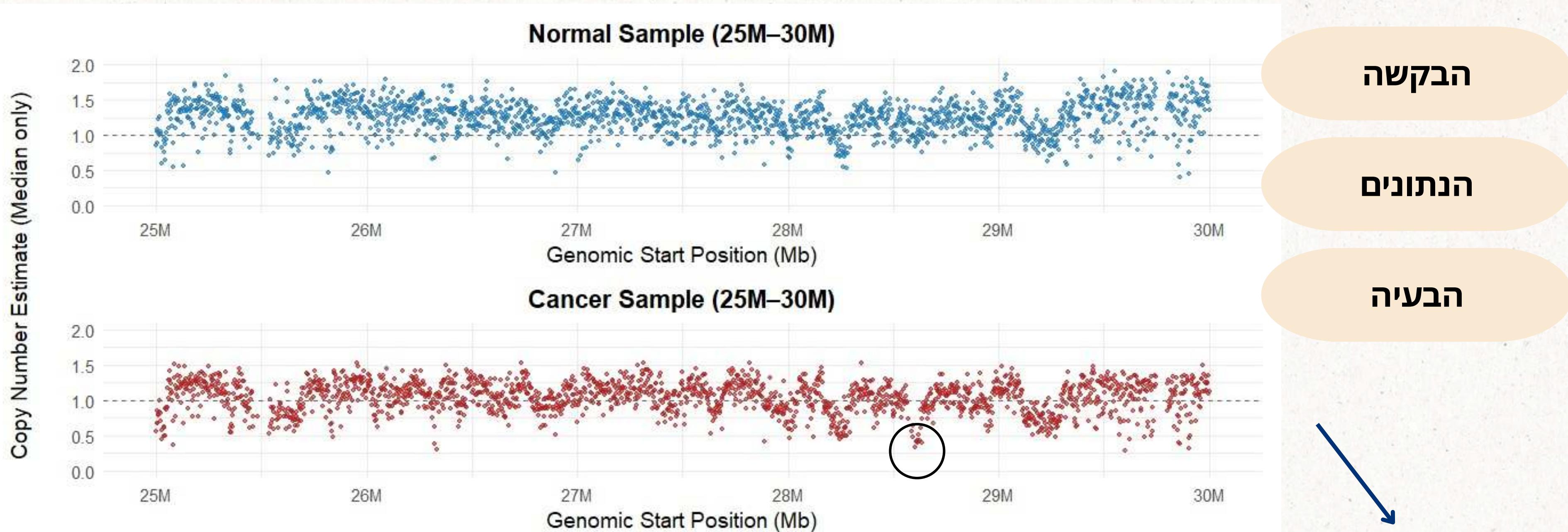


הציגות מסכמת

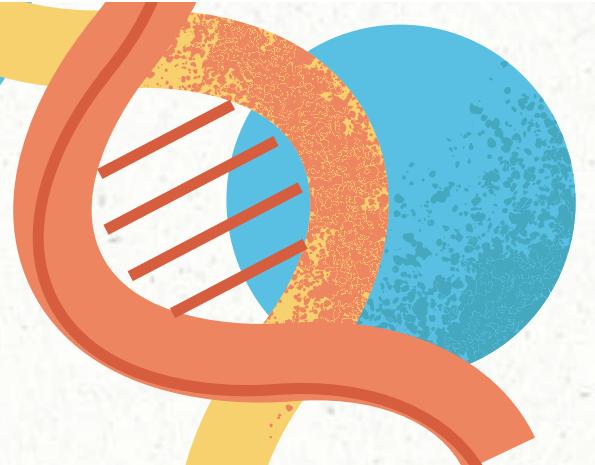
קובוצה B

מציגות: סלומה ברנס, עדן מלכה ועומר אבוחצירה

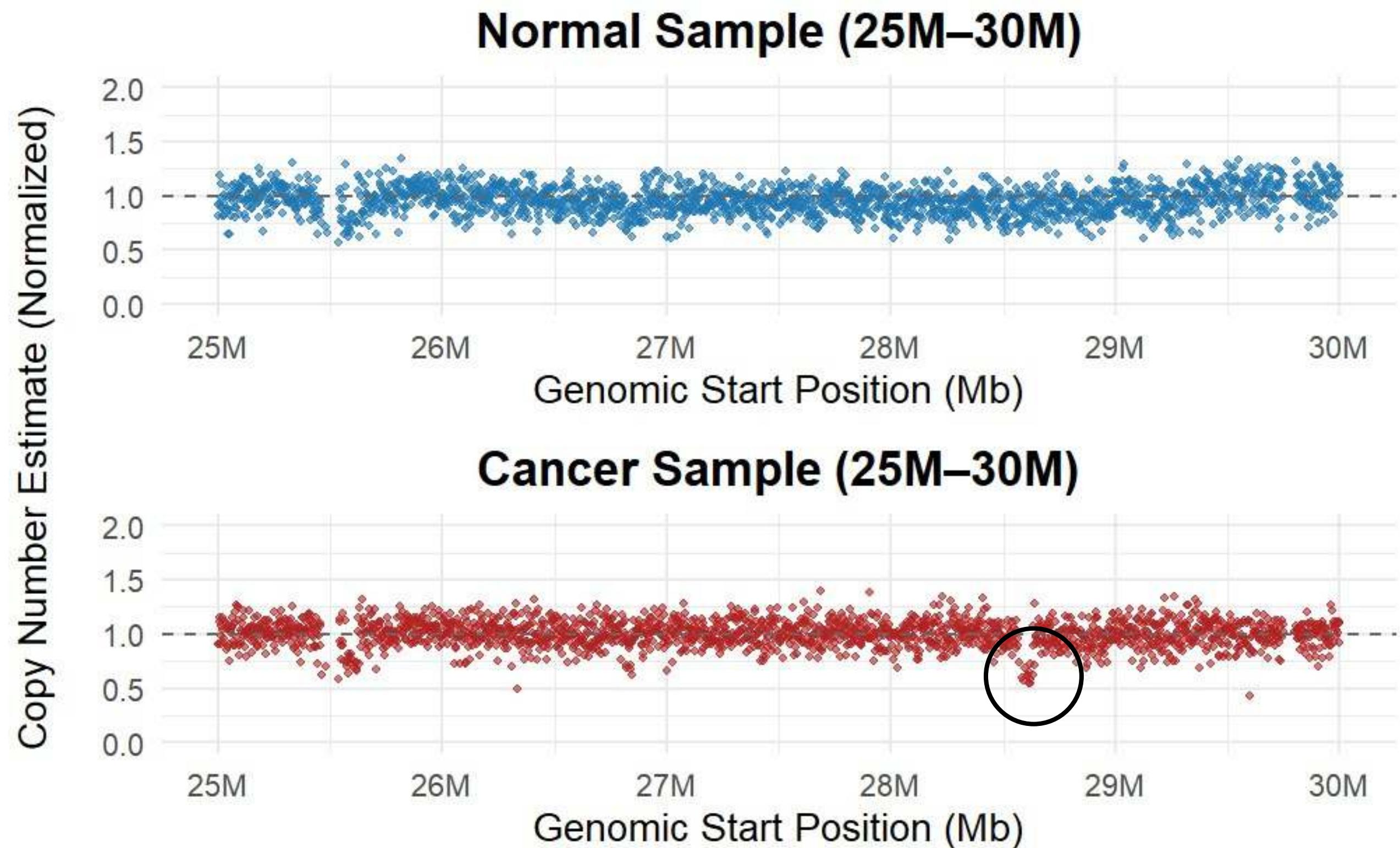
האם יש שינוי במספר העותקים סביר **מ-28.5** בכרומוזום?



כלל תיקון חציון



האם יש שינוי בעותקים סיבי 28.5 בכרומוזום?



הבקשה

הנתונים

הבעיה

ככל תיקון חציו
ותיקון GC

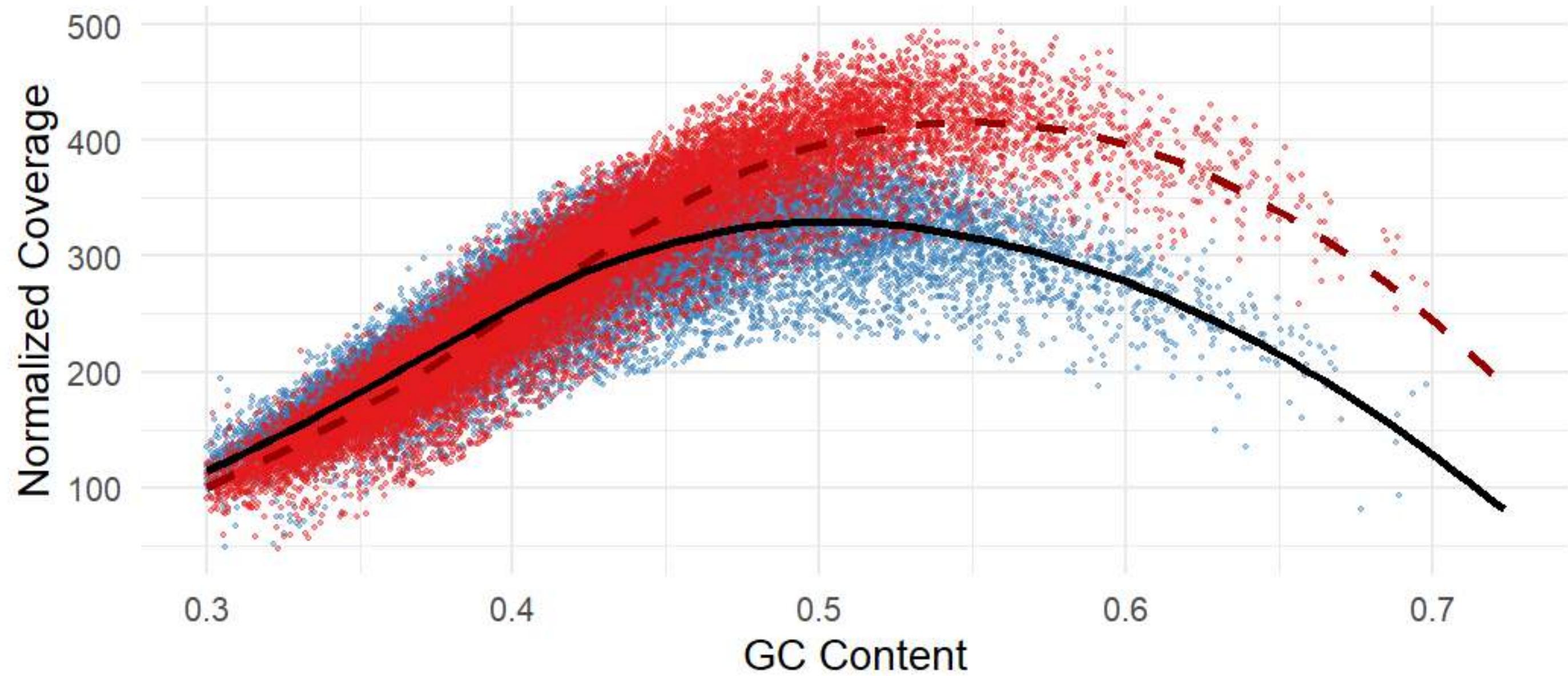
דגימה בריאה למול סרטנית

GC Effect on Coverage (After Median Normalization)

Subset of Points + Full LOESS Curves

50% of the data

Sample • Normal • Tumor



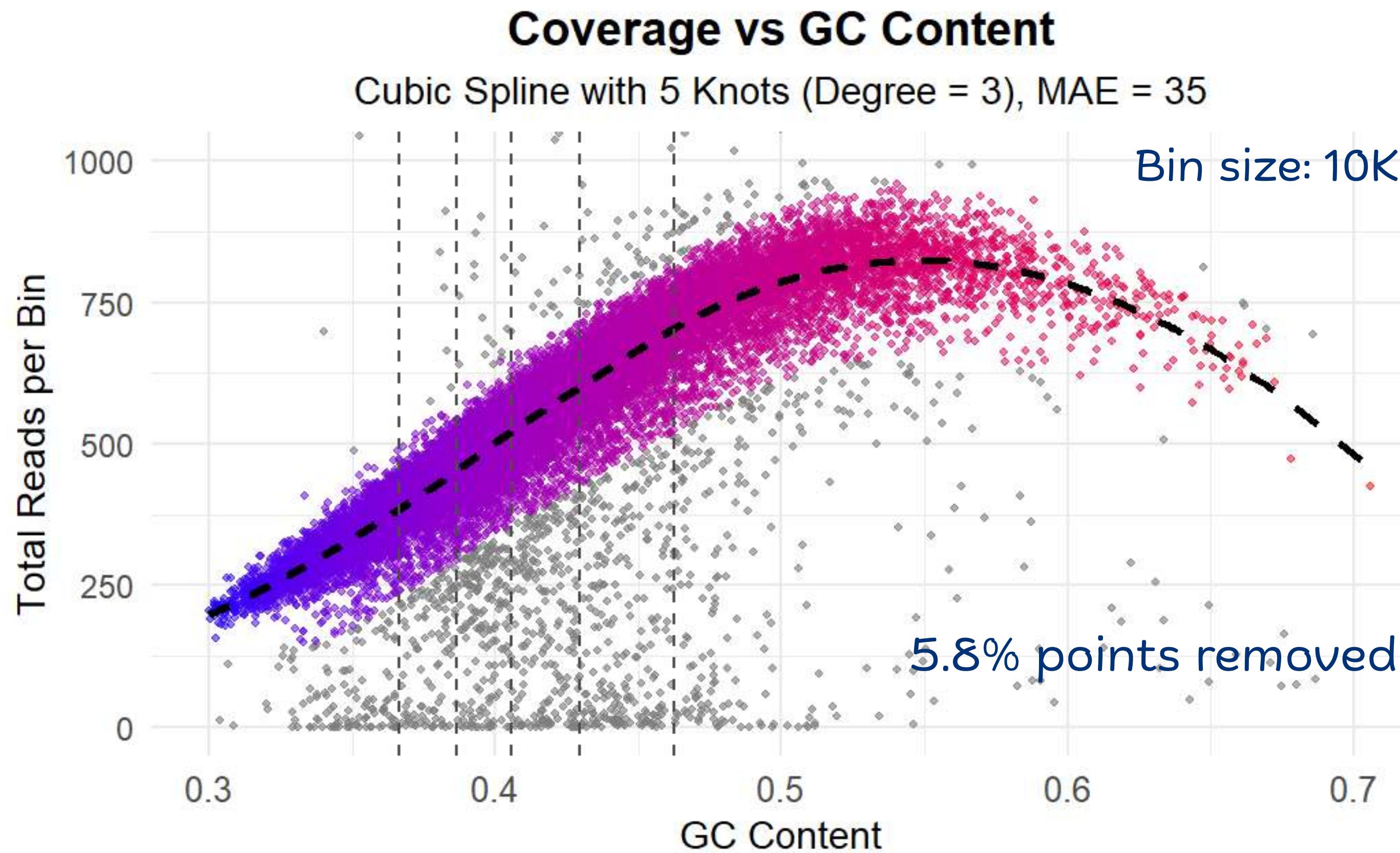
המודל הסטטיסטי באמידת מס' העותקים

$$Y_k^t = a_k \cdot f^t(GC_k) \cdot \gamma_k \cdot \delta_k^t + \eta_k^{t,\text{Pois}}$$

$$Y_k^n = f^n(GC_k) \cdot \gamma_k \cdot \delta_k^n + \eta_k^{n,\text{Pois}}$$

- Y_k^t, Y_k^n : observed read counts in bin k for tumor and normal samples
- $f^t(GC_k), f^n(GC_k)$: GC-bias correction functions (per sample)
- $a_k \in \{0.5, 1, 1.5, 2\}$ is the unknown copy number
- γ_k : shared signal across both samples
- δ_k^t, δ_k^n : sample-specific multiplicative effects
- $\eta_k^{t,\text{Pois}}, \eta_k^{n,\text{Pois}}$: Poisson noise

אודות אפקט ה-GC



Coverage < 5

LOESS \pm 1.5 IQR

LOESS \pm 3.25 IQR

Mean Average Error:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

בחירה אומדן למספר העותקים

Model B

$$\hat{a}_k = \frac{Y_k}{\hat{f}(GC_k)} + C$$

דגם אחד עם תיקון להטיית GC בלבד

Model C

$$\hat{a}_k = \frac{Y_k^T}{Y_k^N} + C$$

יחס קבוע בין דגמה סרטנית לנורמלית: ללא תיקון GC

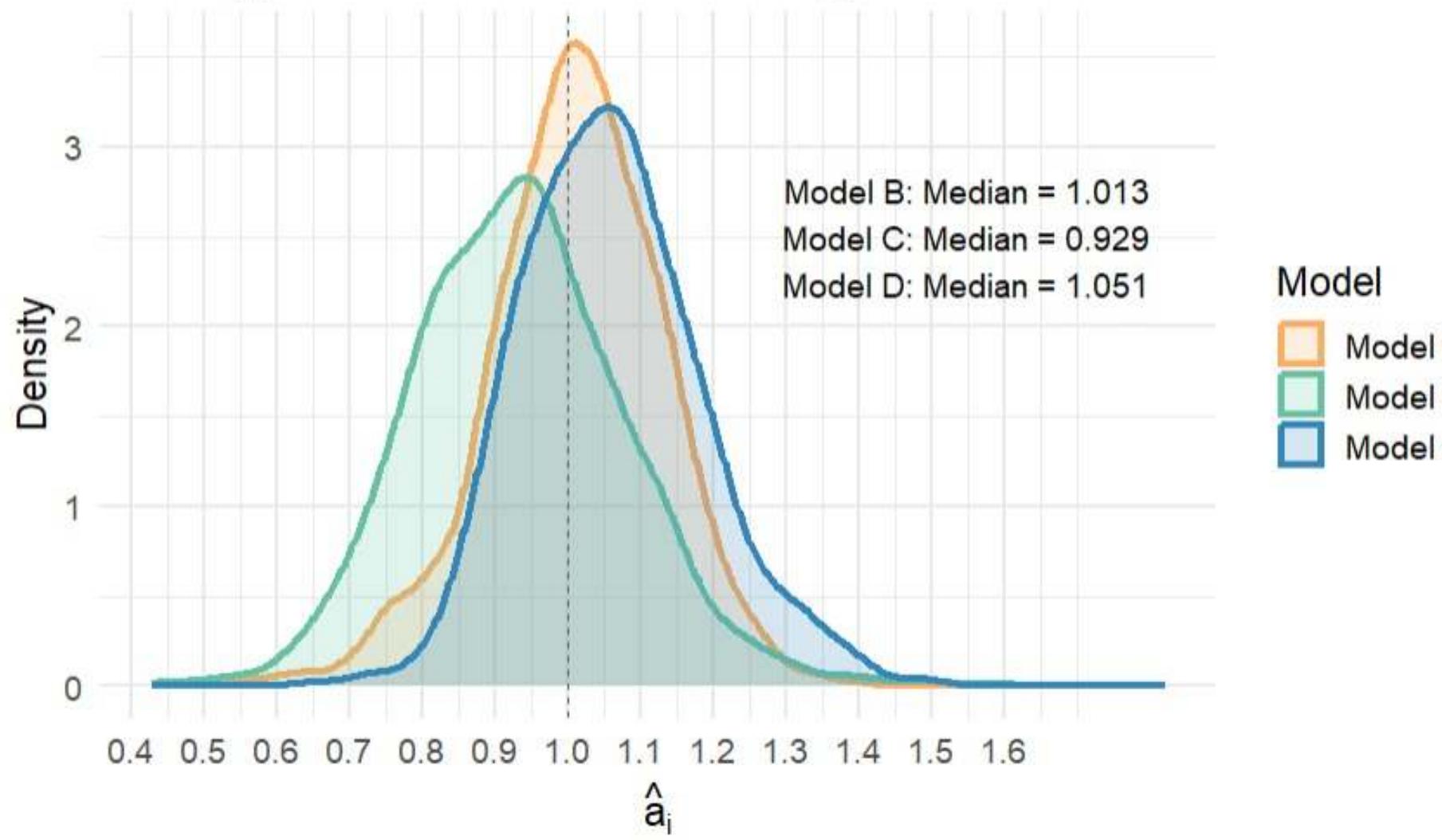
Model D

$$\hat{a}_k = \frac{\left(\frac{Y_k^{(T)}}{\hat{f}_T(GC_k)} + c \right)}{\left(\frac{Y_k^{(N)}}{\hat{f}_N(GC_k)} + c \right)}$$

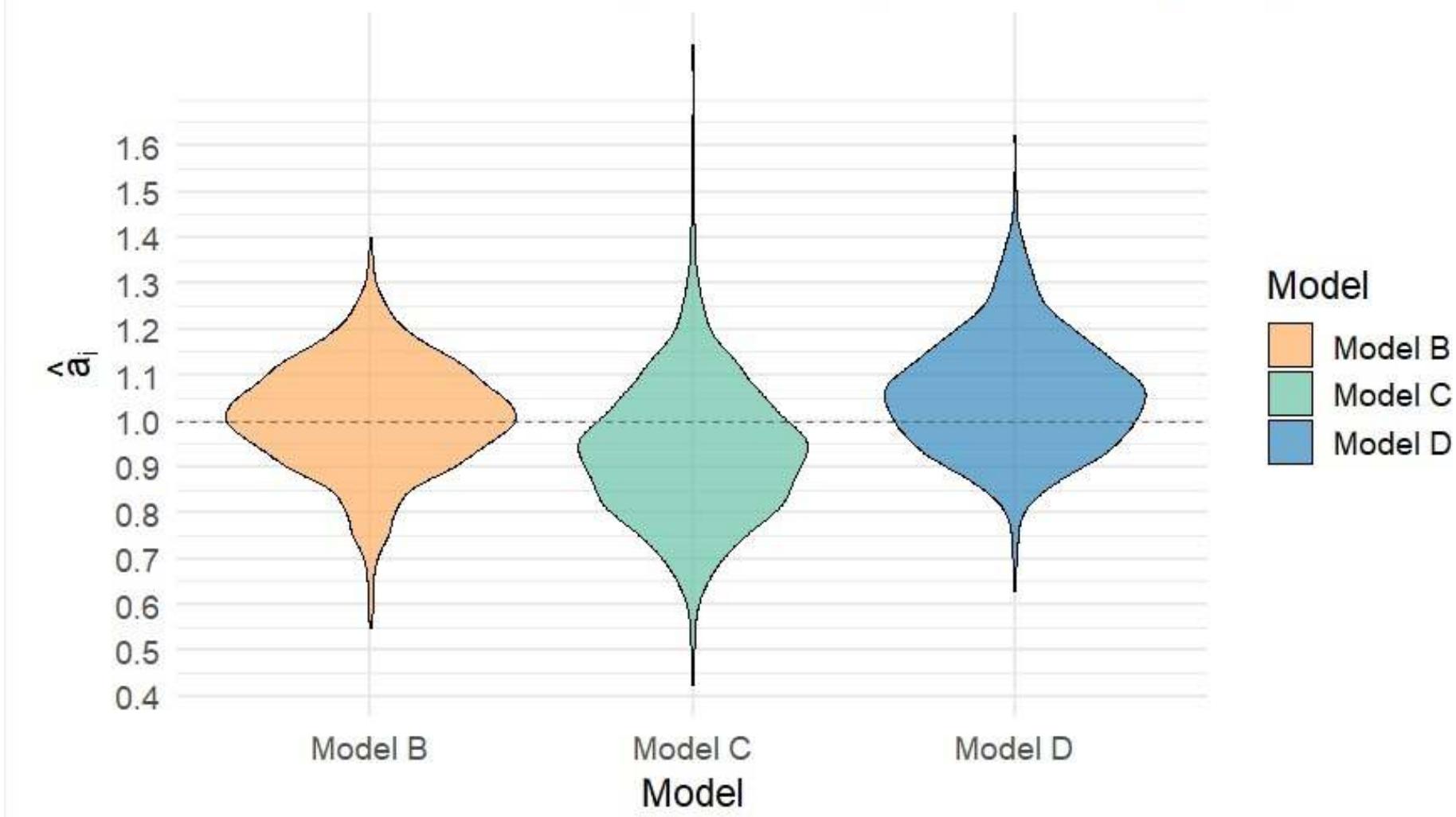
יחס מתוקן בין שתי הדגימות:
עם תיקון להטיית GC +
קבוע יczob $c = 0.1$

מדדי טיב התאמה המודל

Copy Number Estimation – Region 25M–30M



Violin Plot of Estimated \hat{a}_i per Model (Combined Regions)

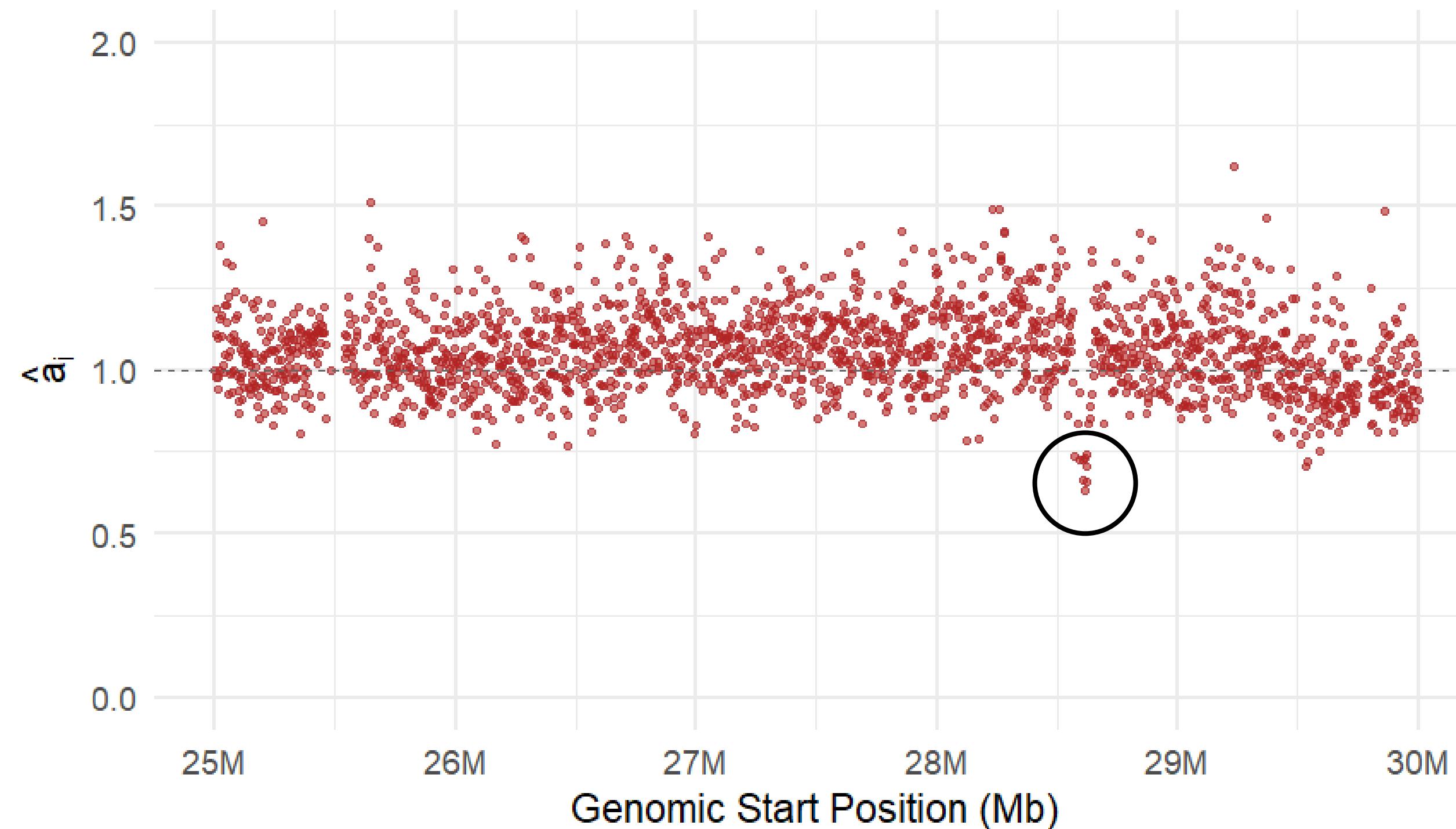


Model	$a^i \in [0.9, 1.1]$
Model B	0.735
Model C	0.482
Model D	0.749

בדיקות פרופורצייה באזוריים
26.5M-27M:

הציגת האזרר המתוקן

Model D: GC-corrected Ratio (25M–30M)



תשובה למתיעץ ומסקנות

נמצאו חריגות ברורות באזור W28.5 בסיסים לאחר תיקון מלא.



אפקטים מקומיים של GC מקשימים על גילוי CNV.



מודל C סיפק את האומדן היציב והמדויק ביותר.

