arXiv:2602.03077v1 [stat.ME] 3 Feb 2026

函数效应的经验贝叶斯收缩：及其在动态表达数量性状位点分析中的应用

Ziang Zhang 人类遗传学系，芝加哥大学，美国伊利诺伊州芝加哥

Peter Carbonetto 人类遗传学系，芝加哥大学，美国伊利诺伊州芝加哥

Matthew Stephens 统计学与人类遗传学系，芝加哥大学，美国伊利诺伊州芝加哥

2026年2月4日

摘要

我们提出函数自适应收缩（FASH），这是一种对观测单位进行联合分析的经验贝叶斯方法，其中每个单元估计一个效应函数，对应于连续条件变量的若干取值。本文的思想源于动态表达数量性状位点（eQTL）研究，旨在刻画遗传效应对基因表达如何随时间或其他连续条件而变化。FASH 将一大类通过线性微分算子定义的高斯过程纳入经验贝叶斯收缩框架，从而实现自适应平滑，并在各单位之间借用信息。这改进了效应函数的估计，并支持规范的假设检验，从而能够直接计算显著性度量，例如局部错误发现率和错误符号率。 为鼓励保守推断，我们提出一种简单的先验调整方法；该方法具有理论保证，并可更广泛地用于其他经验贝叶斯方法。 我们通过重新分析来自诱导多能干细胞的心肌细胞分化动态 eQTL 数据，展示了 FASH 的优势。FASH 识别出新的动态表达数量性状位点，揭示了多样的时间效应模式，并且与原始分析相比提高了检验效能。更广泛地说，FASH 为函数型数据的联合分析提供了灵活的统计框架，其应用不限于基因组学。为便于在动态 eQTL 研究及其他设定中使用 FASH，我们提供了配套的 R 包，网址为 https: //github.com/stephenslab/fashr。

关键词：多重假设检验；高斯过程；贝叶斯推断；自适应收缩；函数型数据分析；表达数量性状位点

1 引言

扰动一个系统并测量由此产生的变化，是理解系统的经典科学方法（DeRisi 等人，1997；Heller 等人，1997；Hughes 等人，2000）。因此，现代基因组学实验常常会测量许多基因组单元在各种扰动或处理下的行为。其中一些实验会测量行为如何随某个潜在连续变量而变化。例如，近期的表达数量性状位点（eQTL）和等位基因特异性表达（ASE）研究，测量遗传变异对基因表达的效应如何随时间变化，或如何随氧气水平等连续条件变化（Cuomo 等人，2020；Elorbany 等人，2022年；Francesconi 和 Lehner，2014；Gutierrez-Arcelus 等人，2020；Kang 等人，2023；Nathan 等人，2022年；Soskic 等人，2022年；Strober等，2019）。我们将这类研究称为“动态 eQTL 研究”。这类研究通常会在连续条件的多个设定下，对基因组单元（基因-变异对）的 eQTL 效应进行含噪测量。动态 eQTL 研究的目标包括估计潜在的 eQTL 效应函数（例如，刻画 eQTL 效应如何随时间变化），以及识别哪些 eQTL 效应偏离了某种“零假设”或“基线”行为。零假设或基线可以定义为效应随时间没有变化，或随时间线性变化。

动态 eQTL 研究涉及海量数据，会从数千乃至数百万个基因组单元中产生带有噪声的测量值。因此，与通常采用的逐个单独处理各单元的做法相比，通过在不同单元之间共享信息，可以显著改善统计推断。经验贝叶斯（EB）（Efron，2008，2009）为此提供了一个强大的框架。经验贝叶斯方法利用所有单元学习一个共同的先验分布，该分布通常以零假设或基线为中心，

然后通过将每个单元的效应估计值向该零假设收缩，改善其估计结果。经验贝叶斯的一个优势在于，收缩程度具有自适应性，取决于先验分布（例如，零假设单元较多的实验会得到收缩程度更大的先验分布），也取决于效应估计值的标准误（估计值的噪声越大，收缩程度越强）（Stephens，2017）。经验贝叶斯方法及其便捷的软件实现已广泛推出，并越来越多地用于涉及单一条件（Stephens，2017；Willwerscheid 等人，2025）或多个条件（Urbut 等，2019；Liu 等，2024）的研究。

在这项工作中，我们开发了新的经验贝叶斯方法和软件，用于估计定义在时间等连续取值空间上的效应函数。为了对效应函数建模，我们采用高斯过程（GP）（MacKay，2003），这可能是为函数定义先验分布最简单且应用最广泛的方法。我们利用一类高斯过程，并将其称为“L-GP 族”，它能够编码多种泛函形状和行为。该族中的成员由单个标量参数 σ 定义，用以量化其偏离零假设/基线模型空间的程度（例如常函数空间或线性函数空间）（Lindgren 和 Rue，2008；Yue 等人，2014；Zhang 等人，2024，2025）。我们展示了如何将 L-GP 族与自适应收缩的思想（Stephens，2017）相结合，从而自适应地收缩函数。我们将这些新方法称为“泛函自适应收缩</style”（FASH）。

这些思路的结合还提供了一种灵活的方法来对效应函数进行假设检验；也就是说，评估效应函数是否显著偏离给定的基线模型。通过 FASH 对效应函数进行假设检验，与动态 eQTL 研究中采用的临时拼凑的方法截然不同，而且具有显著优势（Cuomo 等人，2020；Elorbany 等人，2022；Francesconi 和 Lehner，2014；Gutierrez-Arcelus 等人，2020；Kang 等人，2023；Nathan 等人，2022；Soskic

等人，2022；Strober等，2019）。例如，Soskic 等人（2022）通过比较包含和不包含基因型 × 时间交互项的基因表达模型来进行假设检验。这种方法的问题在于，它未能考虑其他可能值得关注的时间依赖效应，而这些效应随时间的变化可能是非线性的。（认识到这一局限后，Soskic 等人（2022）还比较了包含和不包含二次交互项的模型，即基因型 × 时间^2, ，但这仍然对 eQTL 效应随时间变化的方式作出了严格限制。）相比之下，FASH 中的假设检验不作此类假设（参见第 2 节的示例）。因此，我们的方法更加灵活，检验效能也更高。此外，与动态 eQTL 研究中采用的方法相比，FASH 还有另一项优势：以往的方法需要使用个体水平数据拟合模型，而 FASH 使用汇总统计量；当个体水平数据无法获取时（例如，受健康保险流通与责任法案要求或其他数据隐私问题的限制），有时可以获得汇总统计量。因此，FASH 也适用于只能获取汇总统计量的设定。

最后，我们指出，我们的新方法还解决了经验贝叶斯方法一个广为人知却常未得到处理的局限。经验贝叶斯方法通常要求先验分布准确刻画零假设和备择假设，才能实现校准推断。然而，在实践中，准确指定备择假设下的先验往往并不现实，因此所得推断可能失去校准性，或变得反保守。为帮助确保推断保持保守，我们提出了一种基于贝叶斯因子（贝叶斯因子）的先验调整方法。尽管本文仅针对 FASH 实现了这一思路，但我们指出，这是一种通用技术，也有望用于其他基于经验贝叶斯的收缩方法。本文介绍的所有方法均已在一个 R 包中实现， `fashr`（“R 中的函数自适应收缩”），该包可通过 https://github.com/stephen slab/fashr 获取。据我们所知，以往的动态 eQTL 研究并未发布

用于其假设检验流程的专用软件工具，因此 `fashr` 是首个面向动态 eQTL 研究的集成软件。此外，FASH 方法及其软件实现具有相当的通用性，因此我们预计 FASH 有望应用于动态 eQTL 研究以外的其他数据集。该 R 包提供了详尽的文档和易于使用的界面，并附有一个示例，展示了如何在动态 eQTL 数据集中按推荐方式使用 `fashr`。

1.1 论文结构

本文其余部分安排如下。第 2 节通过一个例子说明向基线模型收缩，以及如何通过在观测单位之间借用信息实现自适应收缩。第 3 节介绍我们的 FASH 方法，以及一种用于保守假设检验的基于贝叶斯因子的调整。第 4 节应用 FASH 重新分析 Strober等人（2019）的动态 eQTL 数据集，该数据集涵盖诱导多能干细胞分化为心肌细胞的 16 天过程，包含逾一百万个基因-变异对。这项分析发现了新的动态表达数量性状位点，揭示了多样的时间效应模式，并且比原始研究中的参数交互分析更准确地刻画了效应函数；原始研究采用了更严格的假设。最后，第 5 节总结我们的贡献，概述当前方法的局限性，并探讨将其拓展至其他领域的可能性。

2 一个启发性示例

为阐明我们方法的核心思想，我们首先给出一个动态 eQTL 应用中的示例，下面将对该应用进行更详细的介绍。在此示例中，我们重点关注单个基因-变异对的数据和结果：基因 *FZD6* 和单核苷酸多态性 rs28392906。图 1 各面板中的黑点显示了不同时间点的 eQTL效应估计

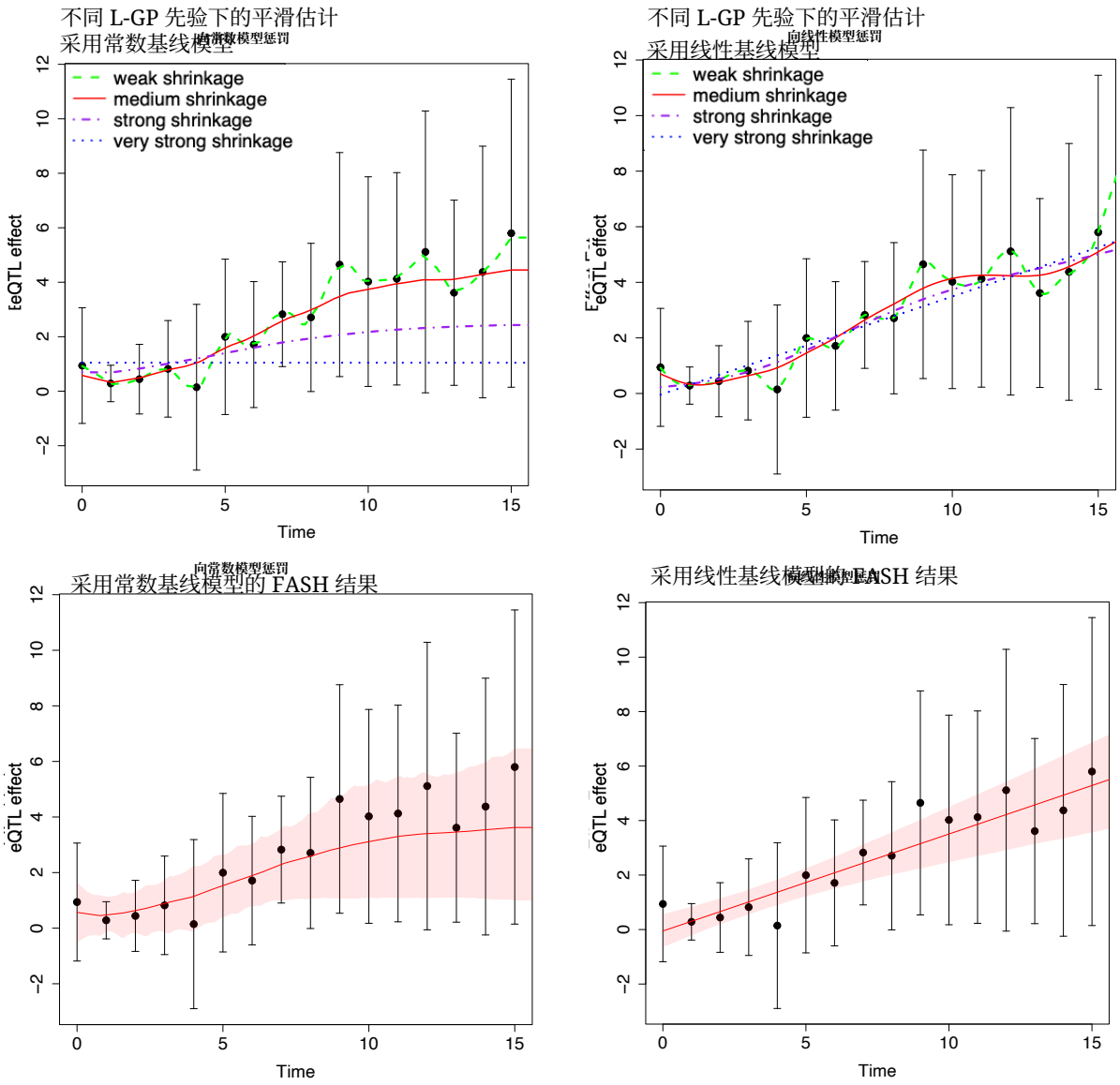


图 1：FASH 分析动态 eQTL 数据集的示例。所有面板中，原始 eQTL 效应估计值以黑点表示，竖直误差线表示 ±2标准误。上方两个面板展示平滑估计，即基于 L-高斯过程方法得到的后验均值；这些不同的估计值是通过向常数和线性基线模型进行不同程度的收缩得到的（分别对应左侧和右侧面板）。下方两个面板概述了所提出的 FASH 方法的结果，其中分别采用常数和线性基线模型（分别对应左侧和右侧面板）。常数和线性基线模型分别使用具有 $L = D^1$和 $L = D^2$的 L-高斯过程定义（见第 3.4 节）。后验均值效应函数以红色实线表示；阴影区域表示 95% 可信区间。

对于该基因-变异对。eQTL 效应的标准误以垂直误差线表示。

统计分析的一部分是通过将这些“嘈杂”的 eQTL 估计值向基线模型收缩，从而提高其准确性。FASH 中的基线模型假设，在该模型下，真实效应函数随时间平滑变化。不过，具体选择哪种基线模型取决于我们关注的科学问题。这里我们考虑两个问题：(i) eQTL 效应是否具有动态性——也就是说，它是否随时间变化？(ii) eQTL 效应随时间的变化是否偏离线性趋势——也就是说，是否随时间呈现非线性变化？与这些基线模型对应的零假设为：(i) 效应函数是常数；(ii) 效应函数是线性的。我们的方法可以将效应函数向上述任一基线模型收缩，也可以向其他基线模型收缩。

首先，为了说明效应估计值的平滑过程，我们展示了应用L-高斯过程模型并采用不同收缩程度的结果，其中基线模型分别为常数模型（图 1，左上图）和线性模型（右上图）。从这些图中可以看出，平滑估计会随收缩程度而大幅变化。在收缩程度最高时，平滑估计会变为一个常数函数（对各时间点取平均），或加权线性回归得到的一条直线。在收缩程度最低时，平滑估计则会变成一条精确插值所有数据点的曲线。因此，确定恰当的收缩程度对分析至关重要。

图 1 最下面的两个面板展示了将 FASH 应用于这些数据后得到的效应函数后验分布。FASH 使用 L-高斯过程先验，但会汇总所有可用基因-变异对的信息，根据数据对这些先验进行调整。此外，还会根据数据与基线模型的匹配程度，分别为每个基因-变异对调整收缩程度；与基线模型接近的数据通常会被更强烈地收缩，而偏离基线模型较远的数据则不会。（第 4 节展示了其他基因-变异对的不同收缩模式。）这种方法有时称为“全局-局部”收缩

（Polson 和 Scott，2010年）。

在这个例子中，数据与效应随时间保持不变的基线模型并不一致。因此，相对于该基线模型，效应估计值的收缩程度较弱（图 1，左下）。基于该基线模型的假设检验用于检验效应是否随时间保持不变；显然，有充分证据拒绝这一假设。（下文将更正式地介绍 FASH 中的假设检验。）请注意，这项检验并未对效应随时间如何变化作出任何具体假设（这与许多已发表的动态 eQTL 研究中的分析不同，后者确实会作出此类假设）。

在图 1 右下角的面板中，我们看到，数据与效应随时间线性变化的模型更加一致。因此，效应估计值会强烈地向线性基线模型收缩。由于这些数据与线性基线模型高度吻合，因此几乎没有证据可以拒绝 eQTL 效应呈线性变化这一假设；据此，我们认为，eQTL 效应随时间大致线性增加，是对这些数据的一种合理描述。这两项检验针对不同的问题：第一项检验评估效应是否会随时间发生变化，第二项则评估是否有证据表明这些变化是非线性的。我们将在第 4 节给出其他假设检验的示例。

3 函数自适应收缩

3.1 符号说明

在数学表达式中，我们采用以下约定。对于整数 J，我们用 $[J] = \{1, \ldots, J\}$ 表示指标集。对于集合 A，$|A|$ 表示其基数。均值为 μ、方差为 σ^2 的正态分布记作 $N(\mu, \sigma^2)$，其密度为

$d\mathcal{N}(\cdot; \mu, \sigma^2)$。我们将“独立同分布”简称为“iid”，将“独立”简称为“ind”。我们用 $\mathbb{Z}^+$ 表示正整数集，用 $\mathbb{R}$ 表示实值向量集，用 $\mathbb{R}^p$ 表示长度为 p的实值向量集，用 $\mathbb{R}^{p\times q}$ 表示大小为 $p \times q$的实值矩阵集，并用 $C^p(\Omega)$ 表示满足p次连续可微条件的函数 $\Omega \to \mathbb{R}$ 的集合。

3.2 问题设定

设 $J \in \mathbb{Z}^+$ 为观测单位的数量。对于每个单位 $j \in [J]$，我们假设在连续条件变量 $t \in \mathbb{R}$ 的若干取值处都有效应估计值。我们将这些效应估计值记为 $\hat{\beta}_j(t_{j1}) \ldots, \hat{\beta}_j(t_{jR_j}) \in \mathbb{R}$，其中 R_j 是单位 j有条件变量 t效应估计值的设定数量。我们的主要目标是估计一个潜在的效应函数 β_j，它是从 $t \in \Omega$ 到 $\mathbb{R}$的映射。我们尤其关注单位数量 J较大的设定——例如数百或数千，甚至更多——这样我们就可以汇集信息，以改进对潜在效应函数的估计。对于每个 $j \in [J]$，在 t 的若干取值处都有观测值也很重要。

在动态 eQTL 研究中，J 是基因-变异对的数量，每个 $\beta_j(t_{jr})$ 是遗传变异在时间点 t_{jr}（或连续变量的设定 t_{jr}）对基因表达的估计效应。若每个基因仅关联一种遗传变异，则 J 就是基因的数量。在下面的动态 eQTL 案例研究（第 4 节）中，J超过 100 万，并且所有 $j \in [J]$ 的 $R_j = 16$。请注意，在一些动态 eQTL 研究中，表达仅在 2 或 3 个时间点测量，这可能不足以让我们的方法发挥作用。

3.3 泛函数据分析的经验贝叶斯

我们的方法假设，在给定底层效应函数的条件下，效应估计值相互独立且服从正态分布：

$$\hat{\beta}_j(t_{jr}) \mid \beta_j, s_{jr} \overset{\text{ind}}{\sim} N(\beta_j(t_{jr}), s_{jr}^2), \quad r \in [R_j], \tag{1}$$

其中，s_{jr} 表示效应估计值 $\beta_j(t_{jr})$ 的标准误。我们进一步假设，每个潜在效应函数 β_j 都是从 Ω 上函数的先验分布中独立同分布抽取的，记作 g_β：

$$\beta_j \overset{\text{iid}}{\sim} g_\beta. \tag{2}$$

我们有两个主要的推断目标：

1. **平滑：** 恢复并可视化潜在效应函数 β_j，或其泛函（例如导数），在 t 的观测值或未观测值处的取值。

2. **假设检验：** 识别偏离零假设 H_0:$\beta_j \in S_0$ 的单元 j，其中 S_0 表示某个预先指定的函数类。在某些应用中，我们可能还会对检验以下形式的假设感兴趣：$H_0 : \mathcal{F}(\beta_j) = 0$，其中某个泛函 $\mathcal{F}$：$C^p(\Omega) \to \mathbb{R}$。例如，我们可能想检验函数的最大值是否超过给定阈值 α；这可以通过设 $\mathcal{F}(\beta_j) = \max_{t \in \Omega} \beta_j(t) - \alpha$ 来实现。

第一个目标（“平滑”）将通过计算每个效应函数 β_j的后验分布来实现：

$$p(\beta_j \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, g_\beta) \propto p(\hat{\boldsymbol{\beta}}_j \mid \beta_j, \boldsymbol{s}_j)\, g_\beta(\beta_j), \tag{3}$$

使得 $\hat{\boldsymbol{\beta}}_j := (\hat{\beta}_j(t_{j1}) \ldots, \hat{\beta}_j(t_{jR_j}))^T$, $\boldsymbol{s}_j := (s_{j1}, \ldots, s_{jR_j})^T$表示单元 $j \in [J]$ 的可用数据。随后可通过其后验均值获得 β_j 的平滑点估计。

第二个目标（“假设检验”）通过计算局部错误发现率和局部错误符号率来实现（见下文）。

指定一个适用于所有数据集的单一先验 g_β 通常并不现实，因此，最好能有一种机制，根据数据自动学习或调整先验。我们采用经验贝叶斯方法来学习先验：给定一个预先定义的先验分布族$\mathcal{G}_\beta$ （将在下文定义），我们选择一个使所有数据的对数似然最大化的 $g_\beta \in \mathcal{G}_\beta$ ，

$$\hat{g}_\beta = \underset{g_\beta \in \mathcal{G}_\beta}{\operatorname{argmax}} \sum_{j=1}^{J} \log \int \prod_{r=1}^{R_j} p(\hat{\beta}_j(t_{jr}) \mid \beta_j, s_{jr}) \times g_\beta(\beta_j)\, d\beta_j, \tag{4}$$

随后，所有后续推断都使用这个估计先验$\hat{}g_\beta$。

尽管我们的主要目标是对未知效应函数进行推断，这需要对函数上的概率分布进行推理，但实际的底层统计计算很直接，可归结为有限维空间上的概率分布。令 $\boldsymbol{\beta}_j := (\beta_j(t_{j1}) \ldots, \beta_j(t_{jR_j}))$表示在效应估计值可用的 t 取值处的底层效应函数，并令 $\boldsymbol{\beta} := \{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J\}$，则上述建模假设意味着存在一个先验 $p(\boldsymbol{\beta}; g_\beta)$，使得

$$p(\boldsymbol{\beta}; g_\beta) = \prod_{j=1}^{J} p(\boldsymbol{\beta}_j; g_\beta), \tag{5}$$

其中， $p(\boldsymbol{\beta}_j; g_\beta)$ 表示由先验 g_β 在 $\boldsymbol{\beta}_j$ 上诱导的有限维分布。例如，对先验 (4) 进行最大似然估计所需的统计计算可简化为有限维积分：

$$\hat{g}_\beta = \underset{g_\beta \in \mathcal{G}_\beta}{\operatorname{argmax}} \sum_{j=1}^{J} \log \int p(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{\beta}_j, \boldsymbol{s}_j)\, p(\boldsymbol{\beta}_j; g_\beta)\, d\boldsymbol{\beta}_j, \tag{6}$$

其中

$$p(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{\beta}_j, \boldsymbol{s}_j) = \prod_{r=1}^{R_j} d\mathcal{N}(\hat{\beta}_j(t_{jr}); \beta_j(t_{jr}), s_{jr}^2). \tag{7}$$

3.4 函数自适应收缩先验族 s

现在，我们来介绍作用于效应函数的“函数自适应收缩”先验分布族，以实现上述目标。该分布族的构造始于指定一个“基线模型”， S_0。具体做法是指定一个 p 阶线性微分算子 $L = \sum_{i=0}^{p} c_i D^i$，其中 D^i 表示第 i 阶导数算子， $c_i \in \mathbb{R}$ 为指定的系数。随后，我们将基线模型定义为

$$S_0 = \text{Null}\{L\} = \{\beta \in C^p(\Omega) : L\beta = 0\}. \tag{8}$$

例如，若 $L = D^1$，则 $S_0 = \text{span}\{1\}$ 对应于常函数类；若 $L = D^2$，则 $S_0 = \text{span}\{1, t\}$ 对应于线性函数类。更一般地，对于 $L = D^p$，基线模型对应于 $p-1$阶多项式空间。

给定 L，我们将 L-高斯过程（“L-GP”）定义为满足以下条件的高斯过程

$$L\beta = \sigma\xi, \tag{9}$$

其中， ξ 是标准高斯白噪声，而 $\sigma > 0$ 是一个参数，用于控制β 偏离基线模型 S_0的程度。为简洁起见，我们将其写作

$$\beta \sim L\text{-GP}(\beta; \sigma).$$

当 $L = D^p$时， L-GP对应于 p阶积分维纳过程（IWP$_p$）（Shepp，1966年）。IWP先验与平滑样条密切相关，因为其后验均值与样条估计量一致（Wahba，1978）。下文中，我们重点讨论IWP$_p$ 形式的L-GP，同时指出该方法也适用于其他类型的L-GP；更多背景和示例参见 Lindgren 和 Rue（2008）； Yue 等人（2014）； Zhang 等人（2024, 2025）。

对 β 使用 L-高斯过程($\beta; \sigma$)先验时， L 的选择决定收缩所趋向的基线模型，而 σ 的选择决定收缩程度。当 σ 逐渐

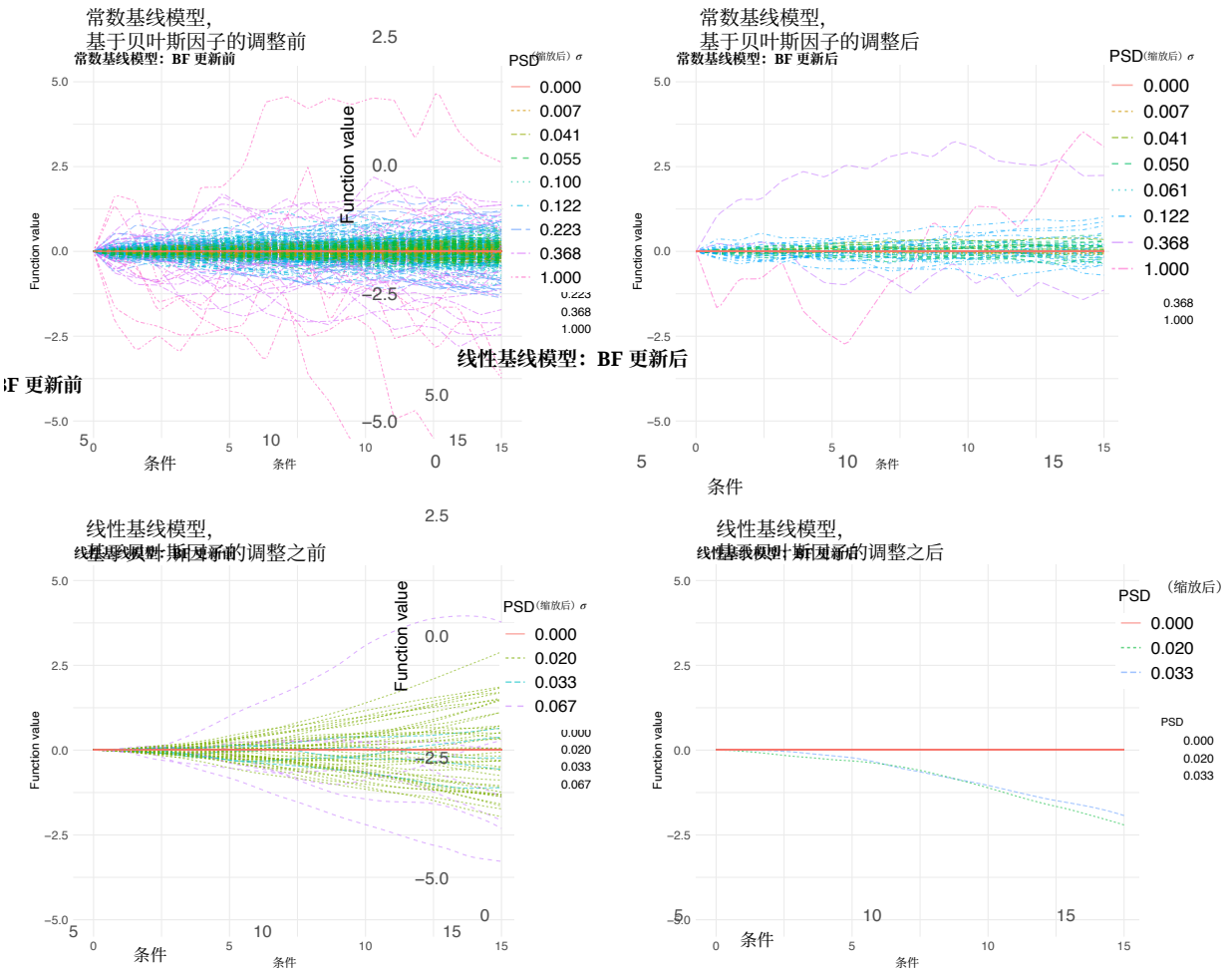


图 2：从第 4 节中基于动态 eQTL 数据估计的拟合 FASH 先验中随机抽取的样本路径，分别对应 $L = D^1$ （上）和 $L = D^2$ （下），以及基于贝叶斯因子的调整前（左）和调整后（右）的情况。为便于可视化，我们将样本路径的初始条件设为零；也就是说，$\beta(0) = 0$ 位于上排，$\beta(0) = \beta'(0) = 0$ 位于下排。每幅图均展示了随机抽取的 5,000 条路径。

趋近于零时，L-高斯过程会受到越来越多的约束，使其保持接近基线模型。图 2 展示了 σ 的不同设定如何产生不同的效应函数先验分布。

对于一个能够根据数据适当调整收缩程度的灵活先验分布族，我们采用混合 L-高斯过程先验，

$$\mathcal{G}_\beta = \left\{ g_\beta : g_\beta = \sum_{k=0}^{K} \pi_k \, L\text{-GP}(\beta; \sigma_k) \right\}, \tag{10}$$

其中 $\{\sigma_k\}_{k=0}^K$ 表示由小到大排序的固定网格值，小值（$\sigma_0 = 0$）对应于偏离 S_0为零，大值为（σ_K）； π_k 表示混合权重（$\pi_k \geq 0$， $\sum_{k=0}^K \pi_k$ =1）。我们将（10）中的先验族称为“函数自适应收缩先验”族，因为其构造方式仿照了 Stephens（2017）和 Urbut 等人（2019）提出的“自适应收缩先验”。

由于形式为 (10) 的先验完全由混合权重 $\boldsymbol{\pi} := (\pi_k)_{k=0}^K$ 指定，因此学习先验 $g_\beta \in \mathcal{G}_\beta$ 可简化为学习先验权重：

$$\begin{aligned} \hat{\boldsymbol{\pi}} &= \operatorname{argmax}_{\boldsymbol{\pi}} l(\boldsymbol{\pi}), \\ l(\boldsymbol{\pi}) &:= \sum_{j=1}^{J} \log \left\{ \sum_{k=0}^{K} \pi_k \ell_{jk} \right\}, \end{aligned} \tag{11}$$

其中， ℓ_{jk} 表示单元 j 在先验的第 k 个分量下的边际似然；也就是说， ℓ_{jk} 是边际似然 $p(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j) = \int p(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{\beta}_j, \boldsymbol{s}_j)\, p_j(\boldsymbol{\beta}_j;\, g_\beta)\, d\boldsymbol{\beta}_j$，其先验形式为 (10)，其中 $\pi_k = 1$，其余混合权重均为零。

3.5 FASH 中的后验推断

除了计算效应函数的后验均值、方差及其他后验矩外，我们还使用后验分布 (12) 对效应函数进行假设检验。所有这些计算都涉及求取关于效应函数后验分布 (3) 的期望。由于似然 (1) 与先验 (10) 具有共轭性，后验分布也为混合分布：

$$p(\beta_j \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, \hat{\boldsymbol{\pi}}) = \sum_{k=0}^{K} \tilde{\pi}_{jk}\, p_k(\beta_j \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j), \tag{12}$$

其中， $p_k(\beta_j \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j)$ 表示在第 k 个 L-高斯过程组分下单元 j 的后验分布，而 $\tilde{\pi}_{jk}$ 表示后验混合权重，

$$\tilde{\pi}_{jk} := \frac{\hat{\pi}_k\, \ell_{jk}}{\sum_{k'=0}^{K} \hat{\pi}_{k'} \ell_{jk'}}. \tag{13}$$

此外， L-高斯过程的马尔可夫结构有助于高效计算这些后验分布。我们将在补充材料的附录 A 中对此进行详细说明。

3.5.1 假设检验

检验 $H_0 : \beta_j \in S_0$ 时，一种常见方法是在 J 个单位中控制错误发现率（FDR），既可以采用经典的频率学派表述（Benjamini和Hochberg，1995），也可以采用与之密切相关的贝叶斯表述（Storey，2002；Efron，2008）。对于被标记为显著的子集 $\Gamma \subseteq [J]$，其估计错误发现率根据局部错误发现率（lfdr）（Efron，2008）计算如下

$$\widehat{\mathrm{FDR}}(\Gamma) = \frac{1}{|\Gamma|} \sum_{j \in \Gamma} \mathrm{lfdr}(j), \tag{14}$$

其中

$$\mathrm{lfdr}(j) := p(\beta_j \in S_0 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, \hat{\boldsymbol{\pi}}). \tag{15}$$

由于混合先验（10）的 $k = 0$ 分量对应于 S_0，因此在 FASH 中，lfdr 可简化为

$$\mathrm{lfdr}(j) = \tilde{\pi}_{j0}. \tag{16}$$

对于形式为 $H_0 : \mathcal{F}(\beta_j) = 0$ 的假设检验，其中 $\mathcal{F} : C^p(\Omega) \to \mathbb{R}$ 是一个泛函，可以考虑使用局部错误符号率（lfsr）和错误符号率（FSR）；与 lfdr 和 FDR 相比，它们通常更稳健（Stephens，2017）。对于双侧备择假设$H_1 : \mathcal{F}(\beta_j) \neq 0$，lfsr 为

$$\mathrm{lfsr}(j) = \min\left\{p(\mathcal{F}(\beta_j) \geq 0 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, \hat{\boldsymbol{\pi}}),\, p(\mathcal{F}(\beta_j) \leq 0 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, \hat{\boldsymbol{\pi}})\right\}. \tag{17}$$

对于单侧备择假设，例如 $H_1 : \mathcal{F}(\beta_j) > 0$，我们将局部错误符号率定义为

$$\mathrm{lfsr}(j) = p(\mathcal{F}(\beta_j) \leq 0 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j, \hat{\boldsymbol{\pi}}). \tag{18}$$

累积错误符号率可以根据局部错误符号率，以类似于计算错误发现率的方式计算（Stephens，2017）。

3.6 基于贝叶斯因子的调整 $\hat{}\pi_0$

局部错误发现率和错误发现率可能对零假设比例的估计值 π_0 相当敏感。尽管局部错误符号率和错误符号率通常比局部错误发现率和错误发现率更稳健，但仍可能受到估计值 $\hat{}\pi_0$ 的影响。特别是，如果 $\hat{}\pi_0$ 低于真实的 π_0，推断可能变得反保守，导致错误发现率或错误符号率膨胀。因此，更可取的是采用保守估计，即 $\hat{}\pi_0 \geq \pi_0$。

FASH 中低估 $\hat{}\pi_0$ 的一个主要原因是先验族设定不当；如果混合先验族 (10) 不够灵活，无法逼近真实的边缘分布，那么 π_0 的最大似然估计就可能不准确。在较简单的单变量设定中（Stephens，2017），先验误设或许只是一个轻微的问题；但在 FASH 中，如同其他多元设定（Urbut 等，2019；Liu 等，2024），由于难以在高维下充分建模数据，这一问题更加突出。

为避免这一问题，我们介绍一种简单而有效的基于贝叶斯因子的调整方法，用于调整 $\hat{}\pi_0$。这种调整并非 FASH 专属，也可能适用于先验可能设定不当的其他情形。该调整仅利用这样一个事实：在零假设下，支持备择假设的贝叶斯因子的期望为 1（也就是说，贝叶斯因子是e变量；参见 Vovk 和 Wang，2021）。贝叶斯因子的这一性质在下面的引理中有更正式的表述。

引理 1： 设单元 j 的贝叶斯因子为

$$\mathrm{BF}_j := \frac{p_1(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)}{p_0(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)},$$

其中，p_1 和 p_0 分别表示备择假设（H_1）和零假设（H_0）下的边缘似然。在零假设 H_0 下，无论如何指定备择假设 H_1，我们都有 $\mathbb{E}_0(\mathrm{BF}_j) = 1$。

算法1 对 π_0进行基于贝叶斯因子的调整。

要求： 一个 $J \times (K+1)$ 矩阵 $\mathbf{L}$，其中元素（j,k）为 ℓ_{jk}，即单元 j 在混合先验的第 k 个分量下的边际似然；估计先验权重 $\hat{\boldsymbol{\pi}} = (\hat{\pi}_0, \ldots, \hat{\pi}_K) \in \mathbb{R}^{K+1}$；一组候选截断值 $\mathcal{C} \subset \mathbb{R}^+$；一个"缓冲区"调节参数 $\epsilon > 0$，通常接近于零。

对备择权重进行归一化，$\hat{\pi}_k^* = \hat{\pi}_k / \sum_{k'=1}^{K} \hat{\pi}_k'$, $k = 1, \ldots, K$。

计算"折叠"似然，$\ell_{j0}^c = \ell_{j0}$ 和 $\ell_{j1}^c = \sum_{k=1}^{K} \ell_{jk} \hat{\pi}_k^*$, $j = 1, \ldots, J$。

计算 $j = 1, \ldots, J$ 对应的贝叶斯因子 $\mathrm{BF}_j = \ell_{j1}^c / \ell_{j0}^c$。

对于 $c \in \mathcal{C}$ **执行**

$J_0 = \sum_{j=1}^{J} \mathbb{I}\{\mathrm{BF}_j < c\}$

$\hat{\pi}_0(c) = J_0 / J.$

$\mu(c) = \sum_{j=1}^{J} \mathrm{BF}_j \times \mathbb{I}\{\mathrm{BF}_j < c\} / J_0.$

循环结束

$c^* = \inf\{c \in \mathcal{C} : \mu(c) \geq 1 + \epsilon\}.$

调整零假设先验权重，$\hat{\pi}_0 = \hat{\pi}_0(c^*)$。

调整备择先验权重，$\hat{\pi}_k = \hat{\pi}_k^*(1 - \hat{\pi}_0(c^*))$，$k = 1, \ldots, K$。

返回调整后的权重，$\hat{\boldsymbol{\pi}} = (\hat{\pi}_0, \ldots, \hat{\pi}_K)$。

证明。

$$\mathbb{E}_0(\mathrm{BF}_j) = \int \frac{p_1(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)}{p_0(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)} \times p_0(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)\, d\hat{\boldsymbol{\beta}}_j = \int p_1(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j)\, d\hat{\boldsymbol{\beta}}_j = 1.$$

□

令 J_0 和 J_1 分别表示零假设单元和备择假设单元的数量，其中 $J = J_0 + J_1$且 $\pi_0 = J_0/J$。根据大数定律，当 J_0 较大时，我们有

$$\frac{1}{J_0} \sum_{j \in \mathcal{H}_0} \mathrm{BF}_j \;\approx\; 1,$$

其中$\mathcal{H}_0 \subseteq \{1, \ldots, J\}$ 表示零假设单元。这一观察促使我们提出一种简单的保守程序来估计 J_0，进而估计 π_0：寻找尽可能大的单位集合，该集合

平均而言与全部为零假设相一致（平均贝叶斯因子 ≤ 1）。算法1实现了这一程序，我们称之为 π_0的基于贝叶斯因子的调整。下面的定理1确立了调整后的$\hat{}\pi_0$的保守性。

定理1 （基于贝叶斯因子的调整给出保守估计）。 假设 $\hat{}\pi_0$ 是通过算法1得到的，且 J_0 零效应独立同分布于先验所指定的零分布。那么，对于任意备择分布以及任意 $\epsilon > 0$,

$$\hat{\pi}_0 \geq \pi_0 \text{ almost surely as } J_0 \to \infty.$$

该定理的证明见补充材料的附录 B。

重要的是，定理*1*仅要求零分布设定正确 （即 FASH 先验的$k = 0$ 混合成分， g_β），而备择分布（ g_β中从混合成分 1 到 K ）可以设定不正确。例如，这一结果不依赖于收缩值 σ_1，.....， σ_K，这意味着即使使用粗网格的取值来减少计算量，基于贝叶斯因子的调整也应当有效。

零假设必须被正确设定，这一要求类似于校准经典 p 值时通常所需的条件，而且远没有正确设定完整分布的要求那么苛刻。尽管如此，在这种设定下，这一要求也并非完全无害，因为零分布并非仅仅是集中在零点的点质量，而是在维数为 $p \geq 1$ 的函数类 S_0 上的弥散分布。（弥散先验可能会给贝叶斯因子的计算带来问题，因为它们会使边际似然的尺度任意变化；但在这里，所有混合成分共享同一个弥散成分，因此计算比值时这些任意尺度会相互抵消，所得贝叶斯因子仍然有效；参见 Cheng和Speckman 2016; Servin和Stephens 2007。）如果真实的零函数在 S_0 上并非弥散分布，那么严格来说，这会导致先验误设，定理1的条件也就不再成立。尽管如此，在我们的

实验中，即使先验误设，基于贝叶斯因子的调整也始终给出了对 $\hat{}\pi_0$ 的保守估计（见补充材料的附录 C）。

由 (13) 和 (15) 可知，在 FASH先验中高估 $\hat{}\pi_0$ 会使后验概率质量向基线模型 S_0 偏移，从而在针对 H_0 时做出更保守的决策。图 2 对此进行了说明。在补充材料的附录 B 中，我们通过考察比较零假设与备择假设的后验优势比，进一步详述了这一点。补充材料的附录 C 还通过实证研究考察了相关量（如局部错误发现率和局部错误符号率）的保守性。

4 iPSC 心肌细胞分化研究中的动态表达数量性状位点分析

我们在模拟数据集中通过多种性能指标评估了 FASH。这些模拟实验在补充材料的附录 C 中有详细说明。值得注意的是，我们的模拟证实，基于贝叶斯因子的调整会对π_0产生保守估计，因此即使先验设定错误，也会对局部错误发现率和错误发现率产生保守估计。令人欣慰的是，这些保守估计并未显著降低检验效能。现在，我们将重点转向一个真实数据应用，即对 iPSC 心肌细胞分化研究（Strober等，2019）中的动态表达数量性状位点进行重新分析。该研究在 16 天期间每日测量基因表达，研究对象为来自 19 个正在分化为心肌细胞的约鲁巴 HapMap 细胞系的诱导多能干细胞（iPSC）。分析目标是识别动态表达数量性状位点——也就是说，识别其对基因表达的效应随时间变化的遗传位点——并刻画基因表达的遗传调控如何随着分化过程而变化。我们表明，FASH 能够以一种有原则的方式实现这些目标，并将 FASH 的结果与

基于参数交互模型的原始分析进行比较。

按照 Strober等人（2019）所述的数据预处理和质量控制流程处理后，我们获得了 $J = 1$,009,173 个基因-变异对（6,362 个基因）的数据。本分析考虑的所有遗传变异，都是位于基因转录起始位点（TSS）50 kb 范围内的单核苷酸多态性。对于每个基因-变异对 $j \in [J]$，我们通过在每个时间点分别拟合线性回归模型，在 16 个时间点获得了 eQTL效应估计 $\beta_j(t_r)r \in [16]$。随后，我们对 eQTL效应估计的标准误 s_{jr} 进行了调整，以解决样本量较小可能导致 I类错误膨胀的问题（详情见补充材料的附录 D）。

具体而言，我们考虑了两个推断目标：

1. 识别动态表达数量性状位点。我们在 FASH 中使用带有 $L = D^1$ 的 L-高斯过程先验混合分布来实现这一目标。我们将本分析中拟合的模型称为“FASH-IWP$_1$”模型。

2. 识别非线性动态 *eQTL*。 这是通过在 FASH 中使用带有 $L = D^2$ 的 L-高斯过程先验混合分布实现的。我们将本分析中拟合的模型称为“FASH-IWP$_2$”模型。

所有结果均使用我们 R 包的 0.1.42 版本生成， `fashr`。实现我们分析的 R 代码可在 https://github.com/stephenslab/fashr-paper 获取。在所有分析中，FASH 先验均定义在对数尺度上等间距的网格上；详情见附录 D。在通过最大似然估计混合权重 $\boldsymbol{\pi}$ 后，我们应用了基于贝叶斯因子的调整，然后使用调整后的权重 $\hat{\boldsymbol{\pi}}$ 进行推断[^1]，基于后验分布 $p(\beta_j \mid \boldsymbol{\beta}_j, \ \boldsymbol{s}_j, \ \hat{\boldsymbol{\pi}})$。我们使用 0.05 的错误发现率阈值来识别“显著”的动态表达数量性状位点。[1] 从头到尾完成每项 FASH 分析

[1]我们使用错误发现率，是为了使 FASH 分析与原始分析更具可比性，但局部错误发现率通常更可取；关于在假设检验中使用错误发现率和错误符号率，而非局部错误符号率和局部错误符号率的进一步讨论，请参见第 5 节

使用 `fashr`耗时约 11 小时。计算在搭载 Intel Xeon Gold 6248R（“Cascade Lake”）处理器的 Linux 机器（Scientific Linux 7.4）上进行，使用了 16 个并行线程。

4.1 动态表达数量性状位点的发现

详细考察 FASH 使用 $L = D^1$ 识别出的部分动态表达数量性状位点，可以展示 FASH 能捕捉到的各种动态变化。图3顶行（A、B）中的示例，其效应似乎随时间逐渐增强。在这些示例中，Strober等人（2019）使用的线性交互作用（$G_c \times t$，其中 G_c 是单核苷酸多态性基因型）和二次交互作用（$G_c \times t^2$）模型也很好地拟合了这些示例中的 eQTL 效应；事实上，这些基因-变异对也被 Strober 等人（2019）识别为动态表达数量性状位点。相比之下，图3中排在中期和底部的示例（C–F）被 FASH 识别为动态表达数量性状位点，但未在原始分析中识别出来。事实上，在所有这些示例中，线性和二次模型似乎都无法很好地拟合这些 eQTL 的非线性动态变化；例如，C 和 D 中的效应在第 5 天左右突然增强，而 E 和 F 中的效应则在第 10 天左右突然增强。FASH 对动态表达数量性状位点 C 和 D 的识别非常明确（局部错误发现率为 0.01 或更低）；E 和 F 则是“临界”动态表达数量性状位点（局部错误发现率 > 0.1），其时间动态较为微妙。补充材料中的图S5、S6和S10还提供了其他说明性示例，其中包括 FASH 未识别为动态表达数量性状位点的基因-变异对。

因此，鉴于 FASH能够捕捉多种时间模式，FASH识别出的动态表达数量性状位点远多于原始分析（图4）：当错误发现率阈值为 $\alpha = 0.05$时，FASH-IWP$_1$ 在 9,205 个基因-变异对（涉及 1,177 个

在 FASH 中进行测试。

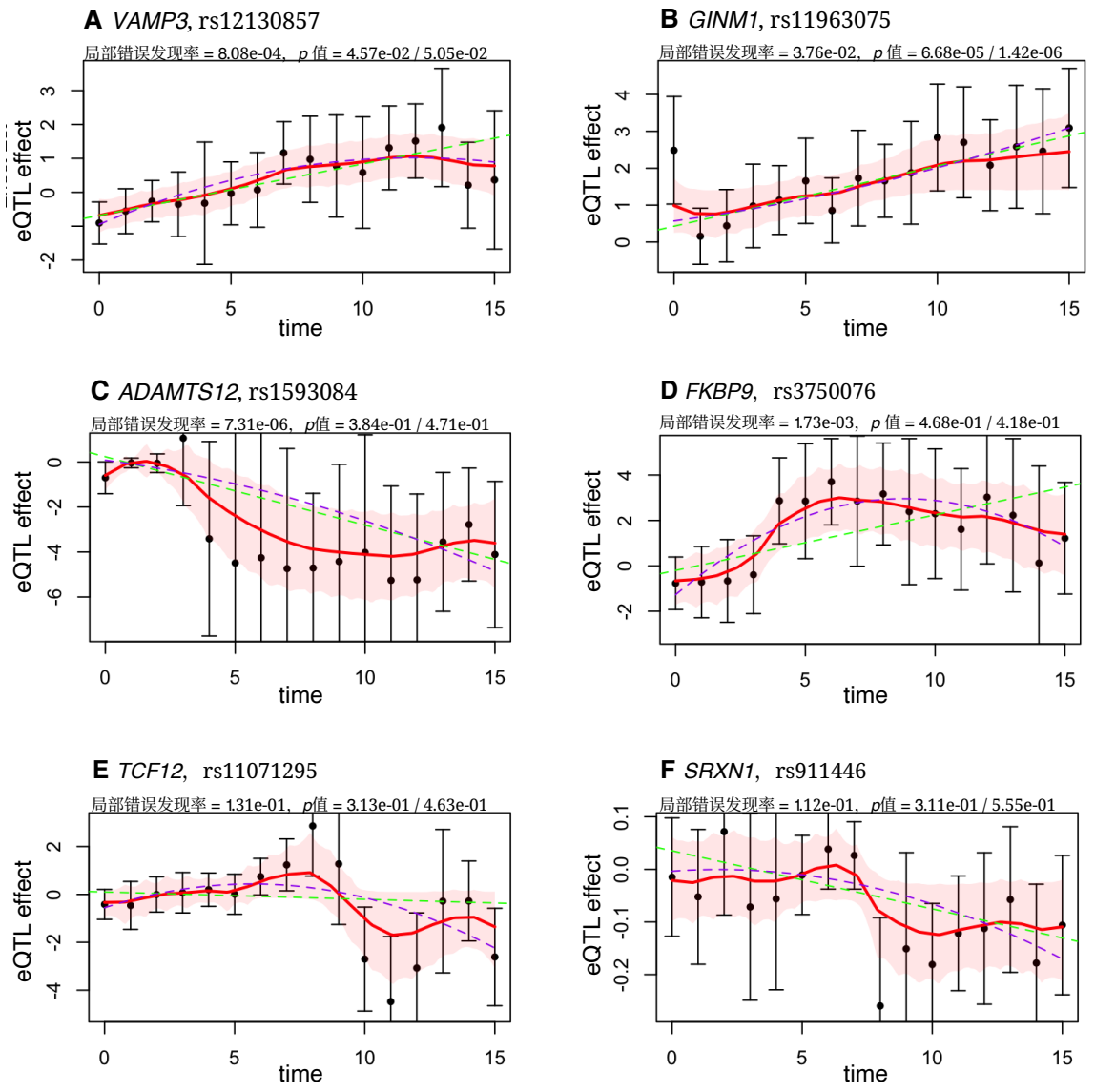
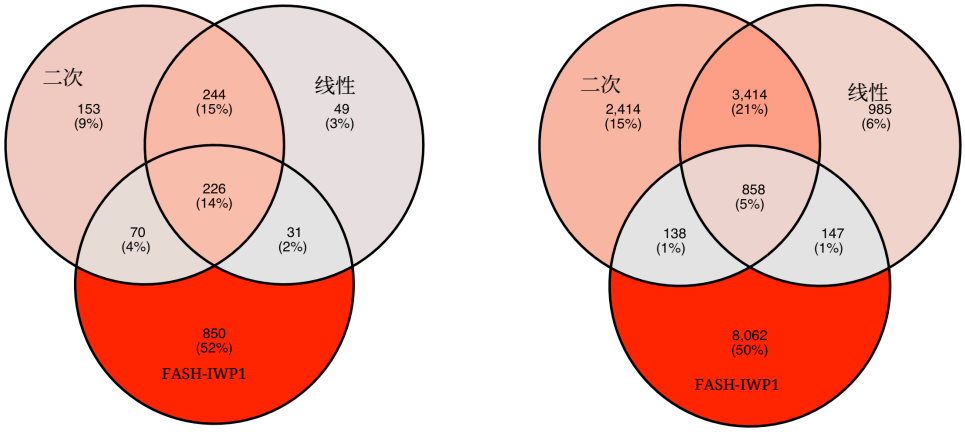


图3：由FASH（$L = D^1$）识别的动态表达数量性状位点示例，以及与参数建模方法（Strober等，2019）的比较。在每幅图中，β_j 的后验均值以红色实线表示，95% 可信区间由˜阴影区域描绘。观测到的效应大小估计值 β_j 以黑点表示，垂直误差线表示 ± 2（校正后的）标准误 $\tilde{s}_{jr}$。局部错误发现率是贝叶斯因子校正后FASH-IWP$_1$ 模型的局部错误发现率；p-值来自Strober等人（2019）使用线性/二次交互模型的分析。线性（$G_c \times t$）和二次（$G_c \times t^2$）参数交互模型的逆方差加权最小二乘估计以虚线表示（绿色 = 线性，紫色 = 二次）。

基因）中识别出动态表达数量性状位点，约占所有受检基因-变异对的 1%（占所有基因的 19%）；在经验错误发现率阈值为 0.05 时，Strober等人（2019）使用线性交互模型，在 5,404 个基因-变异对（涉及 550 个基因）中识别出动态表达数量性状位点；使用



(a)：发现的独特基因数量

(b)：发现的配对数量

图4：使用 FASH 与线性/二次参数交互模型发现动态表达数量性状位点：（a）基因；（b）基因-变异对。

二次交互模型。尽管发现的动态表达数量性状位点有所增加，FASH 推断仍然是“保守的”，因为这些推断是通过对 FASH 先验进行基于贝叶斯因子的调整后得出的。（图 2 展示了这种调整对 FASH 先验的影响。）

需要指出的是，由于两项分析之间除了FASH具有更高的灵活性之外还存在其他若干差异，这些差异也很可能促成了发现结果的差异，并且可能部分解释了为何线性交互模型和二次交互模型识别出的许多动态表达数量性状位点未能在我们的分析中重现。例如，考虑到较小的样本量可能导致I类错误膨胀，我们按照补充材料中所述的简单程序对标准误进行了调整，而原始分析未对此问题加以处理。因此，原始分析未处理这一问题，可能导致了更多的假阳性。另一个重要考虑因素是，两项分析在估计错误发现率方面采用了截然不同的方法：FASH采用经验贝叶斯方法，在模型拟合过程中估计（先验）零假设比例，随后利用该比例估计局部错误发现率和错误发现率；Strober等人（2019）的分析则采用基于置换的方法

来估计零分布p-值，然后应用Gamazon 等（2013）的方法估计错误发现率。有关两项分析更详细的比较，请参见补充材料中的图S8和S9，其中比较了FASH的局部错误发现率值（及其相应的假发现率估计值）与原始分析中的p-值（及其相应的eFDR估计值）。

4.2 动态表达数量性状位点的表征

现在，我们从发现动态表达数量性状位点转向表征这些动态表达数量性状位点的动态变化；特别是，我们希望表征基因表达上的遗传效应如何在细胞分化过程中演变。Strober等人（2019年）也试图表征动态表达数量性状位点，但由于现有方法的局限，其统计分析变得复杂。在这里，我们展示了在 FASH建模框架内，这一过程更容易实现；而且由于同时提供了经过适当校准的后验统计量（如局部假发现率和局部假符号率），结果也更加可靠。

通过使用线性基线模型而非常数基线模型重新分析数据——也就是说，采用带有 $L = D^2$的 FASH先验——FASH识别出了对基因表达具有非线性动态效应的基因-变异对。在错误发现率阈值为 $\alpha = 0.05$时，这个“FASH-IWP$_2$”模型发现了少量非线性动态 eQTL：9个基因中的44个基因-变异对（表表 1）。（未进行基于贝叶斯因子的调整时，FASH-IWP$_2$ 在 $\alpha = 0.05$时识别出37个基因中的159个基因-变异对，数量仍远少于具有动态表达数量性状位点的基因总数；见图4。）

图5展示了几个这类非线性动态 eQTL 的示例。（另请参见图 S7，其中展示了未被归类为“非线性”的动态 eQTL 示例。）非线性效应的多样性相当显著。显然，尽管二次交互模型

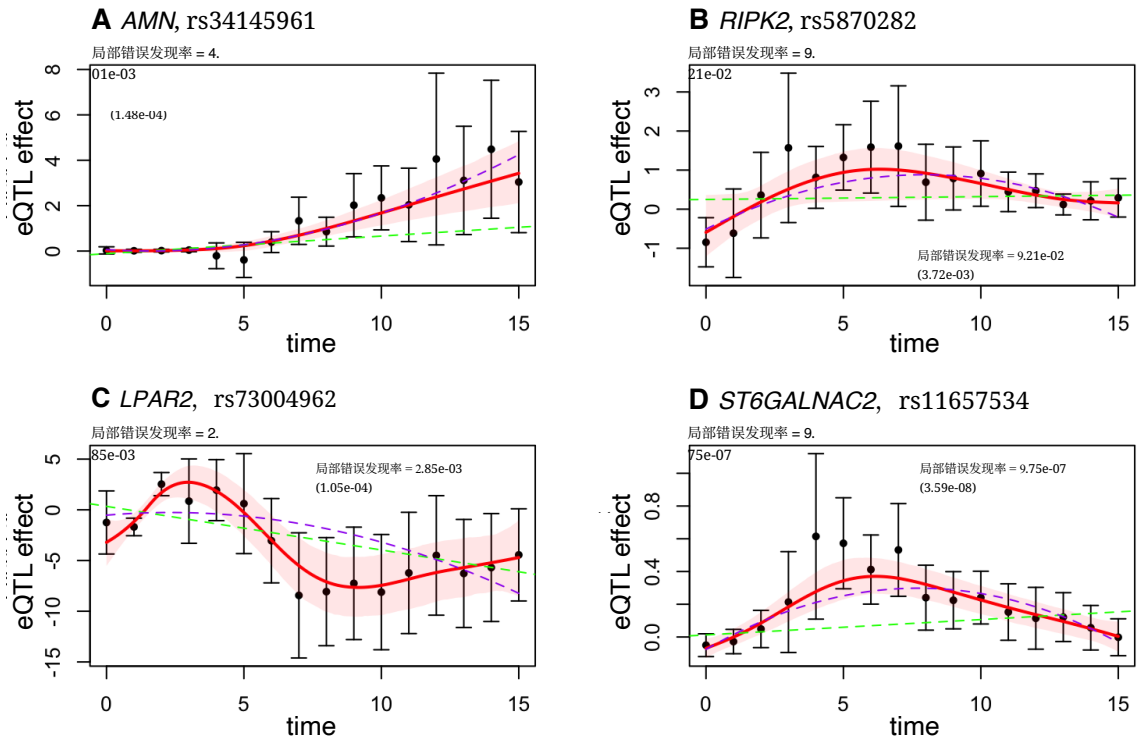


图5：FASH 识别出的非线性动态 eQTL 示例，其中包含 $L = D^2$。在每个图中，β_j 的后验均值以红色实线表示，95% 可信区间则由阴影区域表示。观测到的效应大小估计值 β_j 以黑色圆点表示，垂直误差线表示 ± 2（校正后的）标准误 $\tilde{s}_{jr}$。lfdr 是 FASH-IWP$_2$ 模型在贝叶斯因子校正后的 lfdr。线性（$G_c \times t$）和二次（$G_c \times t^2$）参数交互模型的逆方差加权最小二乘估计分别以虚线表示（绿色 = 线性，紫色 = 二次）。

旨在识别 Strober等人（2019）中的非线性动态效应，但它不够灵活，无法涵盖各种非线性效应。

FASH 能够检验形式为$\mathcal{F}(\beta) = 0$或$\mathcal{F}(\beta) > 0$（第 3.5.1 节）的假设，且$\mathcal{F}$ 可以是任意泛函，这为系统地刻画动态效应提供了新方法。例如，Strober等人（2019）希望找出那些对表达的效应会改变方向的动态表达数量性状位点。为将其表述为假设检验，我们定义一个泛函$\mathcal{F}(\beta)$，用于衡量基因型效应是否呈现符号改变的“开关”现象，且其幅度是否超过阈值 $c > 0$:

$$\mathcal{F}(\beta) = \min\left\{\max_t \beta^+(t),\ \max_t \beta^-(t)\right\} - c, \tag{19}$$

类别	基因变异	基因	泛函，$\mathcal{F}(\beta)$
动态	44	9	（不适用）
早期	124	8	$\max_{t\le 3}\lvert\beta(t)\rvert - \max_{t>3}\lvert\beta(t)\rvert$
中期	24	5	$\max_{4\le t\le 11}\lvert\beta(t)\rvert - \max_{t<4 \text{ 或 } t>11}\lvert\beta(t)\rvert$
late	20	12	$\max_{t\ge 12}\lvert\beta(t)\rvert - \max_{t<12}\lvert\beta(t)\rvert$
开关	984	250	最小$\{$最大$_{0\le t\le 15}\beta^{+}(t)$, 最大$_{0\le t\le 15}\beta^{-}(t)\} - c$, $c = 0.25$

表 1：根据时间效应模式对动态表达数量性状位点进行分类：在错误发现率（首行）或错误符号率（其余各行）为 0.05 时识别出的基因-变异对数量和独特基因数量。对于开关类别，设定 $c = 0.25$ 可确保在基因型剂量范围（0–2）内，最大的效应量差异至少为 1，因为 $2 \times c \times 2 = 1$。图 6 展示了各类别的示例。

Hallmark基因集	基因 全部/开关	p-值 全部/开关	q-值 全部/开关
缺氧反应中上调的基因	25/11	0.017/0.00067	0.436/0.025
KRAS 激活所上调的基因	11/7	0.24/0.0022	0.840/0.035

表 2：动态表达数量性状位点基因与开关动态 eQTL 基因的基因集富集分析。“基因”列给出了属于 Hallmark基因集的动态基因 eQTL 或开关动态 eQTL 基因数量（Liberzon 等人，2015a,b；Subramanian 等人，2005）。GSEAp-值和q-值使用 clusterProfiler R 包中的“enricher”函数计算（Wu 等人，2021；Yu 等人，2012）。更详细的 GSEA 结果见补充材料的表 3。

其中，β^{+} 和 β^{-} 分别表示 β的正部和负部，因此 $\beta = \beta^{+} - \beta^{-}$。当$\mathcal{F}(\beta) > 0$时，存在时间点 t_{+} 和 t_{-}，使得 $\beta(t_{+}) > c$ 且$\beta(t_{-}) < -c$，这意味着基因型每改变一个等位基因，这两个时间点之间的效应差至少为 $2c$。对于 $c = 0.25$，这对应于两种纯合基因型之间的最小差异为$4c = 1$。

除此之外　　e开关类别，我们使用假设检验框架将

动态表达数量性状位点分为另外三类：

- 早期： 最强效应发生在细胞分化的前 3 天中的某个时间。

- 中期： 最强效应发生在第 4 天至第 11 天之间的某个时间。

- 晚期： 最强效应发生在最后 4 天中的某个时间。

这些类别分别对应于如下形式的不等式$\mathcal{F}(\beta) > 0$（表 1）。（注意，切换类别并不排斥其他类别；例如，动态 eQTL 应同时归类为早期 和开关。）在 FSR 阈值为 0.05 时（对先验应用基于贝叶斯因子的调整后），归入这些类别的基因变异和基因数量见表 1。图 6 展示了每个类别中的动态 eQTL 示例。只有极少数动态 eQTL 被归类为早期、中期或晚期，但许多动态 eQTL 的效应被识别为方向发生切换。为更好地理解开关动态 eQTL 的生物学意义，我们使用 clusterProfiler（Wu 等人，2021；Yu 等人，2012）对 Hallmark 基因集（Liberzon 等人，2015a,b；Subramanian 等人，2005）进行了基因集富集分析（GSEA），并比较了具有开关动态 eQTL 的 250 个基因与具有任意类型动态 eQTL 的 1,177 个基因的 GSEA 结果。有趣的是，开关动态 eQTL 中排名前两位的 Hallmark 基因集（按p-值排序）分别是低氧水平（即缺氧）下上调的基因和受 K-Ras 上调的基因；它们也是所有动态 eQTL 中排名靠前的基因集，但仅考虑开关动态 eQTL 时，富集程度要强得多（表 2）。众所周知，缺氧和 K-Ras 都会对干细胞的增殖和分化产生显著影响，因此，开关动态 eQTL 在这些通路中的富集程度更高，可能具有重要意义。更详细的 GSEA 结果见补充材料中的表 3。

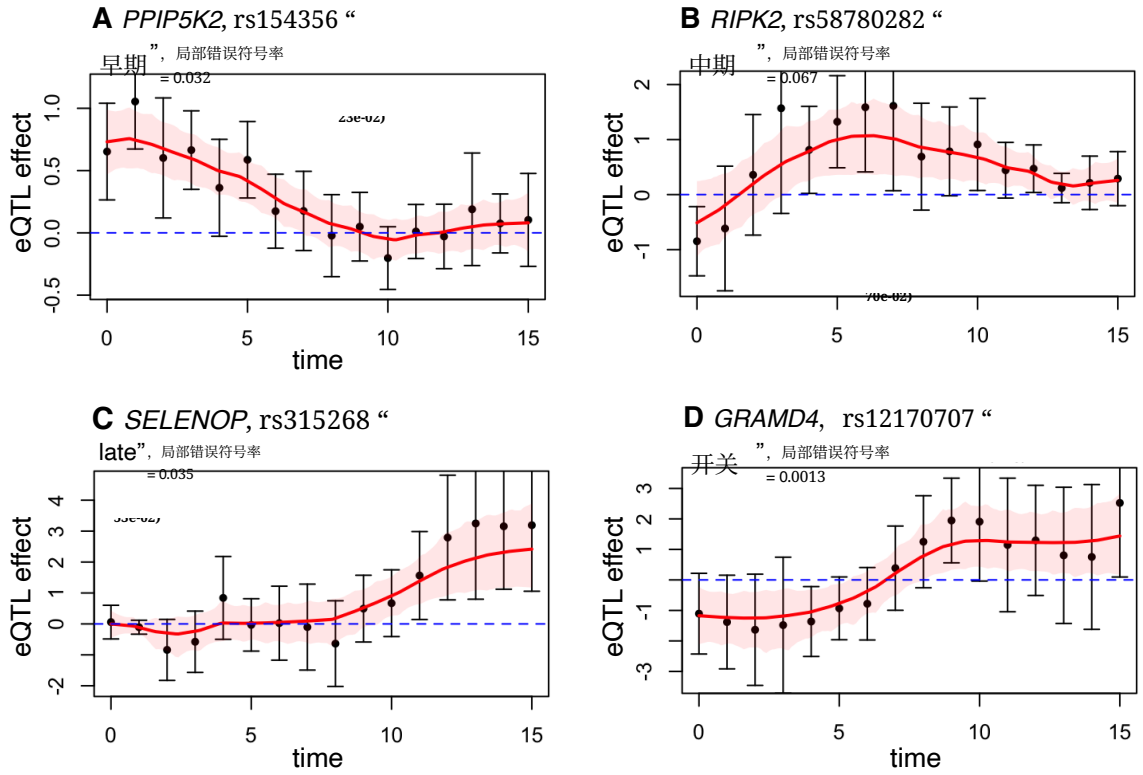


图 6：FASH 将动态表达数量性状位点归类为 (A) 早期、(B) 中期、(C) 晚期 和 (D)开关的示例。（这些类别的定义见表 1。） β_j 的后验均值以红色实线表示，95% 可信区间以 ˆ 阴影区域表示。观测到的效应大小估计值 β_j 以黑点表示，垂直误差线表示 ± 2 （调整后的）标准误 $\tilde{s}_{jr}$。“lfsr”是使用 BF调整先验得到的局部错误符号率。

5 讨论

本文将经验贝叶斯（EB）的思想扩展到函数后验分布的推断。 由此形成了一个强大而灵活的建模框架 FASH（“函数自适应收缩”），可用于检验关于函数的各种假设。FASH 通过在所有观测单位之间借用信息，自动调整先验，因此随着可用数据增多，后验推断应会更加准确。在案例研究中，我们使用 FASH 重新分析动态表达数量性状位点；FASH 尤其吸引人的一点是，它能够将动态 eQTL 归为不同类型（表 1）。我们还提出了一种简单的先验调整方法，以解决对误校准的担忧

在先验误设时，错误发现率和其他后验统计量的估计。这种调整会使先验估计更加保守——也就是说，赋予零假设比例 π_0 更大的权重——因此推断也更加保守。此提议并非 FASH 所独有，因此也可能用于其他经验贝叶斯方法，尤其是在高维多变量数据设定中（Urbut 等，2019；Liu 等，2024），因为这类设定中很可能出现先验误设。

尽管 FASH 在函数上定义后验分布，实际计算涉及有限维空间上的概率分布，因此计算是可行的。为了使这些可行的后验计算能够扩展到大型数据集，我们选择了“L-GP”先验族，因为其后验计算的复杂度随 J（观测单位的数量）和 R_j（每个单位的观测值数量）线性增长；后验计算复杂度的更多细节见补充材料的附录 A（标准 GP 的后验计算复杂度随 R_j 呈立方增长）。除了计算效率之外，“L-GP”先验族还为假设检验提供了灵活性。原则上，FASH 可以扩展到其他高斯过程先验族，以满足其他应用的需求。例如，基于空间高斯随机场的先验族（Lindgren 等，2011）可能适用于涉及空间数据的应用。

与研究基因表达随时间变化的遗传效应（“动态表达数量性状位点”）类似，使用单细胞 RNA 测序（Stegle 等，2015年）研究 eQTL 效应如何沿着连续的细胞环境（如细胞分化轨迹）变化，也引起了相当大的兴趣（van der Wijst 等，2020）。对这类数据进行 FASH 分析，可以先将细胞映射到一条一维轨迹上，然后将 FASH 应用于沿该轨迹不同位置估计的效应。

最后，关于 FASH 的一个实践要点是，通常更倾向于使用局部显著性度量（如 lfsr 和 lfsr），而不是累积显著性度量（如错误发现率和错误符号率）。

我们发现，当许多观测值的局部度量接近零时，这会拉低所有观测单位的累积平均值，从而使错误发现率（或错误符号率）较低，即使局部错误发现率（或局部错误符号率）很高也是如此。在这种情况下，与累积显著性度量相比，使用局部显著性度量可以减少错误发现。

6 利益披露声明

作者没有需要报告的利益冲突。

7 数据可用性声明

动态 eQTL 分析所用的汇总统计量见在线补充材料，复现结果的代码可在 GitHub 上获取。

8 致谢

作者感谢 Kenneth Barr 为第 4 节分析的心肌细胞数据提供支持。

9 资助

本研究部分经费由美国国家科学基金会（NSF）（DMS-2235451）和西蒙斯基金会（MPS-NITMB-00005320）拨款资助，受资助方为美国国家科学基金会—西蒙斯生物学理论与数学国家研究所（NITMB）。

补充材料

补充文本和图表：更多证明和推导、FASH 在模拟数据集上的评估、FASH 实现细节以及补充图表。（PDF 文件）

补充数据：第 4 节分析的所有时间点上所有基因-变异对的 eQTL 效应估计汇总统计量。（ZIP 文件）

参考文献

Benjamini, Y. 和 Hochberg, Y.（1995）。控制错误发现率：一种实用且有效的多重检验方法。*Journal of the Royal Statistical Society, SeriesB*，57(1)：289–300。

Cheng, C.-I. 和 Speckman, P. L.（2016）。平滑样条方差分析的贝叶斯因子。*Bayesian Analysis*，11(4)：957–975。

Cuomo, A. S., Seaton, D. D., McCarthy, D. J., Martinez, I., Bonder, M. J., Garcia-Bernardo, J., Amatya, S., Madrigal, P., Isaacson, A., Buettner, F. 等（2020）。对分化中的诱导多能干细胞进行单细胞 RNA 测序，揭示了基因表达的动态遗传效应。《自然通讯》，11：810。

DeRisi, J. L., Iyer, V. R. 和 Brown, P. O.（1997）。在基因组尺度上探索基因表达的代谢调控和遗传调控。《科学》，278(5338):680–686。

Efron, B.（2008）。微阵列、经验贝叶斯与双组模型。《统计科学》，23(1):1–22。

Efron, B.（2009）。大规模预测问题的经验贝叶斯估计。美国统计学会会刊，104(487):1015–1028。

Elorbany, R., Popp, J. M., Rhodes, K., Strober, B. J., Barr, K., Qi, G., Gilad, Y., 和 Battle, A.（2022年）。单细胞测序揭示了人类心肌细胞分化过程中谱系特异性的基因表达动态遗传调控。*PLoS Genetics*，18(1)：e1009666。

Francesconi, M. 和 Lehner, B.（2014）。遗传变异对发育过程中基因表达动态的影响。《自然》，505(7482)：208–211。Gamazon, E. R., Huang, R. S., Dolan, M. E., Cox, N. J., 和 Im, H. K.（2013）。整合基因组学：利用多个来源的高通量数据量化表型-基因型关系的显著性。《遗传学前沿》，3。Gutierrez-Arcelus, M., Baglaenko, Y., Arora, J., Hannes, S., Luo, Y., Amariuta, T., 等（2020）。在人类白细胞抗原及其他自身免疫位点，T细胞活化过程中等位基因特异性表达会动态变化。《自然遗传学》，52(3)：247–253。Heller, R. A., Schena, M., Chai, A., Shalon, D., Bedilion, T., Gilmore, J., Woolley, D. E., 和 Davis, R. W.（1997）。利用cDNA微阵列发现并分析炎症性疾病相关基因。《美国国家科学院院刊》，94(6)：2150–2155。

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., 等（2000）。通过表达谱汇编发现泛函。《细胞》，102(1):109–126。

Kang, J. B., Shen, A. Z., Gurajala, S., Nathan, A., Rumker, L., Aguiar, V. R. C. 等（2023）。在单细胞分辨率下绘制人类白细胞抗原基因的动态遗传调控架构。《自然遗传学》，55(12):2255–2268。Kim, Y., Carbonetto, P., Stephens, M. 和 Anitescu, M.（2020）。使用序列二次规划进行混合比例最大似然估计的快速算法。*Journal of Computational and Graphical Statistics*，29(2):261–273。Li, G. H., Shi, Y., Chen, Y., Sun, M., Sader, S., Maekawa, Y. 等（2009）。凝溶胶蛋白通过DNase I介导的细胞凋亡调控心肌梗死后的心脏重构。《循环研究》，104(7):896–904。Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P. 和 Tamayo, P.

（2015a）。分子特征数据库标志性基因集合集。*Cell Systems*，1(6):417–425。Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P. 和 Tamayo, P.（2015b）。分子特征数据库标志性基因集合集。*Cell Systems*，1(6):417–425。Lindgren, F. 和 Rue, H.（2008）。关于不规则位置的二阶随机游走模型。斯堪的纳维亚统计学杂志，35(4):691–700。Lindgren, F., Rue, H. 和 Lindström, J.（2011）。高斯场与高斯马尔可夫随机场之间的明确联系：随机偏微分方程方法。英国皇家统计学会会刊*B*辑：统计方法学，73(4):423–498。Liu, Y., Carbonetto, P., Takahama, M., Gruenbaum, A., Xie, D., Chevrier, N. 和 Stephens, M.（2024）。相关计数数据的灵活模型及其在单细胞RNA测序数据多条件差异表达分析中的应用。应用统计学年刊，18(3):2551–2575。Lu, M. 和 Stephens, M.（2019）。考虑估计标准误不确定性的正态均值经验贝叶斯估计。*arXiv*，1901.10679。MacKay, D. J.（2003）。信息论、推断与学习算法。剑桥大学出版社。Mazzotta, S., Neves, C., Bonner, R. J., Bernardo, A. S., Docherty, K. 和 Hoppler, S.（2016）。经典Wnt信号通路和非经典Wnt信号通路在人胚胎心肌细胞发育中的独特作用。*Stem Cell Reports*，7(4):764–776。Nathan, A., Asgari, S., Ishigaki, K., Valencia, C., Amariuta, T., Luo, Y. 等（2022）。单细胞eQTL模型揭示疾病位点的动态T细胞状态依赖性。《自然》，606(7912):120–128。

Polson, N. G. 和 Scott, J. G.（2010年）。全局收缩，局部作用：稀疏贝叶斯正则化与预测。贝叶斯统计，9。Särkkä, S. 和 Solin, A.（2019年）。应用随机微分方程，第10卷。剑桥大学出版社，英国剑桥。Servin, B. 和 Stephens, M.（2007年）。基于插补的关联研究分析：候选区域和数量性状。*PLoS Genetics*，3(7)：e114。Shepp, L. A.（1966年）。高斯测度的拉东-尼科迪姆导数。*Annals of Mathe-matical Statistics*，第321–354页。

Soskic, B.、Cano-Gamez, E.、Smyth, D. J.、Ambridge, K.、Ke, Z.、Matte, J. C. 等（2022年）。免疫疾病风险变异调控 cd4+ t 细胞活化期间的基因表达动态。《自然遗传学》，54(6)：817–826。

Stegle, O., Teichmann, S. A., and Marioni, J. C. (2015年). 单细胞转录组学中的计算与分析挑战. 《自然综述：遗传学》, 16(3):133–145. Stephens, M. (2017年). 假发现率：新方案. 《生物统计学》, 18(2):275–294. Storey, J. D. (2002年). 一种直接处理假发现率的方法. *Journal of the RoyalStatistical Society, Series B*, 64(3):479–498. Strober, B., Elorbany, R., Rhodes, K., Krishnan, N., Tayeb, K., Battle, A., and Gilad, Y. (2019年). 细胞分化过程中的基因表达动态遗传调控.《科学》, 364(6447):1287–1290. Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). 基因集富集分析：一种解读全基因组表达谱的知识驱动方法. 《美国国家科学院院刊》, 102(43):15545–15550. Urbut, S. M., Wang, G., Carbonetto, P., and Stephens, M. (2019年). 灵活的统计

方法，用于估计和检验基因组学研究中多种条件下的效应.《自然遗传学》, 51:187–195. van der Wijst, M. G., de Vries, D. H., Groot, H. E., Trynka, G., Hon, C.-C., Bonder, M., Stegle, O., Nawijn, M., Idaghdour, Y., Van Der Harst, P., et al. (2020). 单细胞 eqtlgen 联盟. ***eLife***, 9:e52155. Vovk, V. and Wang, R. (2021). E 值：校准、组合与应用. *Annalsof Statistics*, 49(3):1736–1754. Wahba, G. (1978). 非正规先验、样条平滑以及回归中防范模型误差的问题. *Journal of the Royal Statistical Society, Series B*, 40(3):364–372. Willwerscheid, J., Carbonetto, P., and Stephens, M. (2025). ebnm：一个使用多种先验族求解经验贝叶斯正态均值问题的 R 包. *Journal ofStatistical Software*, 114(3):1–32. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., and Yu, G. (2021). clusterprofiler 4.0：一种用于解读组学数据的通用富集分析工具. *The Innovation*, 2(3):100141. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012年). clusterProfiler：一个用于比较不同基因簇生物学主题的 R 包. *OMICS: A Journal of IntegrativeBiology*, 16(5):284–287. Yue, Y. R., Simpson, D., Lindgren, F., and Rue, H. (2014). 使用随机微分方程的贝叶斯自适应平滑样条. *Bayesian Analysis*, 9(2):397–424.

Zhang, Z.、Brown, P. 和 Stafford, J.（2025）。使用季节性高斯过程对准周期数据进行高效建模。*Statistics and Computing*， 35:32。

Zhang, Z., Stringer, A., Brown, P. 和 Stafford, J. (2024). 基于模型的平滑：使用积分维纳过程和重叠样条。*Journal of Computational andGraphical Statistics*, 33(3):883–895.

函数效应的经验贝叶斯收缩及其在动态表达数量性状位点分析中的应用：补充文本与图表

Ziang Zhang、Peter Carbonetto 和 Matthew Stephens

附录 A：实现细节

本节将详细介绍 FASH 在各个主要计算步骤中的实现细节。

σ_k的网格选择

我们固定 $\sigma_0 = 0$ 以表示零分量，并按照 Stephens (2017) 中的策略构造其余的 K 网格值$\{\sigma_k\}_{k=1}^K$。目标是创建一个足够稠密、覆盖区间 $[\sigma_{\min}, \sigma_{\max}]$的网格，使推断结果保持稳定，并且在使用更大或更稠密的网格时不会发生明显变化。

从概念上讲，σ_k 控制相对于基线模型 S_0的偏离。其实际解释取决于定义 L-高斯过程的算子 L 以及测量尺度。为了在各种应用中提供更易解释且一致的尺度，我们考虑一种等价的重新参数化方式，以 h-单元预测标准差（功率谱密度，PSD）（Zhang 等人，2024，2025）表示，定义如下：

$$\sigma(h) = \mathrm{SD}[\beta(t+h) \mid \beta(s) : s \le t], \quad \beta \sim L\text{-GP}(\beta; \sigma). \tag{20}$$

功率谱密度 $\sigma(h)$ 是原始 σ 的正比例缩放，但它具有直接的函数变异性解释，并且在不同的 L 选择之间具有可比性。

默认情况下，我们基于单单元功率谱密度 $\sigma(1)$ 选择 $[\sigma_{\min}, \sigma_{\max}]$，网格值$\{\sigma_k\}_{k=1}^K$在对数精度尺度上等距分布，即 $-2\log \sigma_k(1)$ 的取值范围为 0 到 10。虽然可以探索其他设定

根据我们的经验，这一默认设定的表现一直不错。

计算似然矩阵 L

FASH 的主要计算步骤是计算似然矩阵 $\mathbf{L} \in \mathbb{R}^{J\times(K+1)}$。对于每个条目，

$$\begin{aligned}\mathbf{L}_{jk} &= p_k(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j) \\ &= \int p(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{\beta}_j, \boldsymbol{s}_j)\, p(\boldsymbol{\beta}_j \mid \sigma_k)\, d\boldsymbol{\beta}_j \\ &= d\mathcal{N}(\hat{\boldsymbol{\beta}}_j; \mathbf{0}, \mathbf{C}_k + \mathbf{S}_j),\end{aligned} \tag{21}$$

其中，$\mathbf{C}_k$ 是 $\boldsymbol{\beta}_j$ 在 $\beta_j \sim L$-高斯过程$(\beta_j; \sigma_k)$ 下的协方差，$\mathbf{S}_j$ 是平方标准误构成的对角矩阵。

计算公式 (21) 需要精度矩阵 $(\mathbf{C}_k + \mathbf{S}_j)^{-1}$。当 R_j 较大时，直接求逆的计算开销很大，时间复杂度为 $O(R_j^3)$，内存复杂度为 $O(R_j^2)$。然而，与标准高斯过程不同，L-高斯过程由微分算子构造而成，因此具有马尔可夫结构（Särkkä 和 Solin，2019）；利用这一结构可以将计算复杂度降至 $O(R_j)$。实现方法包括将导数与 $\boldsymbol{\beta}_j$ 一并扩增，或采用有限元近似（Lindgren 和 Rue，2008；Zhang 等人，2024、2025）。本文采用后一种方法，默认使用 20 个等间距的 O样条。

优化经验贝叶斯先验 $\hat{}g_\beta$

经验贝叶斯（EB）估计$\hat{}g_\beta$ 通过最大化方程（11）中的对数似然得到。为促使 π_0的估计更为保守，可以添加基于狄利克雷的惩罚（Stephens，2017）：

$$h(\boldsymbol{\pi}; \lambda) = \prod_{k=0}^{K} \pi_k^{\lambda_k - 1}, \quad \lambda_k \geq 1, \tag{22}$$

从而得到惩罚对数似然

$$l(\boldsymbol{\pi}) + \log h(\boldsymbol{\pi}) = \sum_{j=1}^{J} \log\Big(\sum_{k=0}^{K} \pi_k \mathbf{L}_{jk} \Big) + \sum_{k=0}^{K} (\lambda_k - 1) \log \pi_k. \tag{23}$$

这个凸优化问题可以通过约束算法高效求解；我们采用 Kim 等人（2020）的序列二次规划方法，使用 $\lambda_0 = 10$ 和$\lambda_k = 1$ 来处理 $k \neq 0$。

尽管公式 (23) 假设各单元之间相互独立，但在存在依赖关系时，它可以解释为复合似然，而经验贝叶斯估计 $\hat{\boldsymbol{\pi}}$ 通常仍保持一致性（Urbut 等，2019）。

后验计算

给定 $\mathbf{L}$ 和 $\hat{g}_\beta$后，每个 β_j 的后验分布由公式 (12) 中的混合分布给出。每个$\hat{}$分量后验分布 $p_k(\beta_j \mid \boldsymbol{\beta}, \boldsymbol{s})$ 都是一个高斯过程，因此总体后验分布是多个 $K+1$高斯过程的混合分布。实际上，$\hat{\boldsymbol{\pi}}$ 通常是稀疏的，因此只剩下少数非平凡的混合成分，从而减少存储后验过程所需的内存。

为计算泛函$\mathcal{F}$在式(17)中的局部错误符号率，我们从后验混合分布的非平凡分量中抽取 M 条独立样本路径，并通过蒙特卡罗近似相关概率。 L高斯过程的有限元表示使路径模拟即使在较大的 M下也具有计算效率；我们在分析中默认使用 $M = 3000$ 。

Ap附录 B：补充证明与推导

定理1的证明

定理1 （BF 校正给出保守估计）。 假设$\hat{}\pi_0$ 是由算法*1*得到的 $\pi_0 = J_0/J$ 的校正估计，且这 J_0 个零效应独立同分布于自适应先验中指定的零分布。则对于先验中备择分布（分量1 至 K）的任意设定以及任意缓冲区 $\epsilon > 0$,

$$\hat{\pi}_0 \;\geq\; \pi_0 \quad \textit{almost surely as } J_0 \to \infty.$$

证明。 根据引理 1，在零假设下 $\mathbb{E}_0$(BFj) $= 1$ 且 $\mathrm{BF}_j > 0$。因此，根据强大数定律，

$$\lim_{J_0\to\infty} \frac{1}{J_0} \sum_{j\in\mathcal{H}_0} \mathrm{BF}_j \;=\; 1 \quad \text{almost surely.} \tag{24}$$

从现在起，所有渐近论证均默认几乎处处成立，因此我们省略符号说明。

定义由 c^* 诱导的阈值化集合:

$$\mathcal{T}_0 = \{j : \mathrm{BF}_j < c^*\}, \qquad \mathcal{T}_1 = \{j : \mathrm{BF}_j \geq c^*\}.$$

令 $J_{a,\ b} = |\mathcal{H}_a \cap \mathcal{T}_b|$ 对 $a,\ b \in \{0,\ 1\}$成立，并设

$$S = \sum_{j\in\mathcal{H}_0\cap\mathcal{T}_0} \mathrm{BF}_j.$$

则$J_{0,0} + J_{0,1} = J_0$，以及

$$\hat{\pi}_0 = \frac{J_{0,0} + J_{1,0}}{J}, \qquad \pi_0 = \frac{J_{0,0} + J_{0,1}}{J}. \tag{25}$$

从而证明$\hat{}\pi_0 \geq \pi_0$ 等价于表明$J_{1,0} \geq J_{0,1}$。

根据 (24)，存在一个正整数 $J\prime_0$，使得对所有 $J_0 > J'_0$,

$$1 + \frac{\epsilon}{2} \;\geq\; \frac{\sum_{j\in\mathcal{H}_0} \mathrm{BF}_j}{J_0} = \frac{\sum_{j\in\mathcal{H}_0\cap\mathcal{T}_1} \mathrm{BF}_j + S}{J_{0,1} + J_{0,0}} \;\geq\; \frac{c^* J_{0,1} + S}{J_{0,1} + J_{0,0}}, \tag{26}$$

其中最后一个不等式成立，因为 $\mathrm{BF}_j \geq c^*$ 在 $\mathcal{T}_1$ 上。

根据算法1中 c^* 的定义，我们有 $\mu(c^*) \geq 1+\epsilon$，因此

$$1+\frac{\epsilon}{2} \;\leq\; \mu(c^*) = \frac{\sum_{j\in\mathcal{T}_0}\mathrm{BF}_j}{J_{1,0}+J_{0,0}} = \frac{\sum_{j\in\mathcal{H}_1\cap\mathcal{T}_0}\mathrm{BF}_j + S}{J_{1,0}+J_{0,0}} \;\leq\; \frac{c^*J_{1,0}+S}{J_{1,0}+J_{0,0}}, \tag{27}$$

其中最后一个不等式成立，因为贝叶斯因子$_j < c^*$ 在$\mathcal{T}_0$上。

结合(26)和(27)，对所有足够大的 J_0,

$$\frac{c^*J_{0,1}+S}{J_{0,1}+J_{0,0}} \;\leq\; 1+\frac{\epsilon}{2} \;\leq\; \frac{c^*J_{1,0}+S}{J_{1,0}+J_{0,0}}. \tag{28}$$

定义

$$f(x) = \frac{c^*x+S}{x+J_{0,0}}, \qquad f'(x) = \frac{c^*J_{0,0}-S}{(x+J_{0,0})^2}.$$

由于贝叶斯因子$_j < c^*$ 在$\mathcal{T}_0$, 上，我们有$S < c^*J_{0,0}$。因此，对 $x > 0$有 $f'(x) > 0$ ，故 f 严格递增。由(28)， $f(J_{0,1}) \leq f(J_{1,0})$，这意味着$J_{1,0} \geq J_{0,1}$。

最后，由(25)，由此得到$\hat{}\pi_0 \geq \pi_0$，证明完毕。 □

后验优势比的保守性

在本小节中，我们将提供更多细节，说明高估 π_0 会使支持 H_0 的推断更加保守，即使备择假设设定错误也是如此。

为简单起见，假设 $\tilde{\pi}_0$ 与 $0 < \pi_0 \leq \tilde{\pi}_0 < 1$ 固定，并记

$$p_0 = p_0(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j), \quad p_1 = p_1(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j), \quad \tilde{p}_1 = \tilde{p}_1(\hat{\boldsymbol{\beta}}_j \mid \boldsymbol{s}_j),$$

分别为零假设、真实备择假设和错误设定的备择假设的边际密度。

衡量反对 H_0 的证据强度的一种有用指标是后验优势比（PO），其定义为

$$\mathrm{PO} = \frac{p(H_1 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j)}{p(H_0 \mid \hat{\boldsymbol{\beta}}_j, \boldsymbol{s}_j)} = \frac{1-\pi_0}{\pi_0}\,\frac{p_1}{p_0}. \tag{29}$$

~~将 π_0 和 p_1 替换为经过调整/设定错误的对应量后，可得拟合的 PO：

$$\widetilde{\mathrm{PO}} = \frac{1-\tilde{\pi}_0}{\tilde{\pi}_0}\,\frac{\tilde{p}_1}{p_0}. \tag{30}$$

~~首先，很容易证明，在零假设下，拟合的 PO 在期望意义上比真实的 PO 更保守：

引理 S1： 若 $\tilde{\pi}_0 \geq \pi_0$，则

$$\mathbb{E}_{H_0}[\widetilde{\mathrm{PO}}] \;\leq\; \mathbb{E}_{H_0}[\mathrm{PO}].$$

证明。 由 (29)-(30)，并对 p_0取期望，

$$\mathbb{E}_{H_0}[\widetilde{\mathrm{PO}}] = \frac{1-\tilde{\pi}_0}{\tilde{\pi}_0}\,\mathbb{E}_{p_0}\left[\frac{\tilde{p}_1}{p_0}\right] = \frac{1-\tilde{\pi}_0}{\tilde{\pi}_0}, \quad \mathbb{E}_{H_0}[\mathrm{PO}] = \frac{1-\pi_0}{\pi_0}\,\mathbb{E}_{p_0}\left[\frac{p_1}{p_0}\right] = \frac{1-\pi_0}{\pi_0},$$

因为 $\mathbb{E}_{H_0}\left[\frac{r_{(\hat{\boldsymbol{\beta}})}}{p_0(\hat{\boldsymbol{\beta}})}\right] = \int \frac{r_{(\hat{\boldsymbol{\beta}})}}{p_0(\hat{\boldsymbol{\beta}})} p_0(\hat{\boldsymbol{\beta}})\, d\hat{\boldsymbol{\beta}} = \int r(\hat{\boldsymbol{\beta}})\, d\hat{\boldsymbol{\beta}} = 1$ 对任意密度 r均成立。由于 $x \mapsto \frac{1-x}{x}$ 在 (0,1) 上严格递减，且 $\tilde{\pi}_0 \geq \pi_0$，结论得证。 □

在备择假设下， 保守性可通过对 log-PO 应用类似论证得出 .

引理 S2： 若 $\tilde{\pi}_0 \geq \pi_0$，则

$$\mathbb{E}_{H_1}\Big[\log \widetilde{\mathrm{PO}}\Big] \;\leq\; \mathbb{E}_{H_1}\Big[\log \mathrm{PO}\Big].$$

证明。在 H_1的条件下，

$$\begin{aligned}
\mathbb{E}_{H_1}[\log \widetilde{\mathrm{PO}}] &= \log\frac{1-\tilde{\pi}_0}{\tilde{\pi}_0} + \mathbb{E}_{p_1}\left[\log\frac{\tilde{p}_1}{p_0}\right] \\
&= \log\frac{1-\tilde{\pi}_0}{\tilde{\pi}_0} + \mathbb{E}_{p_1}\left[\log\frac{p_1}{p_0}\right] - \mathbb{E}_{p_1}\left[\log\frac{p_1}{\tilde{p}_1}\right] \\
&= \log\frac{1-\tilde{\pi}_0}{\tilde{\pi}_0} + D_{\mathrm{KL}}(p_1\|p_0) - D_{\mathrm{KL}}(p_1\|\tilde{p}_1) \\
&\leq \log\frac{1-\tilde{\pi}_0}{\tilde{\pi}_0} + D_{\mathrm{KL}}(p_1\|p_0) \;\leq\; \log\frac{1-\pi_0}{\pi_0} + D_{\mathrm{KL}}(p_1\|p_0) = \mathbb{E}_{H_1}[\log \mathrm{PO}],
\end{aligned}$$

因为 $D_{\mathrm{KL}}(\cdot\|\cdot) \geq 0$ 且 $x \mapsto \log\frac{1-x}{x}$ 在 (0,1) 上严格递减，且 $\tilde{\pi}_0 \geq \pi_0$。 □

附录 C：模拟研究

我们开展了一项模拟研究，以评估 FASH 结合所提出的基于贝叶斯因子的调整，在估计零假设比例 π_0时的表现。两个假设设定分别是检验效应函数是否为常数或线性，这与第 4 节一致。我们生成了 $J = 1000$ 个相互独立的观测单位，每个单位的效应函数 β_j 均从以下三类之一中抽取：

A. 非动态：$\beta_j = c_j$，其中 $c_j \sim N(0, 1)$。B. 线性动态：$\beta_j = c_j + b_j t$，其中 $c_j \sim N(0, 1)$ 和 $b_j \sim N(0, 1/4)$。C. 非线性动态：β_j，从标准 $\mathrm{IWP}_2(\sigma)$ 中抽样得到，其中 $\sigma(16) = 5$。

每个函数均在十六个等间隔时间点进行观测：$t = 0, 1、\ldots\ldots、16$。标准误 $s_{j,\,r}$ 独立地从{0。1，0。3，0。5} 中等概率抽取。模拟观测值的示例见图S1。为改变基础零假设比例，我们引入了参数 $\rho \in [0.05, 0.5]$，增量为 0。01。对于每个 ρ，我们从类别 A 模拟 $J(1-\rho)$ 个观测值，从 B 类模拟 $J\rho/2$ 个观测值，其余观测值来自 C 类。因此，当检验 S_0 为常函数时，真实零假设比例为 $\pi_0 = 1 - \rho$；当检验 S_0 为线性函数时，真实零假设比例为 $\pi_0 = 1 - \rho/2$。

我们按照第 4 节中的方法拟合 FASH 模型，并比较无惩罚极大似然估计 $\hat{}\pi_0$ 与其 BF调整后的对应估计。结果汇总于图 S2。面板 (a-b) 显示，在各次重复中，BF调整后的估计始终高于真实的 π_0，证实了定理1所述的保守性质。相比之下，原始极大似然估计常常低估 π_0，尤其是在检验 S_0 是否为常函数时。面板 (c-d) 进一步表明，针对 σ 使用更密网格不会影响这一结论：BF调整后的估计仍保持了所需的保守性。

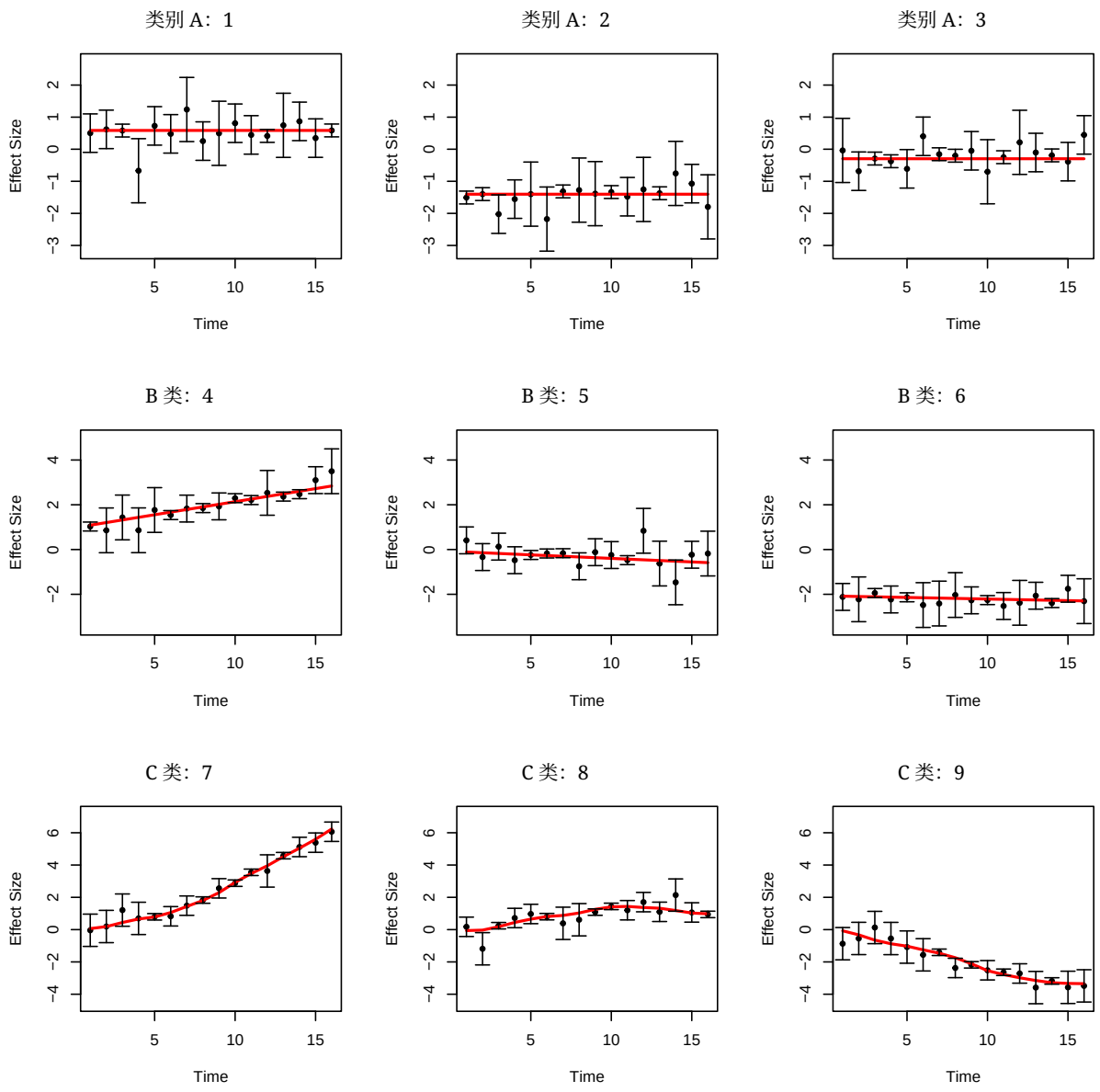


图 S1：模拟观测值示例：点表示观测到的效应估计值；竖线表示 ± 2 标准误；红线表示真实效应函数 β_j。

接下来，我们考察一个具有 $\rho = 0.2$ 的特定设定；在检验动态表达数量性状位点时，其对应的真实零假设比例为 $\pi_0 = 0.8$，在检验非线性动态 eQTL 时则为 $\pi_0 = 0.9$。我们按照第 4 节中的方法应用 FASH，以实现两个推断目标：(i) 检测动态表达数量性状位点；(ii) 检测非线性动态 eQTL。对于一系列名义 FDR 水平 α，我们评估采用和不采用 BF 校正的 FASH 的经验 FDR。如图 S3 所示，当 $\alpha \leq 0.05$ 时，即使不进行校正，经验 FDR 也已得到良好控制。

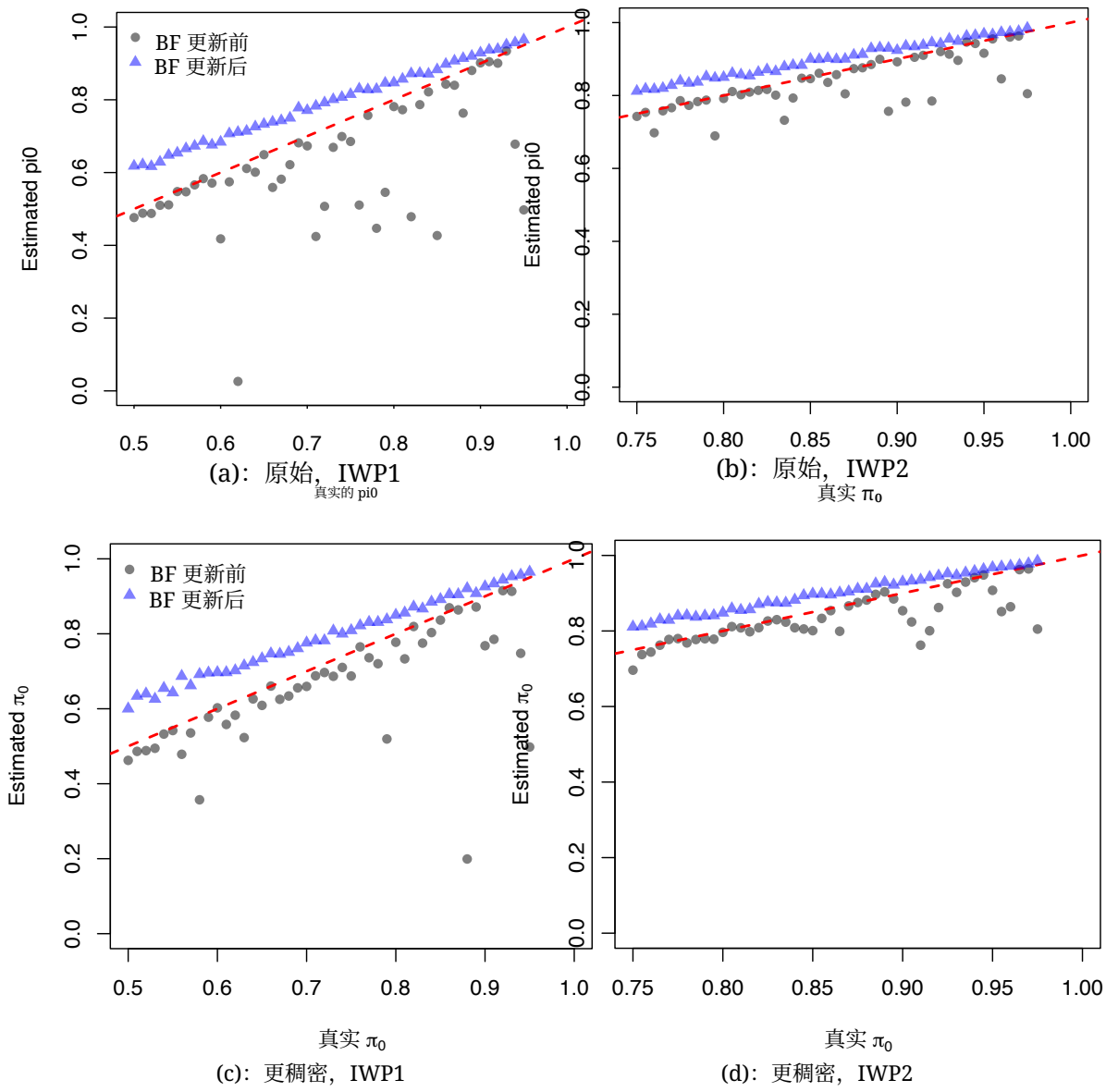


图 S2：模拟研究中估计的零假设比例 $\hat{\pi}_0$。左列对应于检验常函数（$S_0 =$ 常函数，采用 IWP_1 拟合）；右列对应于检验线性函数（$S_0 =$ 线性函数，采用 IWP_2 拟合）。黑色：最大似然估计；蓝色：贝叶斯因子调整估计；红色：45 度线。（a-b）使用第 4 节中的原始网格得到的结果（对数精度尺度上的间距为 0.2）。（c-d）使用更稠密网格得到的结果（间距为 0.1）。

然而，随着 α 增大，未校正的结果表现出膨胀的错误发现率，尤其是在检验非线性动态 eQTL 时。相比之下，BF 校正版本始终将错误发现率控制在名义水平或以下，这与定理 1的理论保证一致。

就检验效能而言，BF 校正会产生略为保守的结果，从而导致检验效能略有下降。然而，这一损失很小——在最坏情况下不超过约5%——并且当 $\alpha = 0.05$ 时，BF校正FASH在两项任务中仍能达到80%以上的检验效能。因此，我们认为这是确保错误发现率保持在名义水平以下的合理权衡，并建议在实践中应用BF 校正。 $\alpha = 0.05$ 处显著发现的示例见图S4，其中将FASH得到的后验效应函数与真实潜在效应进行了比较。

附录 D：iPSC 心肌细胞分化研究的补充细节

对于每个基因-变异对 $j \in [J]$，我们在 16 个时间点分别拟合线性回归模型，从而获得eQTL效应估计 $\hat{}\ \beta_j(t_r)$，每个时间点为 $r \in [16]$。我们使用了 Strober等人（2019）中采用的相同线性回归模型：

$$E_{jc}(t_r) \mid \boldsymbol{z}_{cr}, G_c \overset{\text{ind}}{\sim} N(\boldsymbol{z}_{cr}^T \boldsymbol{b}_{jr} + G_c \beta_j(t_r), \sigma^2_{E_{jr}}), \quad r \in [16], c \in [C_r], j \in [J], \tag{31}$$

其中， $E_{jc}(t_r) \in \mathbb{R}$ 表示细胞系 c 在时间点t_r的基因 j 的标准化表达量， $G_c \in \{0,1,2\}$ 表示细胞系 c中该单核苷酸多态性的基因型剂量， $\boldsymbol{z}_{cr} \in \mathbb{R}^6$ 是协变量向量（截距和前 5 个主成分），而 $\beta_j(t_r) \in \mathbb{R}$, $\boldsymbol{b}_{jr} \in \mathbb{R}^6$ 是待估计的回归系数。主成分（PC）根据维度为 $19 \times 212{,}147$ 的“按细胞系合并”的表达矩阵计算得出（Strober等，2019）。在时间 t_r 观测到的细胞系数量记为 C_r，大多数时间点为 19 个，但第 14 天为 18 个，第 3 天和第 5 天为 16 个。假设时间点 t_r 的残差方差 $\sigma^2_{E_{jr}}$ 在该时间点的各细胞系间保持恒定。根据式 (31)，我们计算了每天每个基因-变异对的 $\hat{}\ t$ 统计量： $T_{jr} = \beta_j(t_r)/s_{jr}$，其中 s_{jr} 是 $\beta_j(t_r)$的 ^标准误。

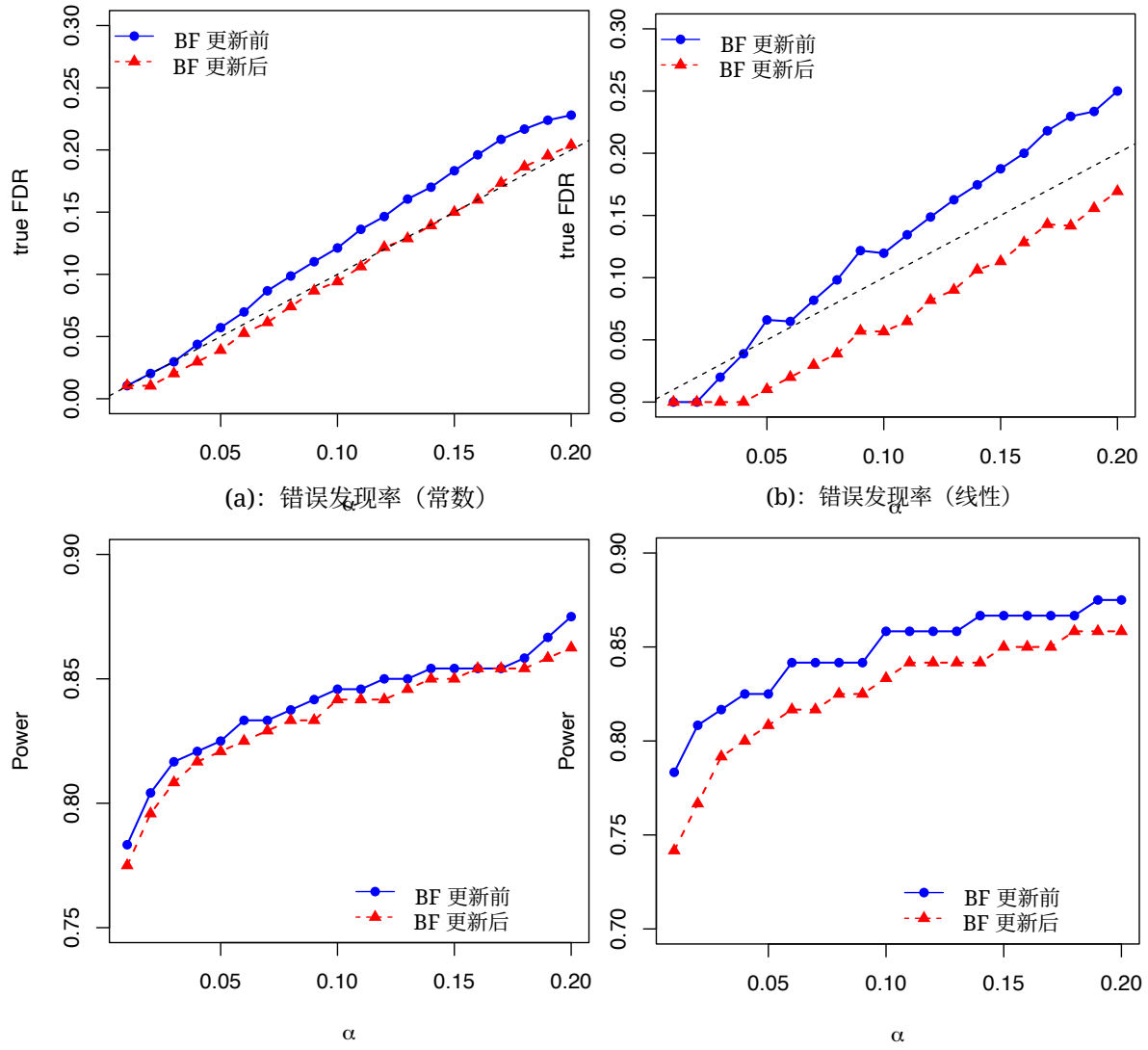
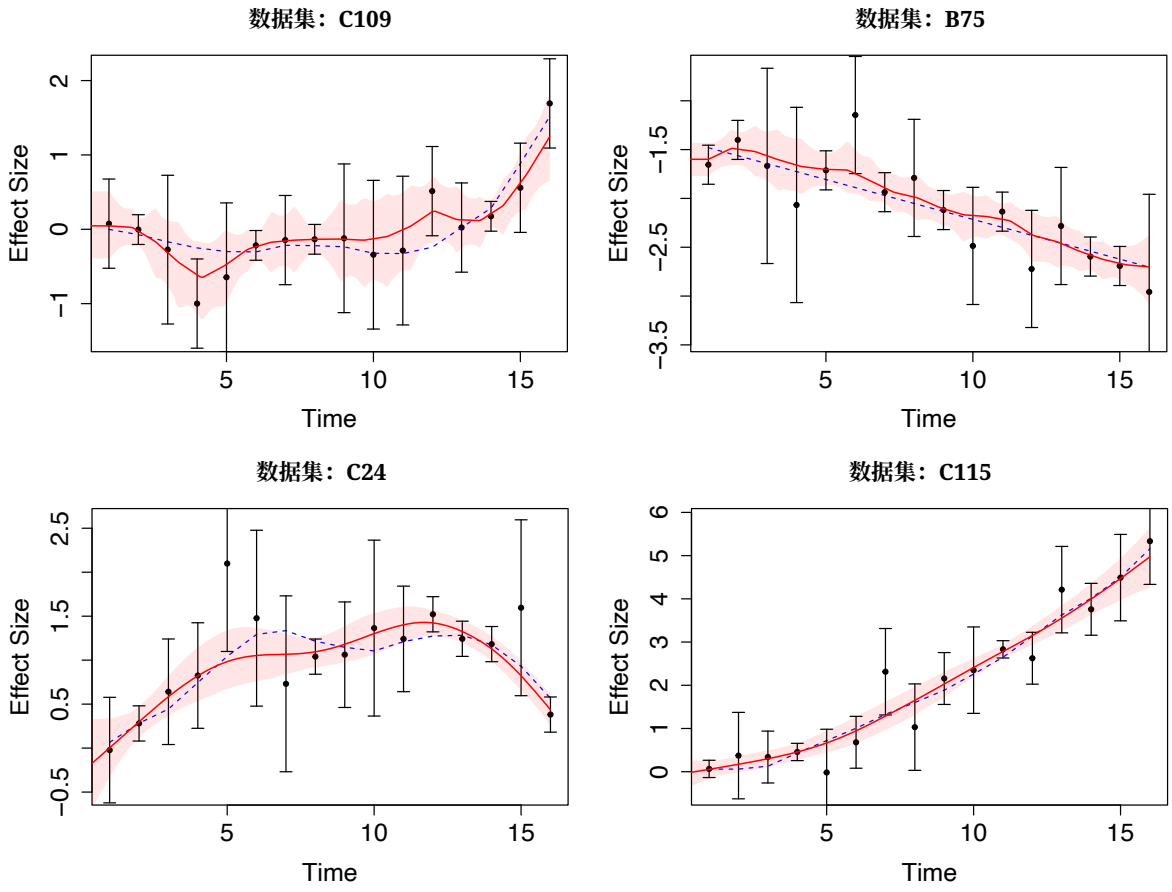


图 S3：在两种推断目标下，FASH 的经验 FDR（a-b）和检验效能（c-d）随名义 FDR α 的变化情况，其中红色表示经过 π_0 的 BF 校正，蓝色表示未校正。左列对应检验动态表达数量性状位点（$H_0 : \beta_j$ 常数），右列对应检验非线性动态（$H_0 : \beta_j$ 线性）。(a-b) 中的黑色虚线为用于错误发现率校准的 45 度参考线。

在零假设 $\beta_j(t_r) = 0$成立时，T_{jr} 服从具有 $\nu_r = C_r - 5$ 个自由度的 t 分布。如果每个时间点测量的细胞系数量较大，极大似然估计量的渐近正态性将意味着 T_{jr} 近似服从标准正态分布，因此 $\hat{\beta}_j(t_r) \sim N(\beta_j(t_r)\ s^2_{jr})$。然而，由于本研究仅纳入了 19 个细胞系，这种近似并不可靠。事实上，膨胀



图S4：在模拟研究中，FASH 在名义错误发现率为 $\alpha = 0.05$ 的水平下识别出的显著 eQTL 示例（上：动态表达数量性状位点；下：非线性<动态表达数量性状位点）。β_j 的后验均值以红色实线表示，95% 可信区间以红色阴影˙表示。观测到的效应大小估计值 β_j 以黑点表示，误差线表示$\pm 2\tilde{s}_{j,r}$。真实的底层效应函数以蓝色虚线表示。

在此设定下报告了 I类错误率（参见 Strober 等人 2019年 补充材料中的图S10）。为以可扩展的方式校正 I类错误率膨胀，我们遵循（Lu 和 Stephens，2019年），将“t-调整后的标准误” $\tilde{s}_{jr}$ 定义为满足

$$\Phi(\hat{\beta}_j(t_r)/\tilde{s}_{jr}) = P_{T_{\nu_r}}(\hat{\beta}_j(t_r)/s_{jr}), \tag{32}$$

其中， Φ 表示标准正态分布的累积分布函数， $P_{T_{\nu_r}}$ 表示具有 ν_r 个自由度的 tˆ分布的累积分布函数。效应估计值 β_{jr} 和调整后的标准误 $\tilde{s}_{jr}$ 随后作为 FASH 的输入。

FASH 分析中使用的先验（10）均定义在对数尺度上的等距网格上，其中 $\sigma_0 = 0$, $\sigma_1 = e^{-5} \approx 0.01$、 $\sigma_{52} = 1$和 log σ_{k+1} = log $\sigma_k + 0.1$、$k = 2, \ldots, 51$。

附录 E：补充图表

表 3：具有动态表达数量性状位点的重点基因（动态）和具有开关动态 eQTL 的基因（开关）的基因集富集结果。所有基因集均来自 MSigDB Hallmark（HALLMARK_*）集合；表中省略了前缀。

类别 基因集	基因比例	背景比例	p-value	q-value
动态 缺氧	25/1177	89/6362	0.0169	0.436
动态 伊利诺伊州6_JAK_STAT3_信号传导	8/1177	21/6362	0.0280	0.436
动态 雌激素_反应_早期	22/1177	82/6362	0.0394	0.436
动态雌激素_反应_晚期	21/1177	79/6362	0.0476	0.436
动态雄激素_反应	16/1177	57/6362	0.0499	0.436
动态 HEME_代谢	22/1177	85/6362	0.0564	0.436
动态 NOTCH_信号传导	5/1177	15/6362	0.1276	0.750
动态 KRAS_信号传导_DN	8/1177	28/6362	0.1306	0.750
动态 糖酵解	23/1177	100/6362	0.1498	0.750
动态 肌发生	16/1177	67/6362	0.1624	0.750
动态 UV_RESPONSE_DN	16/1177	68/6362	0.1780	0.750
动态 KRAS_信号_上调	11/1177	47/6362	0.2415	0.840
动态 TNFA_信号_通过_NFKB	13/1177	58/6362	0.2665	0.840
动态 HEDGEHOG_信号传导	4/1177	15/6362	0.2962	0.840
动态细胞凋亡	15/1177	70/6362	0.3072	0.840
动态干扰素_γ_反应	12/1177	56/6362	0.3359	0.840
动态有丝分裂_纺锤体	28/1177	139/6362	0.3399	0.840
动态 WNT_BETA_CATENIN_SIGNALING	5/1177	21/6362	0.3455	0.840

类别 基因集	基因比例	背景比例	p-value	q-value
动态 凝血	7/1177	31/6362	0.3459	0.840
动态 P53_通路	17/1177	83/6362	0.3626	0.840
动态 脂肪生成	20/1177	106/6362	0.5008	0.945
动态 活性_氧_物种_通路	5/1177	26/6362	0.5410	0.945
动态顶端_表面	3/1177	15/6362	0.5439	0.945
动态 PI3K_AKT_MTOR_信号传导	10/1177	55/6362	0.5794	0.945
动态补体	10/1177	57/6362	0.6278	0.945
动态 伊利诺伊州2_STAT5_信号传导	13/1177	75/6362	0.6499	0.945
动态 异生物质_代谢	12/1177	71/6362	0.6837	0.945
动态干扰素_α_反应	4/1177	27/6362	0.7633	0.945
动态顶端_连接	13/1177	82/6362	0.7739	0.945
动态 DNA_修复	13/1177	82/6362	0.7739	0.945
动态 血管生成	2/1177	15/6362	0.7956	0.945
动态上皮_间质_转化 13/1177		84/6362	0.8031	0.945
动态蛋白质_分泌	9/1177	61/6362	0.8208	0.945
动态同种异体移植物_排斥	5/1177	36/6362	0.8225	0.945
动态 TGF_BETA_信号通路	5/1177	36/6362	0.8225	0.945
动态 未折叠_蛋白质_反应	11/1177	75/6362	0.8442	0.945
动态脂肪酸_代谢	10/1177	72/6362	0.8812	0.945
动态MYC_靶基因_V2	5/1177	41/6362	0.8992	0.945
动态胆汁酸代谢	3/1177	28/6362	0.9132	0.945
动态精子发生	3/1177	28/6362	0.9132	0.945
动态 UV_响应_上调	10/1177	76/6362	0.9174	0.945
动态 炎症_反应	5/1177	43/6362	0.9206	0.945
动态氧化磷酸化_	16/1177	119/6362	0.9445	0.945
动态过氧化物酶体	4/1177	44/6362	0.9739	0.945

类别	基因集	基因比例	背景比例	p-value	q-value
动态	G2M_检查点	16/1177	136/6362	0.9881	0.945
动态	MTORC1_信号传导	13/1177	129/6362	0.9973	0.945
动态	胆固醇_稳态	2/1177	42/6362	0.9981	0.945
动态	E2F_靶标	10/1177	127/6362	0.9998	0.945
动态	MYC_TARGETS_V1	6/1177	124/6362	1.0000	0.945
开关	缺氧	11/250	89/6362	0.000668	0.0246
开关	KRAS信号传导 UP_ _	7/250	47/6362	0.00218	0.0348
开关	肌发生	8/250	67/6362	0.00445	0.0348
开关	P53通路_	9/250	83/6362	0.00501	0.0348
开关	糖酵解	10/250	100/6362	0.00565	0.0348
开关	雄激素反应_	7/250	57/6362	0.00656	0.0348
开关	凝血	5/250	31/6362	0.00662	0.0348
开关	IL6-JAK-STAT3信号通路_ _ _	4/250	21/6362	0.00822	0.0378
开关	NOTCH信号通路_	3/250	15/6362	0.0192	0.0787
开关	顶端连接_	7/250	82/6362	0.0415	0.153
开关	干扰素γ反应_ _	5/250	56/6362	0.0678	0.227
开关	活性氧_ _ _ 通路	3/250	26/6362	0.0802	0.246
开关	雌激素晚期反应_ _	6/250	79/6362	0.0894	0.253
开关	雌激素早期反应_ _	6/250	82/6362	0.1025	0.268
开关	上皮_间质_转化 6/250		84/6362	0.1117	0.268
开关	血红素代谢_	6/250	85/6362	0.1165	0.268
开关	细胞凋亡	5/250	70/6362	0.1398	0.299
开关	外源物质代谢_	5/250	71/6362	0.1459	0.299
开关	PI3K-AKT-mTOR信号通路_ _ _	4/250	55/6362	0.1690	0.328
开关	TNFα经NF-κB信号通路_ _ _	4/250	58/6362	0.1927	0.355
开关	MYC靶基因V2_ _	3/250	41/6362	0.2171	0.381

类别	基因集	基因比例	背景比例	p-value	q-value
开关	炎症反应	3/250	43/6362	0.2380	0.399
开关	紫外线响应下调	4/250	68/6362	0.2778	0.445
开关	KRAS信号传导下调	2/250	28/6362	0.3018	0.463
开关	未折叠蛋白反应	4/250	75/6362	0.3405	0.502
开关	补体	3/250	57/6362	0.3895	0.529
开关	MTORC1信号通路	6/250	129/6362	0.3964	0.529
开关	同种异体移植物排斥	2/250	36/6362	0.4164	0.529
开关	转化生长因子β信号传导	2/250	36/6362	0.4164	0.529
开关	刺猬信号通路	1/250	15/6362	0.4523	0.555
开关	有丝分裂纺锤体	6/250	139/6362	0.4670	0.555
开关	氧化磷酸化	5/250	119/6362	0.5046	0.581
开关	MYC 靶基因 V1	5/250	124/6362	0.5415	0.600
开关	WNT β-连环蛋白信号传导	1/250	21/6362	0.5697	0.600
开关	IL2 STAT5 信号通路	3/250	75/6362	0.5705	0.600
开关	脂肪生成	4/250	106/6362	0.6043	0.618
开关	G2M 检查点	5/250	136/6362	0.6244	0.622
开关	干扰素α反应	1/250	27/6362	0.6620	0.642
开关	脂肪酸代谢	2/250	72/6362	0.7817	0.736
开关	胆固醇稳态	1/250	42/6362	0.8153	0.736
开关	过氧化物酶体	1/250	44/6362	0.8297	0.736
开关	DNA修复	2/250	82/6362	0.8392	0.736
开关	蛋白质分泌	1/250	61/6362	0.9143	0.783
开关	紫外线应答上调	1/250	76/6362	0.9534	0.798
开关	E2F靶基因	1/250	127/6362	0.9942	0.814

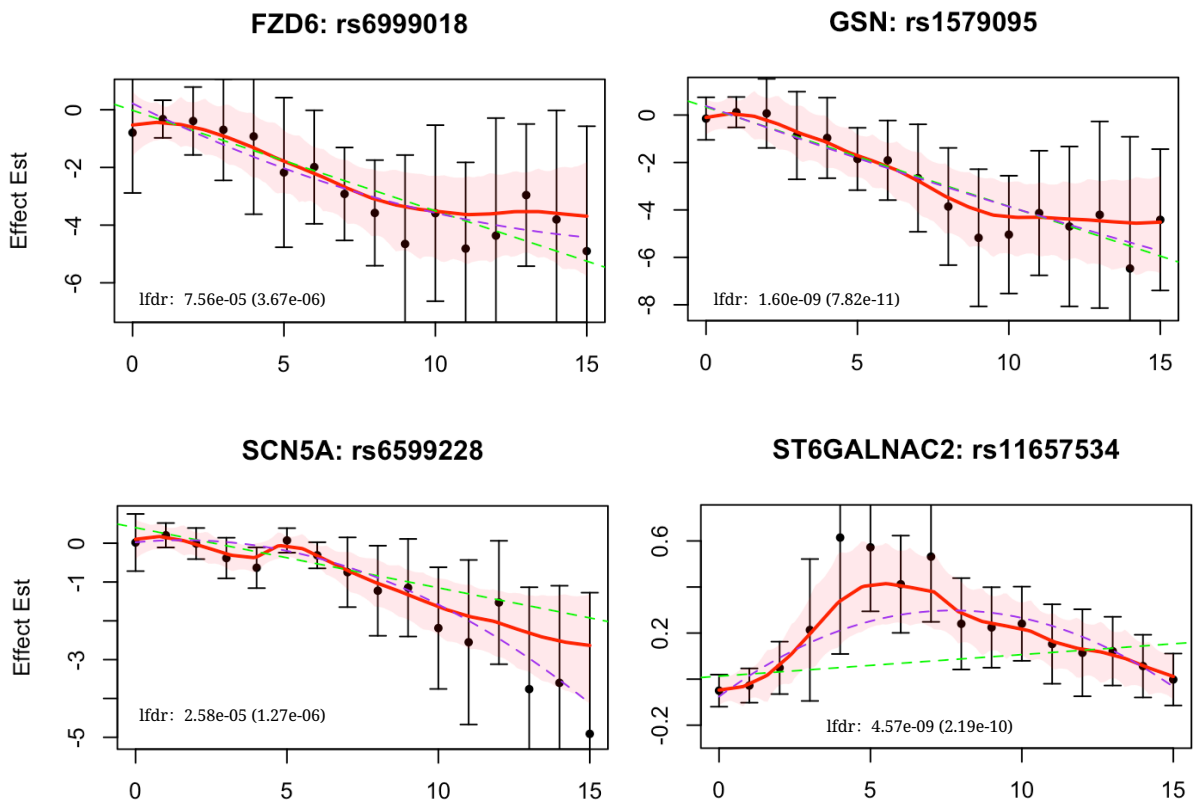


图 S5：基于 5% 错误发现率阈值筛选出的具有显著动态 eQTL 效应的基因-变异对。每个图中，β_j 的后验均值以红色实线表示，95% 可信区间以红色阴影区域表示。观测到的效应量估计值 $\hat{}\beta_j$ 以黑点表示，竖线表示 ± 2 （校正后的）标准误$\tilde{s}_{jr}$。在“lfdr：x (y)”中，x 表示贝叶斯因子校正后的 lfdr，y 表示贝叶斯因子校正前的 lfdr。作为比较，$\hat{}\beta_j$ 的逆方差加权最小二乘估计以绿色（线性）和紫色（二次）虚线表示。变异 rs6999018 和 rs1579095 在分化的前三天似乎对各自基因的表达影响甚微，*FZD6* 和 *GSN*，但在后期开始显著下调表达。这些基因此前已被认为与心肌细胞发育和功能有关。例如，Mazzotta 等（2016）报告称，*FZD6* 通过非经典 Wnt 信号通路在心肌细胞分化中发挥重要的调控作用，而 Li 等（2009）则表明，*GSN* 通过 DNase I 介导的细胞凋亡参与心脏重构。

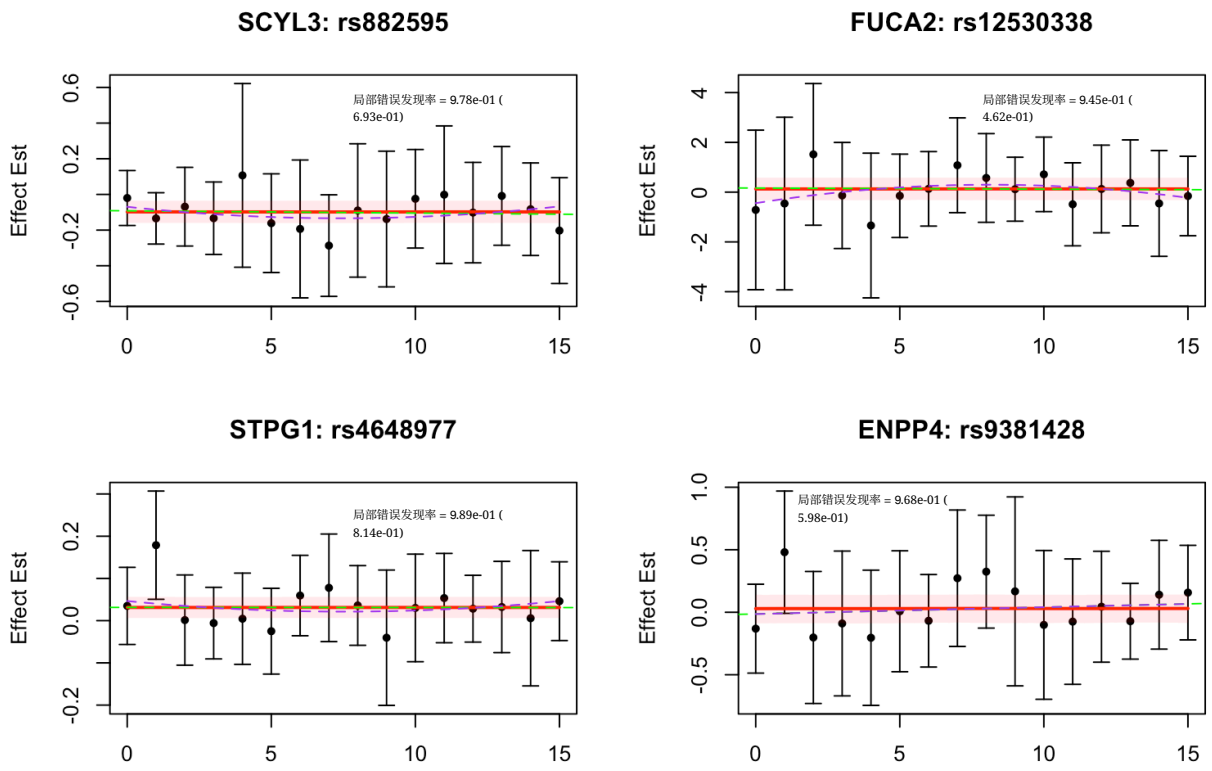


图 S6：根据 5% 错误发现率阈值筛选出的、未显示显著动态 eQTL 效应的变异-基因对。β_j 的后验均值以红色实线表示，95% 可信区间以红色阴影区域表示。观测到的效应大小估计值 β_j 以黑点表示，垂直线段表示$\pm 2\tilde{s}_{j,\,r}$。图中展示了每对的局部错误发现率，括号内为贝叶斯因子校正前的数值。作为比较，β_j 的逆方差加权最小二乘估计以绿色虚线（线性）和紫色虚线（二次）表示。

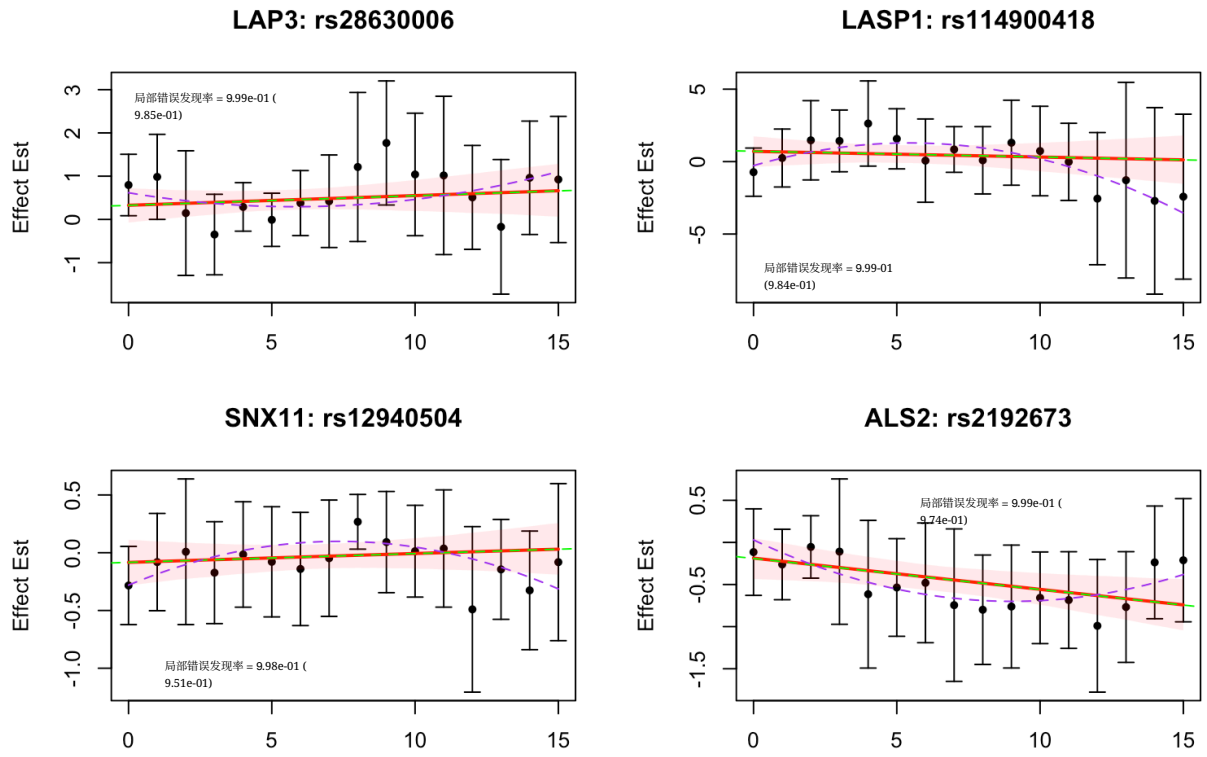
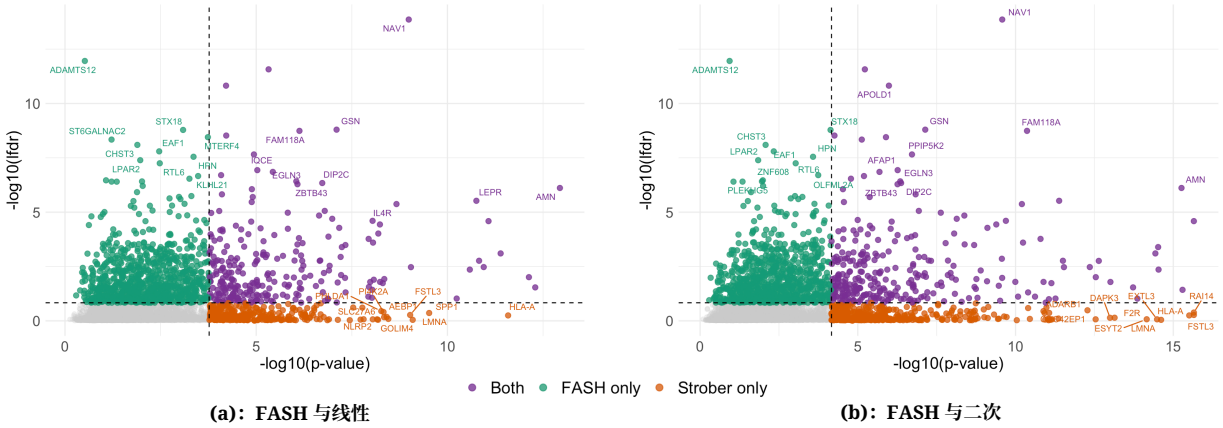
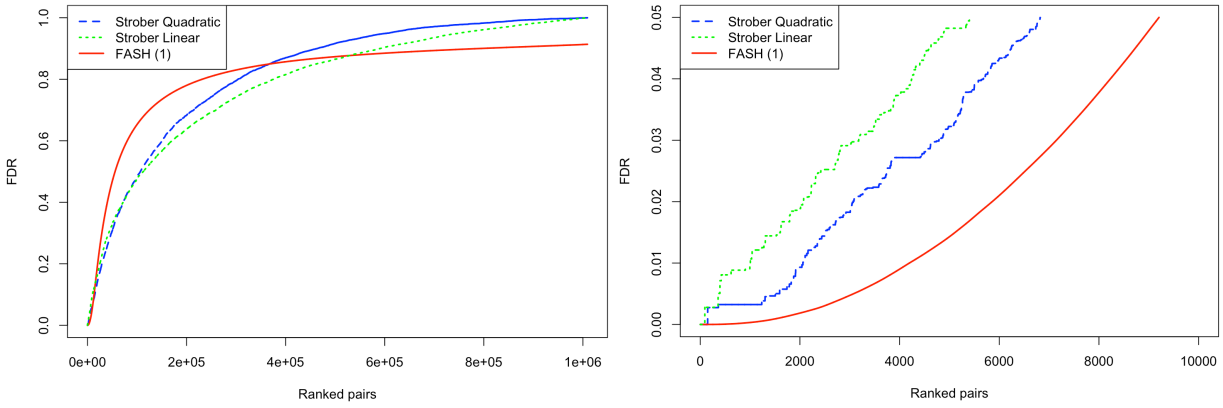


图 S7：根据 5% 错误发现率阈值，选出的未显示显著非线性动态 eQTL 效应的变异-基因对。 β_j 的后验均值以红色实线表示，95% 可信区间以红色阴影区域表示。观测到的效应量估计值 β_j 以黑点表示，竖线表示 $\pm 2\tilde{s}_{j,\ r}$。图中显示了每对的局部错误发现率，括号内为贝叶斯因子校正前的值。作为比较， β_j 的逆方差加权最小二乘估计以绿色虚线（线性）和紫色虚线（二次）表示。



(a)：FASH 与线性 **(b)：FASH 与二次**

图 S8：FASH-IWP$_1$ 所得（$-\log_{10}$）局部错误发现率与 Strober等人（2019）基于线性检验（a）或二次交互作用检验（b）所得（$-\log_{10}$）p 值的散点图。每个点代表一个独特基因；对于每个基因，局部错误发现率和 p值分别取自局部错误发现率最小和 p值最小的变异-基因对。水平虚线表示对应于错误发现率 $= 0.05$ 的局部错误发现率截断值。颜色表示相对于错误发现率阈值的显著性（灰色：两者均不显著；绿色：仅局部错误发现率显著；橙色：仅 p值显著；紫色：两者均显著）。



(a)：总体 **(b)：放大图**

图 S9：沿变异-基因对排序列表（从最显著到最不显著）的错误发现率，由 FASH-IWP$_1$ 产生的局部错误发现率（局部假发现率）（红色实线），以及将 Gamazon 等（2013）的 eFDR 程序应用于 Strober等人（2019）所报告的 p 值后得到的结果计算（线性检验：绿色点线；二次检验：蓝色虚线）。

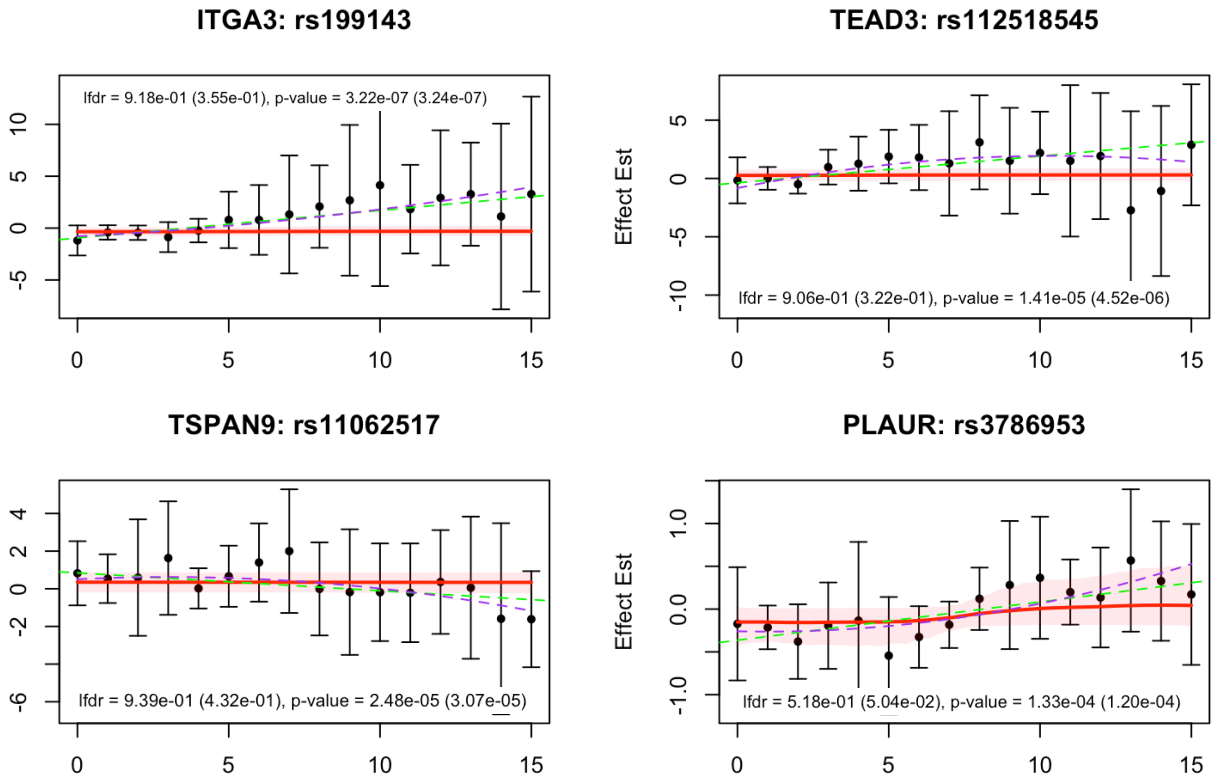


图 S10：Strober等人（2019）识别出但未被 FASH-IWP$_1$检测到的基因示例。对于每个变异-基因对，β_j 的后验均值以红色实线表示，95% 可信区间以红色阴影带表示。观测到的效应ˆ大小估计值 β_j 以黑点表示，竖线表示$\pm 2\tilde{s}_{j,\ r}$。报告了每个变异-基因对的局部错误发现率，并在括号中标出 BF 校正前的数值。Strober等人（2019）报告的（线性/二次）p 值也一并列出，其中二次 p 值显示在括号中。