

DNACopy 分段与 test-v1→test-v2 Liftover 分析报告

科研工作台

1 输入与参数（源自 in/genetic-input.json）

- **分段 (DNACopy CBS)**: 输入 in/bins.csv (180 个 0-based 半开区间 bin, 染色体 chr01, 覆盖 0-18000); mode = log2_ratio; permutations = 1000; seed = 1729; alpha = 0.01; min_width = 2; undo_splits = sdundo (undo_sd = 3)。信号单位为 **synthetic relative coverage**; 参考版本 **synthetic DNA v1**。
- **Liftover**: 输入 in/points.csv (6 个点); chain = in/test.chain (SHA256 facb7476d5b7957efa86c29fb16d52 已实测校验一致); chain 来源声明: *Synthetic exact fixture v1; chrA is UCSC target-field source*; 许可: *Original synthetic fixture*; 源组装 **test-v1**, 目标组装 **test-v2**。该 chain 为原创合成 fixture, 不是 hg19/hg38 链。

2 DNACopy 分段结果（正式运行）

软件: DNACopy 1.72.3 (R 4.2.2 Patched)。CBS 输出 3 个片段:

segment_id	chromosome	start_0based	end_0based	bins	mean_log2_ratio
1	chr01	0	6000	60	7e-04
2	chr01	6000	12000	60	1.0027
3	chr01	12000	18000	60	9e-04

- **坐标系**: 0-based 半开区间 $[start, end)$; CBS 探针位置按 start+1 报告。染色体身份 chr01 全程保留, 未改写。
- **口径声明**: 本分析产出的是 **观测相对信号 (log2 ratio) 的分段**, 不是绝对拷贝数; 未做纯度/倍性/GC/可及性校正、基线居中或临床解释。

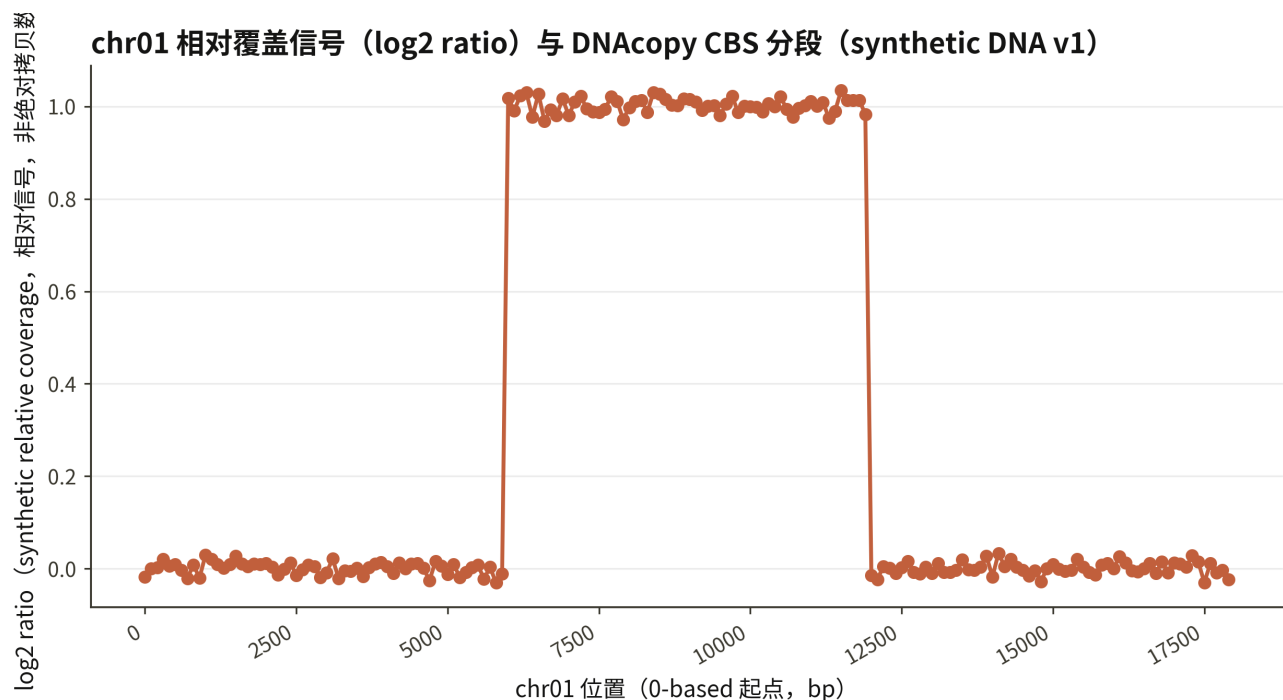


图 1: chr01 逐 bin 相对覆盖信号 (log2 ratio) 与 CBS 分段拟合。纵轴为相对信号，非绝对拷贝数。

3 Liftover 结果 (pyliftover 0.4.1, test-v1 → test-v2)

chain 结构: 2 条链记录、2 个块; 源 contig 1 个 (chrA), 目标 contig 1 个 (chrB)。6 个输入点共产生 8 条输出记录: **unique 1**、**ambiguous 2** (各保留全部匹配, 不自动取舍)、**unmapped 1**、**unknown_contig 1**、**out_of_bounds 1**。坐标系: 两端均为 0-based 点坐标; 链方向在映射中保留/翻转如实记录。

映射表 (out/mapped_points.csv)

point	源染色体	源位置 (0-based)	源链	状态	目标染色体	目标位置 (0-based)	目标链 / score
p1	chrA	110	+	unique	chrB	310	+ / 100
p2	chrA	160	+	ambiguous	chrB	360	+ / 100
p2	chrA	160	+	ambiguous	chrB	1489	- / 90
p6	chrA	160	-	ambiguous	chrB	360	- / 100
p6	chrA	160	-	ambiguous	chrB	1489	+ / 90

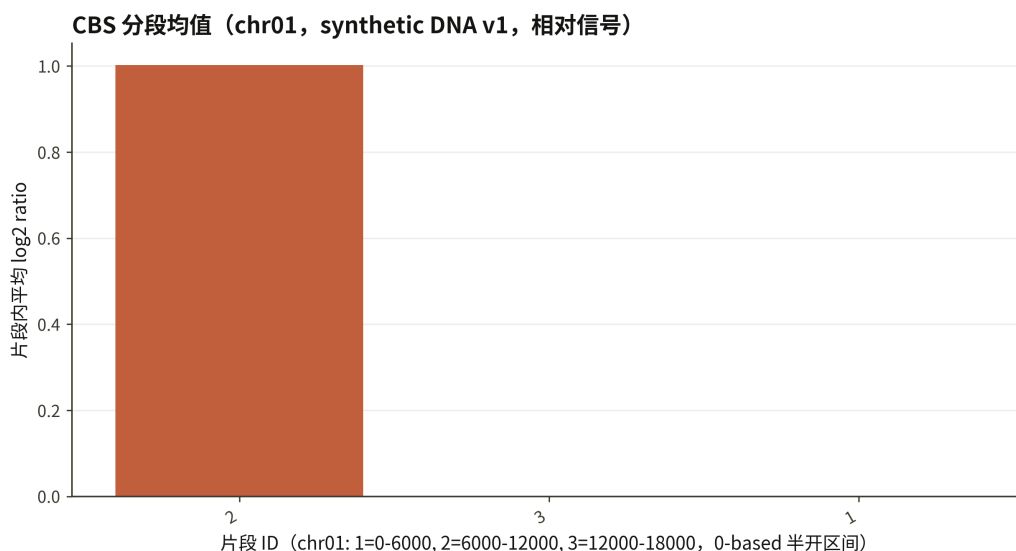


图 2: 三个片段的平均 log2 ratio (相对信号)。

point	染色体	位置 (0-based)	链	状态	未映射原因
p3	chrA	250	+	unmapped	在 chain 的对位块中无匹配 (fixture 中该位点未被任何
p4	missing	100	+	unknown_contig	源染色体名不在 chain 的 source contig 列表中 (chain
p5	chrA	1000	+	out_of_bounds	坐标超出 chain 块在 test-v1 上的覆盖范围

未映射表与原因 (out/unmapped_points.csv)

4 产物清单

- 分段表: out/copy-signal-3943ebb117af/relative_segments.csv; 逐 bin 拟合: per_bin_fitted.csv; 参数: parameters.json; R 过程与诊断: copy_signal.R、R-session-info.txt、DNAcopy_fit.rds; R 原生诊断图: copy_signal.pdf。
- 映射: out/liftover-9f816c744075/point_mappings.csv (全量 8 条); out/mapped_points.csv (映射成功); out/unmapped_points.csv (未映射 + 原因); chain 原件与参数见该目录。
- 图: out/cbs_overlay.png、out/segment_means.png; 本报告: out/analysis-report.pdf。