signatures of incomplete X inactivation, by demonstrating largely monoallelic expression in each cell (**Supp. Fig. 4B; Methods**).

The remaining twelve autosomal genes included *Gadd45g* and *H2bc4*, which we report here as newly identified RME loci (**Fig. 5D; Supp. Fig. 4C**). Notably, most of the autosomal RME genes recovered by ASPEN have been previously validated through classical clonal expansion followed by bulk RNA-seq, demonstrating that ASPEN enables high-resolution, *in vivo* detection of RME (**Supp. Fig. 4D**) [32,39,40]. These results underscore the utility of allelic modeling for revealing complex regulatory landscapes that would be otherwise inaccessible using traditional clonal approaches.

Dynamic Regulation in T-Cell Activation

We next applied ASPEN to dissect the allelic regulatory dynamics accompanying T-cell activation during acute viral infection, during which naïve CD8+ T cells transition through effector and memory states. Among 809 expressed genes, ASPEN identified 90 loci with allelic variance that deviated significantly from expression-level expectations (**FDR < 0.05; Fig. 6A**).

As observed in cortical organoids, under-dispersed genes in T cells are enriched for broadly conserved cellular processes, whereas over dispersed genes are associated with immune-specific functions. These included pathways related to immune system processes (*Trac, Tox, Nkg7, Ccl5*), lymphocyte activation (*Lgals1, Actb, Themis, Lef1*), and leukocyte signaling (*Myo1f, Klrd1, Thy1, Ccnd3*) (**Supp. Fig. 5A**). This suggests that allelic variance may be modulated to support cell type–specific functions while preserving tighter transcriptional control in essential genes.

Allelic variance shifted markedly over the course of T-cell activation. Naïve cells exhibited the highest allelic variance ($\theta = 0.12 \pm 0.04$), which progressively declined at day 7 post-infection ($\theta = 0.05 \pm 0.01$) and further

decreased in memory cells at day 40 ($\theta = 0.02 \pm 0.01$) (**Fig. 6B; Supp. Fig. 5B**). Among 119 dynamically expressed genes, ASPEN identified significant temporal changes in allelic mean ratios for 71 genes (60%) and dynamic allelic variance for 59 genes (50%), with 40 genes showing coordinated changes in both parameters (FDR < 0.05 for both tests; **Fig. 6C; Supp. Table 6**).

Genes exhibiting dynamic allelic regulation spanned diverse functional categories, including transcription factors, kinases, and co-regulators. These genes displayed variable cis-regulatory patterns over time (**Fig. 6D; Supp. Fig. 5C**). For example, *Grap2*, an adaptor protein specific to T lymphocytes, showed high allelic variance early during activation that resolved into a more stable pattern following infection clearance—perhaps suggesting a shift from stochastic to more deterministic expression control. In contrast, *Actb* demonstrated a shift in allelic preference without concomitant changes in variance, while *Trbc1* transitioned from biallelic to maternally biased expression while maintaining stable total transcript levels. Together, these examples highlight the nuanced regulatory programs that govern allelic expression dynamics during immune activation.

Conclusion

ASPEN enables the interrogation of cis-regulatory effects on gene expression by leveraging single-cell, allelically resolved transcriptomic data. Central to its framework is an adaptive shrinkage approach that refines allelic expression modeling by selectively stabilizing dispersion estimates across genes and cells. This strategy enhances statistical power, particularly for detecting subtle biological deviations in allelic balance. Beyond testing for changes in mean allelic expression, ASPEN incorporates statistical models for assessing shifts in allelic variance, capturing regulatory processes that manifest through transcriptional heterogeneity. These include stochastic gene expression [41], epigenetic changes [42], cell state transitions [43], parent-of-origin effects [44], and random monoallelic expression (RME) [35,36].

Applying of ASPEN to mouse brain organoid and T cell differentiation data recovered shifts in transcriptional variances that could reflect how cells explore expression states during differentiation en route to terminal differentiation. Several genes exhibiting altered allelic variance in brain organoids have established roles in neurodevelopment and have been associated with neuropsychiatric disorders, including autism spectrum disorder. This aligns with emerging evidence that expression variance may play a mechanistic role in disease. For example, allelic bursting of *MYBPC3* in cardiomyocytes has been implicated in haploinsufficiency and cardiac dysfunction [45], while incomplete X-inactivation in neural tube cells can result in stochastic failures in X-inactivation [46]. Low-copy transcription factors can exhibit burst-like expression, propagating stochasticity to downstream targets at a systems level [47].

We further demonstrate that ASPEN applied to scRNA-seq can be used to characterize monoallelic expression *in vivo*, enabling the inference of monoallelic expression patterns during cell differentiation. scRNA-seq allows detection of RME *in vivo*, in cases where identification through clonal analysis is not possible, such as in postmitotic cell types, e.g. neurons and cardiomyocytes, or in rare cell populations [32,38,39]. While RME and low frequency transcriptional bursting can both produce apparent monoallelic expression at the single cell level, they are governed by distinct dynamics: where RME reflects a stable, cell-intrinsic allelic bias while bursty genes are expected to transition toward biallelic expression as transcript converge increases. ASPEN can also identify other forms of monoallelic expression, for example, incomplete X inactivation or escape from genomic imprinting or X inactivation, to reveal the diversity of regulatory strategies that shape transcriptional output across cell states.

Methods

We generated F1 alignment indices for four common mouse hybrid strains: *C57BL/6J* (Bl6) × *Mus Castaneus (CAST/Ei)* (Cast), Bl6 × *Mus musculus molossinus (MOLF/Ei)* (Molf), Bl6 × *Mus musculus musculus (PWK/Ph)* (Pwk), and *C57BL/6J* × *Mus Spretus (SPRET/Ei)* (Spret). Reads from hybrid animals were aligned to their parental combined genomes. Only reads that could be allelically resolved were used in gene expression quantitation. By incorporating both SNPs and indels in the genome references, read mapping improved by an average of 10% compared to using SNPs alone (**Supp Fig. 6**).

Combined genome generation

We generated combined genomes containing two sets of chromosomes for each genetic background, which allowed for unambiguous assignment of allelic reads. Variant call files (release 5) for four wild-derived mouse strains (*CAST/Ei, MOLF/Ei*, *PWK/Ph*, and *SPRET/Ei*) were downloaded from the Mouse Genome Project [48]. Strain-specific genomes were created using high-quality calls (filter set to PASS) and incorporating (i) only

SNPs and (ii) both SNPs and insertions and deletions (indels) into the reference *C57BL/6J* genome (GRCm38, release 68) with BCFtools consensus (v1.12) [49]. Each genome was concatenated with the reference *C57BL/6J* genome to produce combined genomes for each mouse strain. The chain files from the BCFtools were used to create strain-specific gene annotations by converting *C57BL/6J* gene annotations (Ensembl release 102) via UCSC liftOver [50]. Combined genome index files were generated using STAR (v2.7.6a) [51].

Beta-binomial parameters estimation

For each gene, we use the beta-binomial distribution to model the ASE counts:

$$P(n_{ij},\alpha_i,\beta_i)=\binom{n_{ij}}{m_{ij}}\frac{\mathrm{Beta}(\alpha_i+\beta_i,\ \beta_i+n_{ij}-m_{ij})}{\mathrm{Beta}(\alpha_i,\ \beta_i)},\ \alpha_i,\ \beta_i > 0$$

where m_{ij} and n_{ij} represent the reference allele and total counts for gene *i* in cell *j*, respectively, and α_i and β_i are the shape parameters of the Beta distribution, which can be described in terms of the Gamma function as $\mathrm{Beta}(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. α_i and β_i are estimated by maximizing the likelihood:

$$L=\binom{n_{ij}}{m_{ij}}\frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\,\Gamma(\beta_i)}\frac{\Gamma(\alpha_i+m_{ij})\,\Gamma(n_{ij}-m_{ij}+\beta_i)}{\Gamma(\alpha_i+\beta_i+n_{ij})}.$$

Maximum likelihood estimation (MLE) was performed with R (v4.3.1)[52] 'optim' function using BFGS quasi-Newton method with constraints. We obtain the beta-binomial mean μ and dispersion θ as $\hat{\mu}_i=\frac{\alpha_i}{\alpha_i+\beta_i}$ and $\hat{\theta}_i=\frac{1}{\alpha_i+\beta_i}$. Here, the mean $\hat{\mu}_i$ represents the gene-wise allelic ratio (mean success probability), while the dispersion $\hat{\theta}_i$ reflects the variability of the allelic ratio around its expected value ($\hat{\mu}_i$) [53].

Modeling the relationship between allelic dispersion and mean expression

To stabilize the excessive variability associated with low gene expression, we model gene-level dispersion, $\hat{\theta}_i$, as a function of gene expression (mean total reads across all cells, $\bar{n}_i$) using the local regression implemented in locfit [54] (v.1.5-9.8):

$$\log(\hat{\theta}_i)=f\log(\bar{n}_i)+\varepsilon_i.$$

The local regression model assumes that log-transformed gene-level dispersion follows a normal distribution $\log(\hat{\theta}_i)\sim N(0,1)$. All genes are included in the model to determine a representative mean-variance trend. The fitted dispersion values, $\bar{\theta}_i$, represent the amount of allelic variation shared between genes expressed at similar levels.

Bayesian shrinkage implementation

We adopt the weighted log-likelihood framework [33,34,48,] (l_{WL}) for each gene:

$$l_{WL}(\theta_i)=l(\hat{\theta}_i)+\delta\, l_{SH}(\bar{\theta}_i),$$

where δ is the weight given to the shared likelihood. $\bar{\theta}_i$ is treated as prior for θ_i under a Bayesian hierarchical framework. We modified the analytical solution for the posterior (shrunken) dispersion, which was derived for the double binomial distribution and extended to the beta-binomial distribution [48]:

$$\hat{\theta}_{WLi}=\frac{N-1}{N}\,\frac{\hat{\theta}_i+\delta\,\bar{\theta}_{SHi}}{1+\delta},$$

where $\bar{\theta}_{SHi}$ is the local dispersion calculated from local regression (not the global dispersion as specified in the analytical solution), $\hat{\theta}_i$ is the individual gene dispersion estimate, N is the effective degrees of freedom and δ is weight given to the shared likelihood. To determine the appropriate degrees of freedom and the weight parameter δ, we assume that dispersion $\hat{\theta}_i$ is approximately Gamma distributed with shape γ and rate ρ parameters expressed in terms of the degrees of freedom and common dispersion, which informs the prior distribution of $\hat{\theta}_i$:

$$\hat{\theta}_i|\bar{\theta} \sim \Gamma\left(\frac{N}{2}, \frac{2}{N-1}\bar{\theta}\right).$$

The gamma parameters for the posterior dispersion $\hat{\theta}_{WLi}$ derived from the weighted likelihood framework are given by:

$$\hat{\theta}_{WLi}|\bar{\theta} \sim \Gamma\left(\delta\frac{N}{2}, \delta\frac{2}{N-1}\bar{\theta}\right),$$

where the posterior distribution follows the same form as the prior distribution, differing only in terms of the shrinkage applied, as controlled by δ. Larger values of δ increase the influence of the shared dispersion $\bar{\theta}_{SHi}$, resulting in greater shrinkage of the posterior estimate toward the common dispersion.

We find estimates of Gamma hyperparameters (γ, ρ) by maximizing the likelihood function

$$L(\gamma, \rho) = (\gamma-1)\sum_{i=1}^{n}\log(x_i) - \frac{1}{\rho}\sum_{i=1}^{n}x_i - n\,\gamma\log(\rho) - n\log\left(\Gamma(\rho)\right),$$

where $\gamma = \delta\frac{N}{2}$, $\rho = \delta\,\frac{2}{N-1}\,\bar{\theta}_{SHi}$ and where $x_i = \hat{\theta}_{WLi_i}$. We used BFGS quasi-Newton method as implemented in the optim function in R (v4.3.1) [52] for MLE. Using T-cell data, based on estimates of the hyperparameters, we get δ = 53 and N = 30. We rounded δ = 50 and used these parameters to obtain shrunken dispersion values in both real and simulated data.

Treatment of genes with low allelic variation

ASPEN's adaptive Bayesian approach applies shrinkage selectively. Genes with very low dispersion estimates that pass a read count threshold (at least five cells with at least five mapped reads each) are not moderated towards the common trend.

From the local regression, we obtain the model residuals, which represent the portion of allelic variation not explained by gene expression

$$\varepsilon_i = \log(\hat{\theta}_i) - f\log(\bar{n}_i)\,.$$

We then calculate median absolute deviation[55]-squared, MAD^2, statistic as

$$MAD^2 = \text{Median}\ (|\varepsilon_i - \text{median}(\varepsilon_i)|)^2$$

and define lowly variable genes as those with $\varepsilon_i < MAD^2$. We repeat the dispersion trend modeling excluding lowly variable genes from Bayesian shrinkage and updating the shared estimates, $\bar{\theta}_i$. Original dispersion estimates, $\hat{\theta}_i$, are used for lowly variable genes.

Statistical framework for testing allelic imbalance

Following the initial estimation of beta-binomial parameters and dispersion modeling, ASPEN conducts a log-likelihood ratio test to identify genes with allelic imbalance by testing the following hypotheses:

$$H_0\text{: AR} \sim \text{BetaBinomial}\left(m_{ij}|n_{ij}, \mu=x, \hat{\theta}_{WLi}\right),$$

$$H_1\text{: AR} \sim \text{BetaBinomial}\left(m_{ij}|n_{ij}, \mu=\hat{\mu}_i, \hat{\theta}_{WLi}\right),$$

where x is the fixed allelic ratio which is determined based on empirical background (except for simulated data where x = 0.5, representing theoretical allelic balance), $\hat{\mu}_i$ is the allelic ratio and $\hat{\theta}_{WLi}$ is the shrunken dispersion that are specific to each gene. For genes with low allelic variation, we use a similar system of hypotheses, only substituting $\hat{\theta}_{WLi}$ with the empirical dispersion estimate, $\hat{\theta}_i$: H_0: AR ~ BB $\left(m_{ij}|n_{ij}, \mu=x, \hat{\theta}_i\right)$, vs. H_1: AR ~ BB$\left(m_{ij}|n_{ij}, \mu=\hat{\mu}_i, \hat{\theta}_i\right)$. The theoretical allelic ratio x is set to beta-binomial mean estimated using genes passing the minimum counts threshold (a minimum of five reads in at least five cells) after excluding genes located on the sex chromosomes and imprinted genes. Likelihood ratio test (LRT), using the χ^2 distribution,

with 1 degree of freedom was used to assess the best fit and p-values were adjusted for multiple testing using a false discovery rate (FDR) [56].

Statistical framework for testing changes in allelic variance

For each gene, we quantified allelic variation via a two-step process. First, we establish the baseline level of expected variation ($\bar{\theta}_{SHi}$) using a local regression model to capture the relationship between expression level and allelic variation. We then compared the stabilized variation ($\hat{\theta}_{WLi}$) against this baseline for each gene, while controlling for its mean allelic ratio ($\hat{\mu}_i$). By analyzing the fit of these models, the test distinguishes genes where ASE patterns vary between cells due to biological processes rather than technical noise.

To detect genes whose allelic variation is different from that expected for genes with similar expression, we formulated a likelihood ratio test based on the following hypotheses:

$$H_0\text{: AR} \sim \text{BetaBinomial}\left(m_{ij} \middle| n_{ij}, \mu=\hat{\mu}_i, \theta=\bar{\theta}_{SHi}\right),$$

$$H_1\text{: AR} \sim \text{BetaBinomial}\left(m_{ij} \middle| n_{ij}, \mu=\hat{\mu}_i, \theta=\hat{\theta}_{WLi}\right),$$

where $\bar{\theta}_{SHi}$ represents the predicted values from the local regression fit and where $\hat{\theta}_{WLi}$ is the stabilized (shrunken) dispersion. For this test, shrinkage is applied to all genes, including those with low variation, that met the minimum count threshold (five UMI counts in at least five cells). We calculated an empirical p-value under the null. Raw unnormalized data was used. We simulated reference allele counts then calculated the likelihood ratio for the two models. We computed the *p*-value by determining the proportion of simulated likelihood ratios that exceeded the observed ratio from the real data. We noted that the simulation-based *p*-value converged around 400 permutations with a tolerance of 0.001 (**Supp. Fig. 7**). Hence, we ran 500 permutations. *p*-values were adjusted using the FDR method[56].

Detecting dynamic changes

ASPEN implements two tests to assess temporal allelic dynamics: dyn-mean (for changes in mean allelic ratios) and dyn-var (for changes in allelic variance) across discrete groups. These tests require a label assignment for each cell barcode (eg. cell types, temporal ordering or low-dimensional embedding clusters). Continuous covariates (eg. pseudotime) must be supplied in the discretized form. Beta-binomial parameters estimation, dispersion modeling and moderation steps are performed at the global level (across all cells) and group level for each gene.

To detect changes in allelic patterns across distinct cell states, we compare the null hypothesis, which assumes that the shared mean and dispersion parameters can describe ASE counts distribution across the groups, against the alternative hypothesis, wherein the beta-binomial parameters are specific to each cellular or temporal state. In the dyn-mean test, the dispersion parameter (set to the stabilized group-level value) is kept constant between the two models to facilitate detection of changes driven by allelic mean. The null hypothesis states that the mean allelic ratio is the same across the groups:

$$H_0\text{: AR} \sim \text{BetaBinomial}\left(m_{ij} \middle| n_{ij}, \hat{\mu}_i, \bar{\theta}_{WLi}^{glob}\right),$$

where $\hat{\mu}_i$ denotes the shared-group mean allelic ratio and where $\bar{\theta}_{WLi}^{glob}$ is the global posterior (shrunken) dispersion, obtained by estimating beta-binomial parameters for the entire cell population for gene *i*. Under the alternative hypothesis, the mean allelic ratio varies among the groups. We define the alternative hypothesis as a mixture of beta-binomial distributions, where the number of individual components matches the number of groups and all components are given equal weights:

$$H_1\text{: AR} \sim \sum_{g=1}^{L} \text{BetaBinomial}\left(m_{gij} \middle| n_{gij}, \hat{\mu}_{gi}, \bar{\theta}_{WLi}^{glob}\right),$$

where $\hat{\mu}_{gi}$ is the mean allelic ratio for the *i*th gene in group *g* and where $\bar{\theta}_{WLi}^{glob}$ is the global posterior (shrunken) dispersion (set to be the same as in H_0).

In the dyn-var test, we maintain the estimates of the mean, μ_i, constant. The value of μ_i is set to the observed shared-group mean allelic ratio value (theoretical value). We compare the null hypothesis (that there are no changes in allelic dispersion):

$$H_0\text{: AR} \sim \text{BetaBinomial}\left(m_{ij} \middle| n_{ij}, \bar{\mu}^{glob}, \hat{\theta}_i\right)$$

where $\bar{\mu}^{glob}$ is the theoretical allelic ratio and $\hat{\theta}_i$ is the shared empirical dispersion estimate, with the alternative hypothesis (that allelic dispersion changes between the groups):

$$H_1\text{: AR} \sim \sum_{g=1}^{L} \text{BetaBinomial}\left(m_{gij} m_{gi} \middle| n_{gij}, \bar{\mu}^{glob}, \hat{\theta}_{gi}\right),$$

where $\bar{\mu}^{glob}$ is the theoretical allelic ratio and $\hat{\theta}_{gi}$ is the empirical dispersion estimate for the *i*th gene in group *g*. As the null and alternative models are nested for both dyn-mean and dyn-var tests, statistical significance was determined by the LRT. p-values are obtained based on χ^2_1 distribution and FDR control was achieved by the Benjamini and Hochberg's method [56].

Generation of simulated data

To simulate single-cell RNA-seq data, we estimated statistical parameters using the brain organoid dataset from male *C57BL/6J* × *SPRET/Ei* F1 hybrids (Bl6 × Spret) [57], which includes three cell types (n = 3,269 cells; data from clone four was used). The top 2,000 most variable genes identified using getTopHVGs from scater (v1.28.0) [58] were retained for parameter estimation using ZINB-WaVE model (zinbwave v.1.22.0) [59]. This was fitted to the total count matrix with the number of unknown cell-level covariates set to K = 2, while all other parameters were maintained at their default settings.

Reference allele counts were simulated using VGAM (v.1.1-9) [60] by drawing beta-binomial counts from the simulated total count matrix. To assess ASPEN's ability to detect genes with allelic imbalance, we generated multiple datasets with allelic ratios set to {0.10, 0.30, 0.35, 0.40, 0.42, 0.45, 0.46, 0.47, 0.48, 0.49, 0.50, 0.51, 0.52, 0.53, 0.54, 0.55, 0.57, 0.60, 0.65, 0.70, 0.90}. The dispersion estimates from real data were kept fixed for the respective gene in each simulation. The simulation for each group was repeated five times to improve technical reproducibility.

For simulating various pseudotime scenarios, we used the same total count matrix generated with zinbwave. This total count matrix was processed using Seurat (v. 4.0.5) [61]. We conducted principal component analysis (PCA) on log-normalized data and scaled it to 10000 counts using all the genes. The top 20 principal components (PCs) were selected for downstream Louvain clustering as implemented in the FindClusters function, and low-dimensional representations were obtained through uniform manifold approximation and projection (UMAP). We rescaled the PC1 coordinates between 0 and 1, using the resulting values as pseudotime. The pseudotime vector was divided into three equally sized bins (based on tertiles), and the cells were grouped according to this temporal assignment.

To simulate allelic counts where the mean allelic ratio varies across pseudotime bins while maintaining constant dispersion, we sampled beta-binomial counts with mean values of {0.3, 0.5, 0.7} and dispersion values of {1×10^{-3}, 1×10^{-3}, 1×10^{-3}}. For the second scenario, the mean allelic ratio was held constant across the bins with varying dispersion, and the mean values were set to {0.5, 0.5, 0.5} while the dispersion values were {1×10^{-3}, 0.05, 0.2}. In the control group, both parameters remained constant: the mean values were {0.5, 0.5, 0.5} and the dispersion was {1×10^{-3}, 1×10^{-3}, 1×10^{-3}}.

Bulk RNA-Seq data processing

Mouse fibroblast RNA-Seq fastq files to evaluate read mapping strategies were downloaded from GSE193728. To compare the distribution of genes with low dispersion between single-cell and bulk RNA-seq data, CD8+ T-cell data from male Bl6 × Spret hybrid data were obtained from GSE126770. The reads were trimmed via Trimmomatic [62] in pair-end mode and mapped to the combined genome index via STAR (v2.7.6a) [51]. To achieve full-length read alignment, the insertions, soft-end clipping and splicing modes were turned off (--outBAMcompression 6 --outFilterMultimapNmax 1 --outFilterMatchNmin 30 --alignIntronMax 1 --alignEndsType EndToEnd --outSAMattributes NH HI NM MD AS nM --scoreDelOpen -1000 --scoreInsOpen -1000). Reads mismatched (flagged as 'NM') and reads overlapping diverse strain-specific regions that were enriched for recently transposed long interspersed nuclear elements (LINEs) and long-terminal repeat (LTR) elements[63] were removed. Reads were counted via Rsubread (v2.4.3104)[64] with the following parameters: isPairedEnd=T, allowMultiOverlap=T, and countMultiMappingReads=F. To normalize for sequencing depth, size factors were estimated via total counts (sum of allelic counts) with DESeq2 (v1.30.193) [65] and used to adjust allelic counts.

scRNA data preprocessing

CD8+ T-cell data for naïve, day 7 and day 40 postinfection samples were obtained from the GSE164978 dataset (SRA files SRR13450127 and SRR13450128). Fastq files for mouse brain organoids generated from female Bl6 × Cast and male Bl6 × Spret F1 hybrids were downloaded from GSE268332. Files were mapped to the respective strain genomes via STARSolo (v.2.7.6a) [66] (--soloType CB_UMI_Simple --soloUMIlen 12 --soloCBstart 1 --soloCBlen 16 --soloUMIstart 17 --soloCBwhitelist <(zcat ${dir}/align/3 M3M-february-2018.txt.gz) --soloFeatures Gene --outSAMattributes NH HI nM AS CR UR CB UB GX GN sS sQ sM NM MD --outBAMcompression 6 --outFilterMultimapNmax 1 --outFilterMatchNmin 30 --alignIntronMin 20 --alignIntronMax 20000 --alignEndsType EndToEnd --outFilterMismatchNoverReadLmax 0.001 --scoreDelOpen -1000 --scoreInsOpen -1000). The total count matrix was produced by combining maternal and paternal allelic counts. Genes that were expressed in fewer than 10 cells were removed from further analysis.

Detecting allelic imbalance in the simulated data

We applied scDALI (v0.1.0) [25] in homogeneous (Hom) mode to the simulated counts to evaluate the global allelic imbalance. In ASPEN, we used the mean test following the dispersion modeling and shrinking steps. For both methods, we tested the observed allelic ratio against the null hypothesis of AR = 0.5 (in the absence of reference allele bias). P-values were adjusted for multiple testing via the false discovery rate (FDR). To assess the performance of the tests, we stratified genes by their expression levels (based on transcript per million (TPM) values) and allelic dispersion. We assigned the top 10% of genes to the high expression group (log(TPM) > 8), the next 50% to the medium expression group (log(TPM) < 8), and the remaining genes to the low expression group (log(TPM) < 3). The TPM values were obtained by running calculateTPM from scuttle (v1.10.2) [58] on the total count matrices. Genes were partitioned based on allelic dispersion as follows: genes whose dispersion was less than 0.005 were assigned to the low-dispersion group, and a threshold of 0.45 was selected to assign genes to either the medium- or high-dispersion group.

Single cell data allelic testing in CD8+ T cells

We conducted ASPEN allelic imbalance and variance tests on data from F1 mouse brain organoids and CD8 T cells. We examined 5,095 genes across 723 (naïve), 400 (day 7), and 582 (day 40) cells. The total count matrices for each time point were merged. Size factors were calculated using computeSumFactors from scran (v.1.28.2) [67] by pooling similar groups of cells together (we used the original time points for cluster assignment). The same scaling was applied to both the total and reference allele counts. Initial beta-binomial parameters were estimated, and dispersion modeling and shrinkage were conducted with raw counts. The expected allelic ratio was adjusted to the null allelic mean of 0.52 during the allelic imbalance test to reflect an alignment bias toward the reference genome (**Supp Fig. 8A**). Testing was then performed on the normalized counts, selecting genes that had at least 5 cells with at least 5 mapped reads. We ordered the datasets by time points (naïve, day 7, and day 40 post-infection) to assess dynamic changes in the allelic ratio distribution and used this as a discrete temporal assignment. P values for the ASPEN tests were FDR-adjusted.

Pseudobulked allelic testing in CD8+ T cells

Merged CD8+ T-cell counts were processed in Seurat (v. 4.0.5) [61]. We performed principal component analysis (PCA) on log-normalized data and scaled it to 10000 counts using all genes. The top 20 principal components (PCs) were selected for downstream Louvain clustering as implemented in the FindClusters function (resolution = 0.9), and low-dimensional representations were obtained through uniform manifold approximation and projection (UMAP). This process produced three clusters per time point, which were then used to aggregate the total counts as pseudoreplicates within each cell state (in lieu of the biological replicates) (**Supp Fig. 8B**). The same cell assignment was used to pool the reference allele counts. Size factors were estimated via total counts with DESeq2 (v1.30.1) [65] and were applied to scale the total and reference allele counts in parallel. Genes with a minimum of 10 reads across all clusters within the sample were retained for allelic imbalance testing. We checked for mapping bias in the pseudobulk dataset, revealing a slight skew toward the Bl6 allele, which was accounted for by adjusting the expected ratio (**Supp Fig. 8C**). P-values for the allelic imbalance test obtained from ASPEN were FDR-adjusted.

Analysis of mouse brain organoid single-cell RNA-seq data

To obtain initial beta-binomial parameters and shrunk dispersion values from CD8 T cells, we used the same procedure as for the male Bl6 × Spret hybrid data. For allelic imbalance testing, we adjusted the allelic ratio value under the null hypothesis to 0.54. Using beta-binomial shape parameters, we defined genes with $\alpha < 1$ OR $\beta < 1$ as monoallelically expressed.

Allelic imbalance in pseudobulked counts from mouse brain organoids

Single-cell Bl6 alleles and total counts from four female Bl6 × Cast F1 hybrids were aggregated by clonal origin to generate pseudobulk replicates. We estimated size factors using total counts with DESeq2 (v1.30.1) [65] and

applied them to scale the total and Bl6 allele counts. Genes with fewer than 10 total counts across the pseudoreplicates were excluded from further analysis.

Identification of random monoallelic expression

We identify RME expression patterns using the shape parameters (α and β) of the beta-binomial distribution. When either α or β falls below one, the distribution becomes skewed towards one allele. The hallmark of RME emerges when both α and β are less than one and are approximately equal, producing a U-shaped allelic density – indicating that individual cells express either the maternal or paternal allele but rarely both. We identified incomplete X chromosome inactivation in genes that exhibited biallelic expression based on the following criteria: mean allelic ratio between 0.25 and mean 0.75 in at least 15% of cells, with a minimum of five cells.

Functional enrichment analysis

Gene Ontology (GO) term enrichment analysis was performed in clusterProfiler (v4.8.3) [68] via the org.Mm.eg.db database (3.17.0) [69]. We used the enrichGO function to identify enriched biological processes in genes with low dispersion ($\hat{\theta}_i < 0.001$) in the T-cell data collected on day 7 post-LCMV infection. All genes with sufficient coverage from the same time point were used as the background for GO enrichment.

Availability of data and materials

ASPEN is available as an R package at https://github.com/ewonglab/ASPEN. Single-cell RNA-seq data for CD8 T cells from naïve, day 7, and day 40 LCMV Armstrong infection samples were downloaded from the GSE164978 dataset (SRA files SRR13450127 and SRR13450128) [29]. Bulk RNA-Seq data for CD8+ T cells 7 days after LCMV Armstrong infection were downloaded from GSE126770 [23]. Fastq files for bulk RNA-seq of mouse fibroblasts from four F1 hybrids (Bl6 × Cast, Bl6 × Molf, Bl6 × Pwk, and Bl6 × Spret) were obtained from GSE193728 [70]. Single-cell RNA-seq data for mouse brain organoids (Bl6 x Cast and Bl6 x Spret) were downloaded from GSE268332 [57]. Variant call files (release 5) between *C57BL/6J* and four wild-derived mouse strains (*CAST/Ei*, *MOLF/Ei*, *PWK/Ph*, and *SPRET/Ei*) were obtained from the Mouse Genome Project [48]. A list of recently transposed long interspersed nuclear elements (LINEs) and long-terminal repeat (LTR) elements in the strain-specific genomes was downloaded from the original publication [63].

Acknowledgments

We thank Dr. Daniel Medina-Cano and Mohammed Islam from the Memorial Sloan Kettering Cancer Center (MSKCC) for generating the mouse brain organoids used for method evaluation. We also thank Muqing Niu for the discussions. This project was undertaken with assistance from the resources and services of the National Computational Infrastructure (NCI), which is supported by the Australian Government.

Funding

V.P. is supported by an Australian Government Research Training Stipend Scholarship. T.V. is supported by startup funding provided by MSKCC, the Josie Robertson Investigator Program, and NIH grants (Project # 5R01NS126921-02, NIH P30 CA008748). E.S.W. is supported by an NHMRC Investigator Grant (GNT2009309), ARC Discovery Project (DP200100250), and a Snow Medical Fellowship.

Authors' contributions

E.S.W. conceived and supervised the project. V.P. wrote the code and performed method evaluation. T.V. supplied data for method evaluation. E.S.W. and V.P. wrote the manuscript. M.N. designed and wrote the code for the variance test. All the authors reviewed the manuscript.

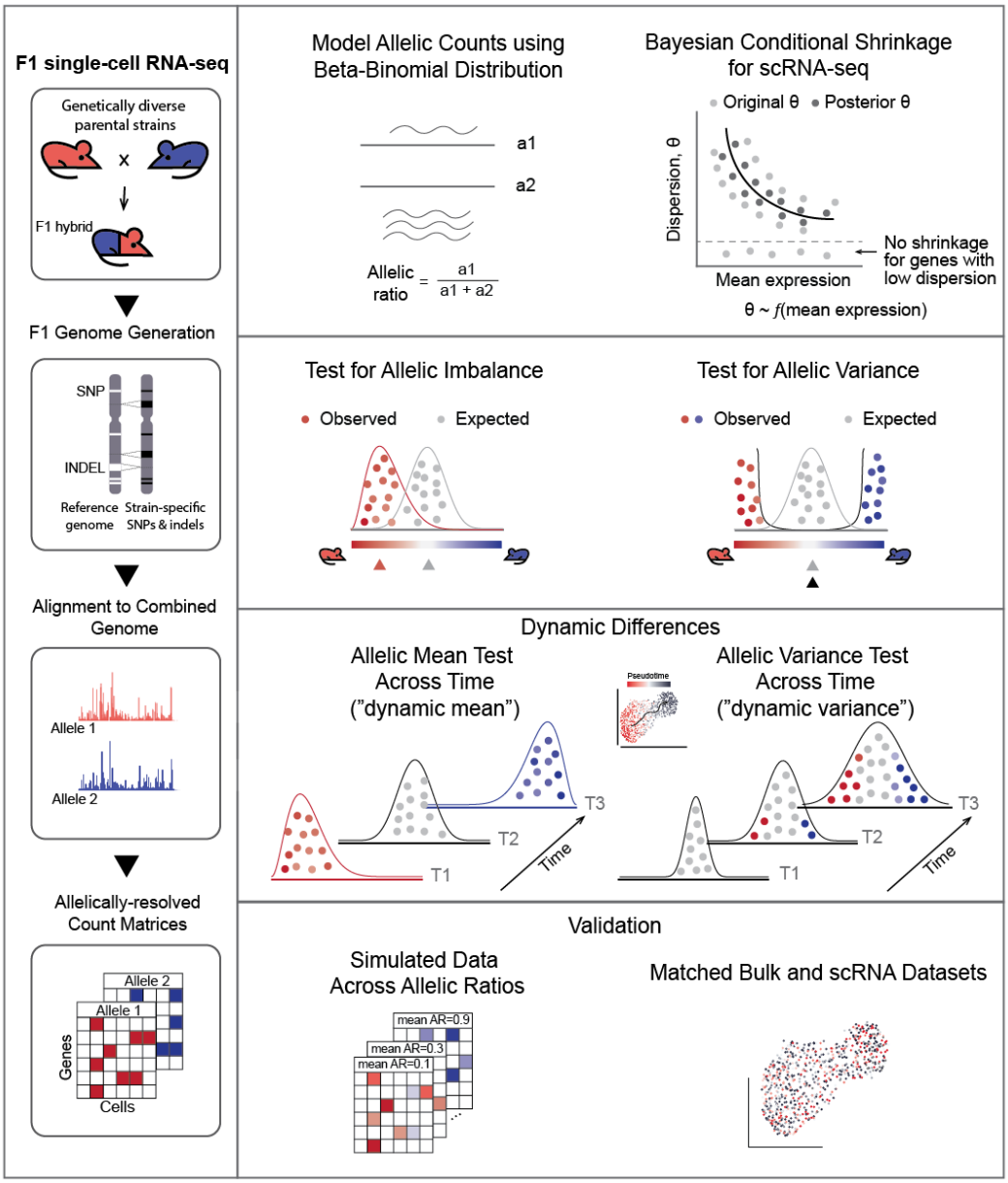

Figure 1. **Overview of ASPEN.** LHS: Raw sequencing reads are mapped to a combined genome created by concatenating the two parental genomes. The C57BL/6J (GRCm38) genome serves as the reference, while strain-specific genomes are produced by integrating variant information (SNPs and indels) into the C57BL/6J genome. RHS: ASPEN employs the Bayesian shrinkage method to stabilize dispersion estimates across genes with varying expression levels. Using weighted log-likelihood, original dispersion estimates are shrunk toward the common trend dispersion — the expected level of variation for genes exhibiting similar expression levels. The posterior (shrunken) dispersion estimates facilitate the detection of static differences and dynamic changes in the allelic ratio distribution.

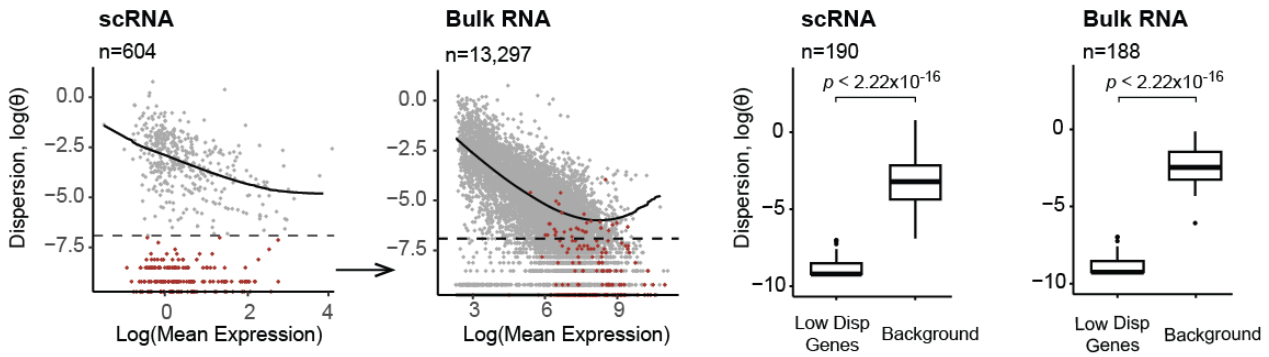

Figure 2. **ASPEN performs adaptive dispersion shrinkage tailored for single-cell ASE analysis.** Evaluating the concordance of low dispersion for genes in matched bulk and single-cell data. Genes with θ < 0.001 in the scRNA counts are marked in red (left panel). The same genes are indicated in red based on the dispersion estimation in the bulk RNA-seq data (right panel). Comparisons were conducted using T-cell data collected on day 7 after LCMV infection. Boxplots illustrate the mean dispersion estimation between low-dispersed genes and the background, a set of genes matched by gene expression (two-sided Wilcoxon rank-sum test).

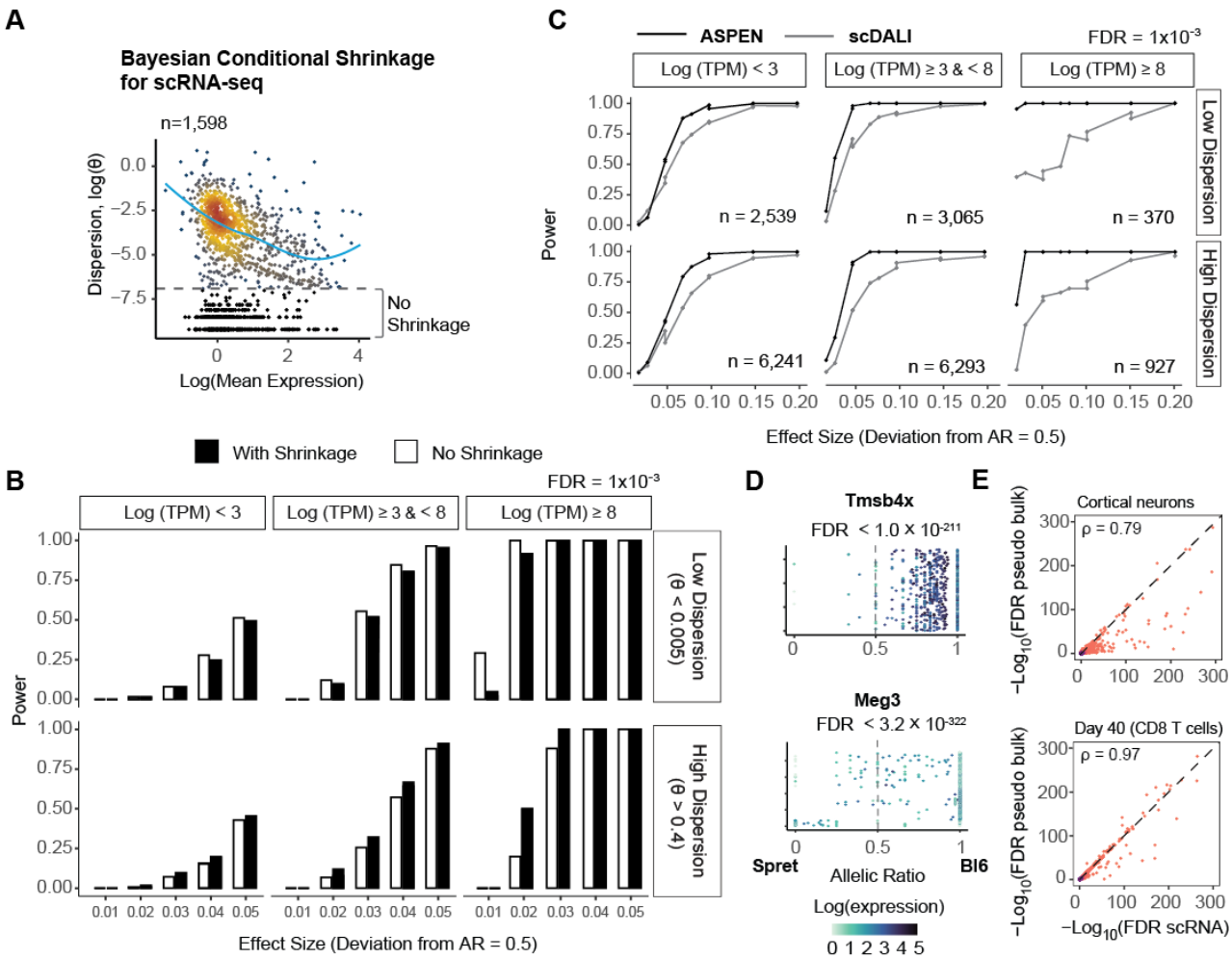

Figure 3. Evaluation and comparison of ASPEN's performance in detecting ASE imbalance. (A) Modeling the relationship between allelic ratio dispersion and mean gene expression revealed a population of genes with persistent low allelic dispersion (T-cell data from male Bl6 × Spret F1 hybrids, combining all cell states. The dashed line separates genes with low variation (in black), which are excluded from the shrinkage procedure. **(B)** Evaluating Bayesian (shrinkage applied to all genes) and non-Bayesian (no shrinkage applied) strategies for testing allelic imbalance. Using simulated data, power was estimated as the fraction of genes detected as allelically imbalanced at FDR = 0.001, stratified by dispersion level: low ($\theta < 0.005$) and high ($\theta > 0.4$); and gene expression levels – low (log(TPM) < 3), medium (log(TPM) ≥ 3 & log(TPM) < 8) and high (log(TPM) ≥ 8). **(C)** Performance comparison between ASPEN and scDALI using simulated data. Line plots illustrate TPRs calculated at FDR = 0.001 for genes categorized by gene expression and allelic dispersion levels (low dispersion – $\theta < 0.005$, medium – $\theta \geq 0.005$ & $\theta < 0.4$, and high – $\theta \geq 0.4$) across groups with varying degrees of deviation from the mean AR = 0.5. Only datasets simulated with a mean AR of 0.5 ±± 0.2 are displayed. No performance differences were identified for groups with deviations from balanced allelic expression > 0.2. **(D)** Examples of the allelic ratio distributions of X-linked genes (*Tmsb4x*) and imprinted genes (*Meg3*). **(E)** Correlation of ASPEN-$\log_{10}$FDR values generated by running ASPEN mean test on the single-cell counts (x-axis) and pseudobulked counts (y-axis). Analyses were conducted in mouse cortical neurons (n = 972) from female Bl6 × Cast F1 hybrids and in $CD8^+$ T cells at day 40 (n = 768) following LCMV infection in male Bl6 × Spret F1 hybrids.

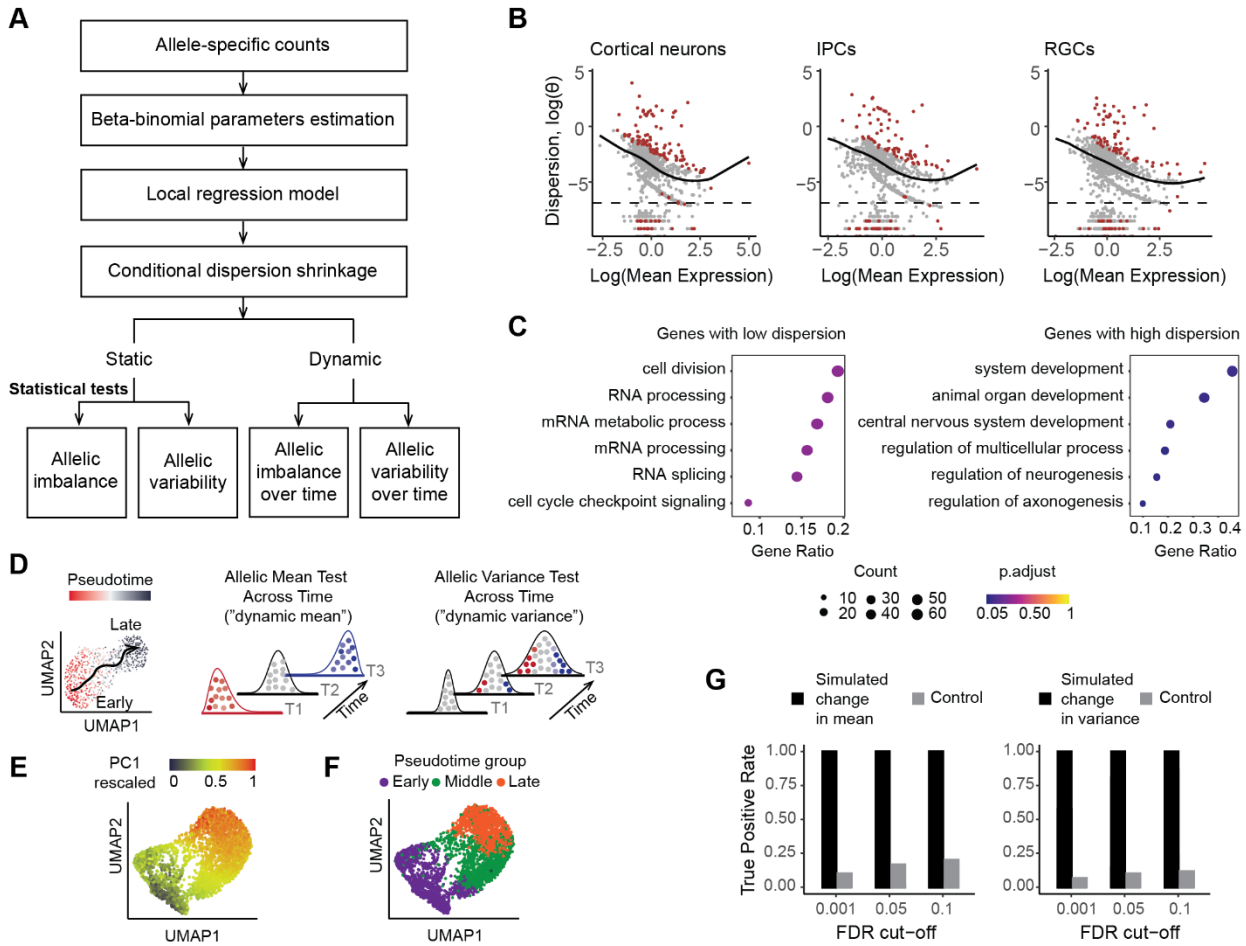

Figure 4. Detect allelic variation and dynamic changes in ASE with ASPEN. (A) Overview of the ASPEN analytical workflow. The main analysis steps involve the initial estimation of the allelic count distribution parameters, modeling allelic count dispersion as a function of gene expression, and conditional shrinkage toward the common dispersion estimates. **(B)** Results of the ASPEN var test in brain organoids data. Genes with significant deviation from expected variation levels are in red (FDR < 0.05): cortical neurons n = 140 (out of 973), IPCs n = 117 (out of 857) and RGCs n = 117 (out of 857). **(C)** Top six Gene Ontology (GO) functional terms enriched in low dispersed genes (below the trend line) and highly dispersed genes (above the trend line). Genes below the dashed line were excluded from the trend modelling. **(D)** Schematic representation of the ASPEN strategy used to detect changes in allelic mean or variance over time. **(E)** UMAP visualization of the simulated dataset used for testing. The cells are colored based on the rescaled PC1 coordinates representing temporal ordering. **(F)** UMAP visualization of cell assignments to discrete pseudotime bins: early, middle, and late (created by dividing the PC1 vector into tertiles). **(G)** TPR values calculated at multiple FDR thresholds: the dyn-mean test performance on the data simulated with changes in allelic mean vs. control (left plot) and dyn-var test performance on the data simulated with changes in allelic variance vs. control (right plot).

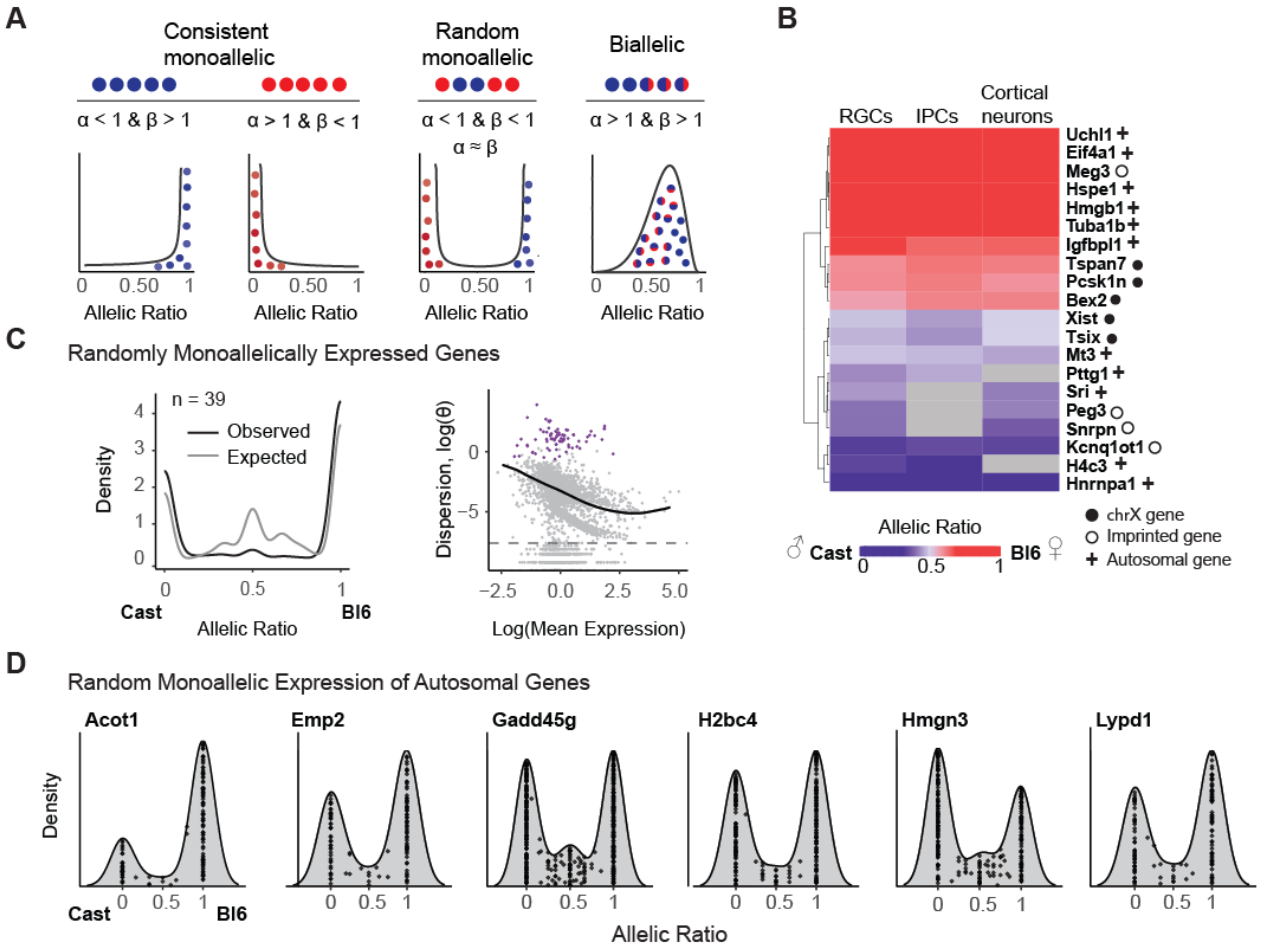

Figure 5. Analysis of ASE dynamics in early neurogenesis. (A) Scenarios where allelic variance deviates from expected levels and the corresponding beta-binomial distribution shape parameters α and β. **(B)** The heatmap displays the mean allelic ratios across cells for genes with monoallelic expression. Genes were classified as monoallelic if their estimated beta-binomial α or β values were less than one. Results are from brain organoid data derived from female Bl6 × Cast F1 hybrids using three cell types from the early neurodevelopmental pathway. Genes with allelic imbalance detected across all three cell types, with a mean expression above 1 in at least one, were selected for plotting (n = 18). Rows were hierarchically clustered using complete linkages. **(C)** Evaluating the expected versus observed dispersion distributions of genes with random monoallelic expression (RME) (total n = 39). The expected allelic ratio is calculated as the ratio of the simulated reference allele count to the total count. The expected gene set comprises genes without evidence of monoallelic expression, matched by expression to RME genes. The departure of the RME genes (in purple) from the common dispersion trend is visualized in the mean-variance plot. **(D)** Allelic ratio distribution plots for a selection of autosomal RME genes.

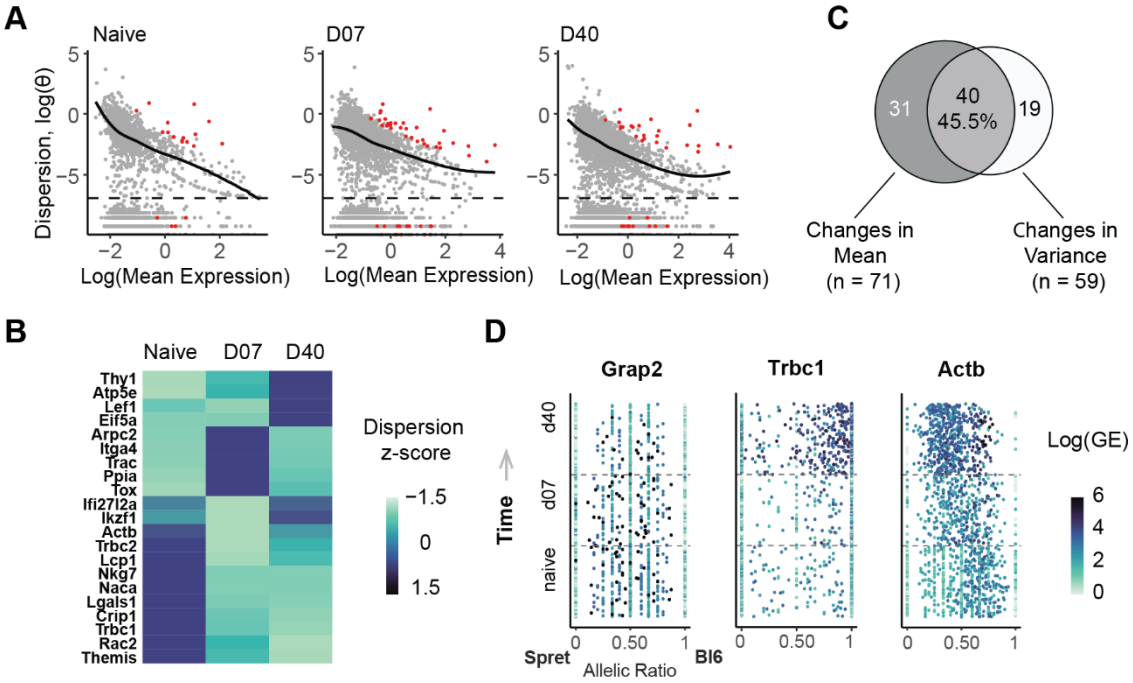

Figure 6. **Evaluating allelic dynamics during T-cell activation.** (**A**) Genes with significant deviations from the expected dispersion levels (in red) determined by the ASPEN var test (FDR < 0.05): Naïve n = 21, day 7 n = 50, day 40 n = 41. **(B)** Heatmap of the dispersion z-scores for the genes with ASPEN FDR var < 0.05 in at least one cell state (n = 21). Z-scores represent the empirical dispersion estimates for each gene scaled by column. **(C)** Venn diagram showing the overlap between the genes with dynamic changes in mean (ASPEN dyn-mean: FDR < 0.05) and variance (ASPEN dyn-var: FDR < 0.05). **(D)** Visualization of dynamic changes in allelic distribution during T-cell activation. ASPEN tests categorize genes into three groups: those changing only in variance and not in mean (*Grap2*, FDR dyn-var = 2.21×10^{-48}, FDR dyn-mean = 1); those with changes in both mean AR and variance (*Trbc1*, FDR dyn-var = 4.82×10^{-69}, FDR dyn-mean = 0.02); and those changing only in mean and not in variance (*Actb*, FDR dyn-var = 1, FDR dyn-mean = 3.86×10^{-63}).

686 **References**
687

688 1. Babak, T. *et al.* Global Survey of Genomic Imprinting by Transcriptome Sequencing.
689 *Current Biology* **18**, 1735–1741 (2008).

690 2. Goncalves, A. *et al.* Extensive compensatory cis-trans regulation in the evolution of mouse
691 gene expression. *Genome Res.* **22**, 2376–2384 (2012).

692 3. Andergassen, D., Smith, Z. D., Kretzmer, H., Rinn, J. L. & Meissner, A. Diverse epigenetic
693 mechanisms maintain parental imprints within the embryonic and extraembryonic lineages.
694 *Developmental Cell* **56**, 2995-3005.e4 (2021).

695 4. Reinius, B. *et al.* Analysis of allelic expression patterns in clonal somatic cells by single-cell
696 RNA–seq. *Nat Genet* **48**, 1430–1435 (2016).

697 5. Keniry, A. *et al.* BAF complex-mediated chromatin relaxation is required for establishment
698 of X chromosome inactivation. *Nat Commun* **13**, 1658 (2022).

699 6. Coolon, J. D., McManus, C. J., Stevenson, K. R., Graveley, B. R. & Wittkopp, P. J. Tempo
700 and mode of regulatory evolution in Drosophila. *Genome Research* **24**, 797 (2014).

701 7. Khoueiry, P. *et al.* Uncoupling evolutionary changes in DNA sequence, transcription factor
702 occupancy and enhancer activity. *eLife* **6**, e28440 (2017).

703 8. Wong, E. S. *et al.* Interplay of cis and trans mechanisms driving transcription factor binding
704 and gene expression evolution. *Nature Communications* (2017).

705 9. Hu, C. K. *et al. cis*-Regulatory changes in locomotor genes are associated with the
706 evolution of burrowing behavior. *Cell Reports* **38**, 110360 (2022).

707 10. Wittkopp, P. J., Haerum, B. K. & Clark, A. G. Evolutionary changes in cis and trans gene
708 regulation. *Nature* **430**, 85–88 (2004).

709 11. Wang, X., Werren, J. H. & Clark, A. G. Allele-Specific Transcriptome and Methylome
710 Analysis Reveals Stable Inheritance and Cis-Regulation of DNA Methylation in Nasonia. *PLoS*
711 *Biology* **14**, e1002500 (2016).

712 12. Rossi, M. *et al.* Visual mate preference evolution during butterfly speciation is linked to
713 neural processing genes. *Nat Commun* **11**, 4763 (2020).

714 13. Springer, N. M. & Stupar, R. M. Allele-Specific Expression Patterns Reveal Biases and
715 Embryo-Specific Parent-of-Origin Effects in Hybrid Maize. *The Plant Cell* **19**, 2391 (2007).

716 14. Zhang, X. & Borevitz, J. O. Global Analysis of Allele-Specific Expression in Arabidopsis
717 thaliana. *Genetics* **182**, 943 (2009).

718 15. Bell, G. D. M., Kane, N. C., Rieseberg, L. H. & Adams, K. L. RNA-Seq Analysis of Allele-
719 Specific Expression, Hybrid Effects, and Regulatory Divergence in Hybrids Compared with Their
720 Parents from Natural Populations. *Genome Biology and Evolution* **5**, 1309–1323 (2013).

721 16. Tirosh, I., Reikhav, S., Levy, A. A. & Barkai, N. A Yeast Hybrid Provides Insight into the
722 Evolution of Gene Expression Regulation. *Science* **324**, 659–662 (2009).

723 17. Bueker, B. *et al.* Meiotic recombination in the offspring of Microbotryum hybrids and its
724 impact on pathogenicity. *BMC Evolutionary Biology* **20**, 123 (2020).

725 18. Tsouris, A., Brach, G., Schacherer, J. & Hou, J. Non-additive genetic components
726 contribute significantly to population-wide gene expression variation. *Cell Genomics* **4**, 100459
727 (2024).

728 19. Jiang, Y., Zhang, N. R. & Li, M. SCALE: modeling allele-specific gene expression by single-
729 cell RNA sequencing. *Genome Biology* **18**, 74 (2017).

730 20. Larsson, A. J. M. *et al.* Genomic encoding of transcriptional burst kinetics. *Nature* **565**,
731 251–254 (2019).

732 21. Pacini, G. *et al.* Integrated analysis of Xist upregulation and X-chromosome inactivation
733 with single-cell and single-allele resolution | Nature Communications. *Nature Communications* **12**,
734 3638 (2021).

735 22. Lentini, A. *et al.* Elastic dosage compensation by X-chromosome upregulation. *Nat*
736 *Commun* **13**, 1854 (2022).

737 23. van der Veeken, J. *et al.* Natural Genetic Variation Reveals Key Features of Epigenetic and
738 Transcriptional Memory in Virus-Specific CD8 T Cells. *Immunity* **50**, 1202-1217.e7 (2019).

739 24. Mu, W. *et al.* Airpart: interpretable statistical models for analyzing allelic imbalance in
740 single-cell datasets. *Bioinformatics* **38**, 2773–2780 (2022).

741 25. Heinen, T. *et al.* scDALI: modeling allelic heterogeneity in single cells reveals context-
742 specific genetic regulation. *Genome Biology* **23**, 8 (2022).

743 26. Saukkonen, A., Kilpinen, H. & Hodgkinson, A. Highly accurate quantification of allelic gene
744 expression for population and disease genetics. *Genome Res* **32**, 1565–1572 (2022).

745 27. Robinson, M. D. & Smyth, G. K. Moderated statistical tests for assessing differences in tag
746 abundance. *Bioinformatics* **23**, 2881–2887 (2007).

747 28. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor
748 RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 4288–4297
749 (2012).

750 29. Pritykin, Y. *et al.* A unified atlas of CD8 T cell dysfunctional states in cancer and infection.
751 *Molecular Cell* **81**, 2477-2493.e10 (2021).

752 30. Zhao, H. *et al.* Genome-wide fitness gene identification reveals Roquin as a potent
753 suppressor of CD8 T cell expansion and anti-tumor immunity. *Cell Reports* **37**, 110083 (2021).

754 31. Metzger, B. P. H., Yuan, D. C., Gruber, J. D., Duveau, F. & Wittkopp, P. J. Selection on
755 noise constrains variation in a eukaryotic promoter. *Nature* **521**, 344–347 (2015).

756 32. Eckersley-Maslin, M. A. *et al.* Random Monoallelic Gene Expression Increases upon
757 Embryonic Stem Cell Differentiation. *Developmental Cell* **28**, 351–365 (2014).

758 33. Arpi, M. N. T. & Simpson, T. I. SFARI genes and where to find them; modelling Autism
759 Spectrum Disorder specific gene expression dysregulation with RNA-seq data. *Sci Rep* **12**, 10158
760 (2022).

761 34. Collins, S. C. *et al.* Large-scale neuroanatomical study uncovers 198 gene associations in
762 mouse brain morphogenesis. *Nat Commun* **10**, (2019).

763 35. Naik, H. C. *et al.* Semicoordinated allelic-bursting shape dynamic random monoallelic
764 expression in pregastrulation embryos. *iScience* **24**, 102954 (2021).

765 36. Balasooriya, G. I. & Spector, D. L. Allele-specific differential regulation of monoallelically
766 expressed autosomal genes in the cardiac lineage. *Nat Commun* **13**, 5984 (2022).

767 37. Deng, Q., Ramsköld, D., Reinius, B. & Sandberg, R. Single-Cell RNA-Seq Reveals
768 Dynamic, Random Monoallelic Gene Expression in Mammalian Cells. *Science* **343**, 193–196
769 (2014).

770 38. Marion-Poll, L. *et al.* Locus specific epigenetic modalities of random allelic expression
771 imbalance | Nature Communications. *Nat Commun* **12**, (2021).

772 39. Gendrel, A.-V. *et al.* Developmental Dynamics and Disease Potential of Random
773 Monoallelic Gene Expression. *Developmental Cell* **28**, 366–380 (2014).

774 40. Kravitz, S. N. *et al.* Random allelic expression in the adult human body. *Cell Reports* **42**,
775 111945 (2023).

776 41. Kim, J. K., Kolodziejczyk, A. A., Ilicic, T., Teichmann, S. A. & Marioni, J. C. Characterizing
777 noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic
778 expression | Nature Communications. *Nat Commun* **6**, (2015).

779 42. Zhou, X. *et al.* Epigenetic modifications are associated with inter-species gene expression
780 variation in primates. *Genome Biology* **15**, 547 (2014).

781 43. Foreman, R. & Wollman, R. Mammalian gene expression variability is explained by
782 underlying cell state. *Molecular Systems Biology* **16**, e9146 (2020).

783 44. Mott, R. *et al.* The Architecture of Parent-of-Origin Effects in Mice. *Cell* **156**, 332–342
784 (2014).

785 45. Burkart, V. *et al.* Transcriptional bursts and heterogeneity among cardiomyocytes in
786 hypertrophic cardiomyopathy. *Front. Cardiovasc. Med.* **9**, (2022).

787 46. Delbridge, A. R. D. *et al.* Loss of *p53* Causes Stochastic Aberrant X-Chromosome
788 Inactivation and Female-Specific Neural Tube Defects. *Cell Reports* **27**, 442-454.e5 (2019).

789 47. Becskei, A., Kaufmann, B. B. & van Oudenaarden, A. Contributions of low molecule
790 number and chromosomal positioning to stochastic gene expression. *Nat Genet* **37**, 937–944
791 (2005).

792 48. Keane, T. M. *et al.* Mouse genomic variation and its effect on phenotypes and gene
793 regulation. *Nature* **477**, 289–294 (2011).

794 49. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008
795 (2021).

796 50. Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids*
797 *Res* **34**, D590-598 (2006).

798 51. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

799 52. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for
800 Statistical Computing, 2023).

801 53. Griffiths, D. A. Maximum Likelihood Estimation for the Beta-Binomial Distribution and an
802 Application to the Household Distribution of the Total Number of Cases of a Disease. *Biometrics*
803 **29**, 637–648 (1973).

804 54. Loader, C. *Locfit: Local Regression, Likelihood and Density Estimation.* (1999).

805 55. Hampel, F. R. The Influence Curve and Its Role in Robust Estimation. *Journal of the*
806 *American Statistical Association* **69**, 383–393 (1974).

807 56. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and
808 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* 289–300 (1995).

57. Medina-Cano, D. *et al.* A Mouse Organoid Platform for Modeling Cerebral Cortex Development and Cis-Regulatory Evolution in Vitro. *Preprint* 2024.09.30.615887 (2024) doi:10.1101/2024.09.30.615887.

58. McCarthy, D. J., Campbell, K. R., Lun, A. T. L. & Wills, Q. F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179–1186 (2017).

59. Risso, D., Perraudeau, F., Gribkova, S., Dudoit, S. & Vert, J.-P. A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat Commun* **9**, (2018).

60. Yee, T. *Vector Generalized Linear and Additive Models: With an Implementation in R | SpringerLink*. (2015).

61. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

62. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

63. Lilue, J., Shivalikanjli, A., Adams, D. J. & Keane, T. M. Mouse protein coding diversity: What's left to discover? *PLOS Genetics* (2019).

64. Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* **47**, e47 (2019).

65. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

66. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. 2021.05.05.442755 Preprint at (2021).

67. Lun, A. T. L., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122 (2016).

68. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141 (2021).

69. Carlson, M. & Falcon, S. *Org. Mm. Eg. Db: Genome Wide Annotation for Mouse*. vol. 3 (R package, 2019).

70. Yang, M. G., Ling, E., Cowley, C. J., Greenberg, M. E. & Vierbuchen, T. Characterization of sequence determinants of enhancer function using natural genetic variation. *eLife* **11**, e76500 (2022).