

AUDIT OBJECTIVE & THE FTO FAILURE MODE PITCH: CRISPR screening and GWAS hits frequently map to non-coding deserts where 1D distance heuristics misassign causal targets. Crucially, a naive computational prior (distance x expression) also fails at complex loci: at FTO, naive priors rank FTO #1 at 92.7% because the target tissue expresses FTO highly (44.5 TPM) and sits 65 kb proximal, missing the true targets (IRX3/IRX5) 0.5 Mb away. Resolving this requires locus-specific measured contacts (4C/Hi-C), which is precisely what TrueLocus Bio's 14-day audit adds over a prior.

BENCHMARK GROUND TRUTH: On 72 held-out CRISPRi enhancer-gene pairs (K562, Schraivogel 2020), TrueLocus achieves **69.4% Top-1 (50/72)** and **84.7% Top-3 accuracy** vs 54.2% nearest-gene (p=0.007, McNemar exact), requiring zero cell-type Hi-C input. All coordinates anchored to Ensembl. Authentic CAGE TPMs extracted from FANTOM5.

TABLE 1: 8q24.21 Oncology Desert Screen Audit (Fulco et al. 2016 Science e1-e4 + rs6983267 Enhancer)

Context: Core 8q24.21 cluster in K562. 1D nearest-TSS assigns inactive lncRNAs (CASC8, CCAT2). TrueLocus tracks active polymer contact to MYC (496.1 TPM; 244-334 kb distal).

Hit ID	Genomic Coord (GRCh38)	1D Nearest TSS	1D Dist	TrueLocus #1 Target	Loop Dist	P_contact	Share	Audit De-Risking Action & Rationale
CR-8q-e1	chr8:127465656-655	CASC8 (lncRNA)	15.5 kb	MYC (Oncogene)	268.8 kb	0.0183	73.3%	[PROMOTE MYC / DEMOTE CASC8] CASC8 silenced in K562. Active promoter contact drives MYC (496.1 TPM; Fulco 2016 Science).
CR-8q-e2	chr8:127471556-055	CASC8 (lncRNA)	9.8 kb	MYC (Oncogene)	263.1 kb	0.0186	73.4%	[PROMOTE MYC / DEMOTE CASC8] Bypasses non-expressed CASC8 body; loops directly to active MYC core.
CR-8q-e3	chr8:127482856-855	CASC8 (lncRNA)	1.7 kb	MYC (Oncogene)	251.6 kb	0.0195	73.5%	[PROMOTE MYC / DEMOTE CASC8] 1.7 kb proximal TSS is epigenetically inert; 3D contact drives MYC.
CR-8q-e4	chr8:127490056-055	CASC8 (lncRNA)	8.9 kb	MYC (Oncogene)	244.4 kb	0.0200	73.6%	[PROMOTE MYC / DEMOTE CASC8] CRISPRi titration validates selective MYC loss (Fulco 2016); CASC8 unaffected.
CR-8q-rs698	chr8:127401060 (SNP)	CCAT2 (lncRNA)	661 bp	MYC (Oncogene)	334.4 kb	0.0147	72.6%	[PROMOTE MYC / DEMOTE CCAT2] Sits in CCAT2 body; CCAT2 promoter silenced in K562. 3D loop drives MYC (Pomerantz 2009 3C).

TABLE 2: Multi-Promoter Kinetic Competition & Boundary Insulation (GWAS & 3C/4C Benchmarks)

Context: Distance x expression priors succeed when competing promoters are closed (8q24) but fail when proximal bystanders are active (FTO 92.7%). Resolving FTO and 9p21 near-ties requires locus-specific 3C/4C conformation.

Locus & Hit	Genomic Coord	1D Nearest TSS	1D Dist	TrueLocus Rank #1 Target	Target Dist	Comp. Share	Tug-of-War Mechanism & Experimental Validation
FTO / IRX3 Obesity / Browning	chr16:53767042 (rs1421085)	FTO (Intron 1)	65.4 kb	FTO (Naive Prior) IRX3 (4C True Target)	65.4 kb / 520.4 kb	FTO: 92.7% (IRX3: 7.3%)	[FAILED NAIVE PRIOR : OVERRIDDEN BY 4C] Naive prior ranks FTO #1 (92.7%) because FTO is active (44.5 TPM) and 65 kb proximal. True causal targets (IRX3/5) are 0.5-1.2 Mb distal. Claussnitzer 2015 NEJM 4C proves physical loop bypasses FTO to IRX3/IRX5.
9p21.3 / CAD Coronary Disease	chr9:22124478 (rs10757278)	CDKN2B (TSS) (in ANRIL body)	115.2 kb (130 kb ANRIL)	CDKN2A / CDKN2B (Near-Tie Ambiguity)	129.2 kb / 115.2 kb	CDKN2A: 39.9% CDKN2B: 36.5% (MTAP: 23.7%)	[AMBIGUOUS : RESOLVED BY 3C] HUVEC CAGE reveals active near-tie: CDKN2A (8.8 TPM, 39.9%) vs CDKN2B (1.6 TPM, 36.5%). Harismendy 2011 Nature 3C proves physical looping to both promoters; IFNA cluster insulated.
TCF7L2 / T2D Type 2 Diabetes	chr10:112998590 (rs7903146)	TCF7L2 (Intron 4)	48.6 kb	TCF7L2 (Pancreas TF)	48.6 kb	100.0% (VT11A: 0%)	[CONCORDANT / INSULATED] Context: Pancreas (FANTOM5 whole-tissue, CNhs11756, 61.7 TPM). 3C confirms intronic loop to TCF7L2 (Gaulton 2010 Nat Genet / Greenwald 2019 Nat Commun). VT11A (551 kb) is insulated across CTCF boundary.

AUDIT BREAKDOWN: Full Candidate Triage Matrices (8q24.21 vs 9p21.3 CAD Competition Fields)

Candidate field deconvolution: 8q24.21 filters inactive lncRNAs to promote MYC (73.3%). 9p21.3 quantifies an authentic near-tie (CDKN2A 39.9% vs CDKN2B 36.5%, MTAP 23.7%), demonstrating why physical conformation assays are required.

Rank	Gene (8q)	Dist	CAGE	Score	Share	Triage Status	Rank	Gene (9p)	Dist	CAGE	Score	Share	Triage Status
#1	MYC	268.8k	496.1	0.0494	73.3%	TARGET (Open/Loop)	#1	CDKN2A	129.2k	8.84	0.0466	39.9%	CO-TARGET (3C active)
#2	PVT1	327.9k	6.8	0.0180	26.7%	DECOY (Active prom.)	#2	CDKN2B	115.2k	1.61	0.0427	36.5%	CO-TARGET (Primary 3C)
#3	CASC8	15.5k	0.00	0.0012	0.0%	FILTERED (Closed TSS)	#3	MTAP	321.9k	55.72	0.0277	23.7%	CO-TARGET (3C loop)
#4	CCAT2	66.3k	0.00	0.0004	0.0%	FILTERED (Closed TSS)	#4	ANRIL	130.3k	0.00	0.0002	0.0%	FILTERED (Closed TSS)
#5	POU5F1B	144.5k	0.00	0.0002	0.0%	FILTERED (Closed prom.)	#5	IFNA1	684.0k	0.00	0.0000	0.0%	FILTERED (Insulated)
#6	LRATD2	908.2k	0.00	0.0000	0.0%	FILTERED (Insulated)	-	Domain	21.7-22M	HUVEC	Sum	100%	Harismendy 2011 3C

SCORING FORMULA & PROVENANCE: In Silico ABC formulation: Score = P_contact(d) * F_TAD * A_promoter. Polymer contact prior: P_contact = [5000 / (dist_bp + 5000)]^gamma (gamma=1.0 inside TAD, 1.75 across boundary). Boundary factor: F_TAD = 1.0 (in convergent CTCF loop), 0.08 (insulated across boundary). Promoter competence: A_promoter = log10(CAGE_TPM + 9) if active (CAGE >= 1 TPM), else 0.005 (silenced). Shares assume a mutually-exclusive contact model; measured contacts (3C/4C) discriminate competition from sharing. CAGE TPMs extracted from authentic FANTOM5 phase 1/2 (CNhs12334, CNhs10872, CNhs12065, CNhs11756).

14-DAY PILOT AUDIT SPECIFICATION: 20 to 50 candidate screening hits audited per engagement. Deliverables include: (1) Target De-Risking TSV with full candidate rankings, (2) Promoter Competition & Decoy Matrix, (3) Vector-mapped CTCF boundary profiles. Turnaround: 14 Business Days. Contact: contact@truelocus.bio | truelocus.bio