ELECTRONIC GENOME MAPPING IN ASSESSMENT OF PROCESS-INDUCED GENETIC ALTERATIONS IN HUMAN T CELL-DERIVED iPSCs

Ruud Hulspas, Carolina Sasso, Kristin Cabral, Carolyn Lutzko, Jose Cancelas, Catherine Brownstein, Jerome Ritz
Connell & O'Reilly Families Cell Manipulation Core Facility
Dana Farber Cancer Institute, Harvard Medical School, Children's Hospital, Boston, Massachusetts, USA

INTRODUCTION

Recently published OncoPanel quality control results from our T cell-derived iPSC manufacturing process revealed a recurrent deletion of the cancer-associated gene serine protease 1 gene (PRSS1) with loss of at least one allele across all T cell-derived iPSC lines. The PRSS1 deletion was also detected in pre-programming controls and mature T cells from other donors, indicating this gene loss is intrinsic to T cell differentiation and not related to the iPSC manufacturing process. (Hulspas R et al. Cytotherapy.2026 Apr;28(4)). Given the role of PRSS1 in enzymatic digestion of food, T-cell derived iPSC would not be preferred iPSC lines to generate exocrine pancreatic cells or tissues.

OncoPanel is a next-generation sequencing (NGS) assay that surveys exonic DNA sequences of 447 cancer genes and 191 regions across 60 genes for rearrangement detection. We used the same set of genes in assessing structural variants (SVs) by means of Electronic Genome Mapping (EGM). EGM provides an approach complementary to NGS by analyzing intact DNA molecules exceeding 100 kb to generate genome-wide structural maps, enabling detection of copy number changes and structural variants down to ~300 bp.

In this initial study, we established that the heterozygous as well as homozygous loss of PRSS1 found in T cells are also identified by EGM as large deletions. Furthermore, 3 iPSC lines were used to explore the presence of additional SVs.

METHOD

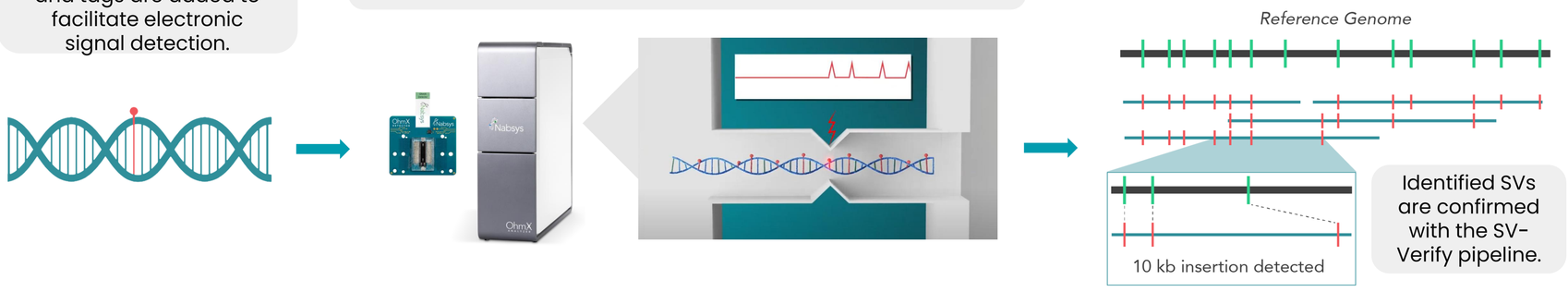
Samples: DNA extracts from T cell-derived iPSC lines E, I and Q as well as pre-reprogrammed T cells (control).

Donor 1	Pluripotency (FCM)	T cell clonality	G Banding Karyotyping	OncoPanel
line	OCT4/TRA-1-60	productive	karyotype	variant
E	94%	100%	normal male	loss PRSS1
I	97%	47%	normal male	2 copy del PRSS1
Q	96%	100%	normal male	2 copy del PRSS1

Sequence-specific nickases create nicks ~4 kb apart genome-wide, and tags are added to facilitate electronic signal detection.

As tagged DNA passes through the detector, they create a signal. The time between tags is converted to distance in base pairs to create sequence-specific molecules or maps.

Read maps are assembled *de novo* and aligned to a human reference genome to discover SVs.



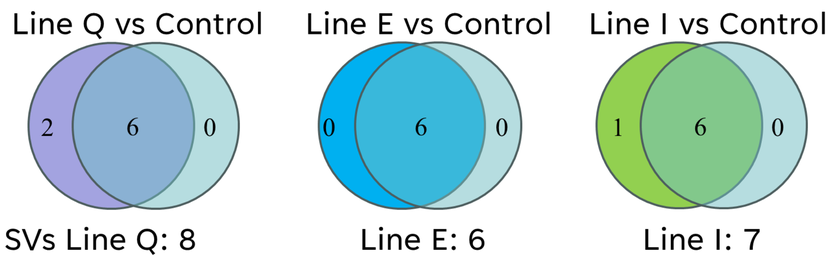
447 cancer genes screened in OncoPanel

ABCB11	CASP8	DKC1	FGFR3	KAT6B	MYBL1	POLD1	RNF8	TCF3
ABL1	CBFA2T3	DMC1	FGFR4	KCNQ1	MYC	POLE	ROS1	TCF7L2
ACVR1	CBFB	DNMT3A	FH	KDM5A	MYCL1	POLH	RPA1	TDG
AKT1	CBL	DOCK8	FLCN	KDM5C	MYCN	POLQ	RPTOR	TERC
AKT2	CBLB	EGFR	FLT1	KDM6A	MYD88	POT1	RSPO2	TERT
AKT3	CCND1	EGLN1	FLT3	KDR	NBN	PPARG	RSPO3	TET1
ALK	CCND2	ELANE	FLT4	KEAP1	NEIL1	PPM1D	RUNX1	TET2
APC	CCND3	EME1	FOXA1	KIF1B	NEIL2	PPP2R1A	RUNX1T1	TFE3
AR	CCNE1	ENG	FOXL2	KIT	NEIL3	PRDM1	SBDS	TLX3
ARAF	CD274	EP300	FUS	KLF2	NF1	PRF1	SDHA	TMEM127
ARHGAP35	CD79B	EPCAM	GALNT12	KLF4	NF2	PRKAR1A	SDHAF2	TMPRSS2
ARHGEF12	CDC73	ERBB2	GATA2	KLLN	NFE2L2	PRKCI	SDHB	TNFAIP3
ARID1A	CDH1	ERBB3	GATA3	KMT2A	NFKBIA	PRKDC	SDHC	TOPBP1
ARID1B	CDH4	ERBB4	GATA4	KMT2D	NFKBIE	PRSS1	SDHD	TP53
ARID2	CDK12	ERCC1	GATA6	KRAS	NFKBIZ	PTCH1	SERPINA1	TP53BP1
ASXL1	CDK4	ERCC2	GBA	LIG4	NKX2-1	PTEN	SETBP1	TRAF3
ATM	CDK6	ERCC3	GEN1	LMO1	NKX3-1	PTK2B	SETD2	TRAF7
ATR	CDK8	ERCC4	GLI1	LMO2	NOTCH1	PTPN11	SF3B1	TRIM37
ATRX	CDKN1A	ERCC5	GLI2	MAF	NOTCH2	PTPN14	SH2B3	TSC1
AURKA	CDKN1B	ERCC6	GNA11	MAFB	NOTCH3	PVRL4	SH2D1A	TSC2
AURKB	CDKN1C	ERG	GNAQ	MAP2K1	NPM1	QKI	SLC25A13	TSHR
AXIN2	CDKN2A	ESR1	GNAS	MAP2K2	NR0B1	RAC1	SLC34A2	U2AF1
AXL	CDKN2B	ETV1	GPC3	MAP2K4	NRAS	RAD21	SLX1A	UBE2T
B2M	CDKN2C	ETV4	GREM1	MAP3K1	NRG1	RAD50	SLX1B	UIMC1
BABAM1	CEBPA	ETV5	H19	MAPK1	NSD1	RAD51	SLX4	UROD
BAP1	CHEK1	ETV6	H3F3A	MAX	NT5C2	RAD51C	SMAD2	USP28
BARD1	CHEK2	EWSR1	H3F3B	MBD4	NTHL1	RAD51D	SMAD4	USP8
BCL11B	CIC	EXO1	HABP2	MCL1	NTRK1	RAD52	SMARCA4	VEGFA
BCL2	CIITA	EXT1	HELQ	MCM8	NTRK2	RAD54B	SMARCB1	VHL
BCL2L1	COL7A1	EXT2	HFE	MDM2	NTRK3	RAF1	SMARCE1	WAS
BCL2L12	CREBBP	EZH2	HIST1H3B	MDM4	OGG1	RARA	SMC3	WHSC1
BCL6	CRKL	FAH	HIST1H3C	MECOM	PALB2	RASA1	SMO	WHSC1L1
BCOR	CRLF2	FAM175A	HMBS	MED12	PARK2	RB1	SOCS1	WRN
BCORL1	CRTC1	FAM46C	HNF1A	MEF2B	PAX5	RBBP8	SOS1	WT1
BLM	CSF3R	FAN1	HOXB13	MEN1	PAXIP1	RBM10	SOX2	XPA
BMPR1A	CTCF	FANCA	HRAS	MET	PBRM1	RECQL4	SOX9	XPC
BRAF	CTLA4	FANCB	ID3	MGA	PDCD1LG2	REL	SPOP	XPO1
BRCA1	CTNNA1	FANCC	ID4	MITF	PDGFRA	RELA	SRSF2	XRCC1
BRCA2	CTNNB1	FANCD2	IDH1	MLH1	PDGFRB	RET	SRY	XRCC2
BRCC3	CUX1	FANCE	IDH2	MLH3	PHF6	RHBDF2	SS18	XRCC3
BRD3	CXCR4	FANCF	IGF1R	MPL	PHOX2B	RHEB	STAG2	XRCC4
BRD4	CYLD	FANCG	IGF2	MRE11A	PIK3C2B	RHOA	STAT3	XRCC5
BRE	DAXX	FANCI	IKZF1	MSH2	PIK3CA	RHOH	STAT6	XRCC6
BRIP1	DCLRE1C	FANCL	IL7R	MSH6	PIK3R1	RHOT1	STK11	YAP1
BUB1B	DDB1	FANCM	ITK	MTA1	PIM1	RICTOR	SUFU	ZNF217
C17ORF70	DDB2	FAS	JAK1	MTAP	PML	RIF1	SUZ12	ZNRF3
C19ORF40	DDR2	FAT1	JAK2	MTOR	PMS1	RINT1	TAL1	ZRSR2
C1ORF86	DICER1	FBXW7	JAK3	MUS81	PMS2	RIT1	TAL2	
CALR	DIS3	FGFR1	JAZF1	MUTYH	PNKP	RMRP	TAZ	
CARD11	DIS3L2	FGFR2	KAT6A	MYB	POLB	RNF43	TCEB1	

After *de novo* assembly, iPSC line samples as well as a pre-reprogrammed T cell control sample were aligned to a human reference genome (hg38). Subsequently identified SV's in the iPSC lines were compared to SV's identified in the control.

RESULTS

Figure 1. Venn diagrams showing differences and similarities in SV's between pre-reprogrammed T cells (control) and iPSC lines, Q, E and I.



total SVs Line Q: 8 Line E: 6 Line I: 7

Gene	Location	Genic Status	Gene Region	Variant Type	Zygosity	iPSC line	Similar variants in control data bases	Notes
PRSS1	7q34	Genic	exonic	DEL	HOM, HET	In all T cells and T cell-derived lines	yes	published: R Hulspas et al. Development of a practical GMP-compliant manufacturing process for T cell-derived iPSCs. Cytotherapy. 2026 Apr;28(4):102036.
ARHGAP30	1q23.3	Genic	exonic*	INS	HOM	In all T cells, lines E, I, Q	yes	Downstream of NECTINA4 (oncopanel target). *) putatively exonic - label interval overlaps exons
POLE	12q24.33	Genic	exonic*	INS	HOM	In all T cells, lines E, I, Q	no	*) putatively exonic - label interval overlaps exons
DNMT3A	2p23.3	Genic	intronic	INS	HOM	In all T cells, lines E, I, Q	yes	
FGFR2	10q26.13	Genic	intronic	INS	HOM	In all T cells, lines E, I, Q	no	
PRKN	6q26	Genic	intronic	DEL	UNK	In all T cells, lines E, I, Q	yes	
CCND1	11q13.3	Intergenic	-	INS	HOM	Q, I	no	
MYCL	1p34.2	Intergenic	-	INS	HET	Q	no	

Table 1. Select SVs detected by EGM overlapping or near onco-panel genes

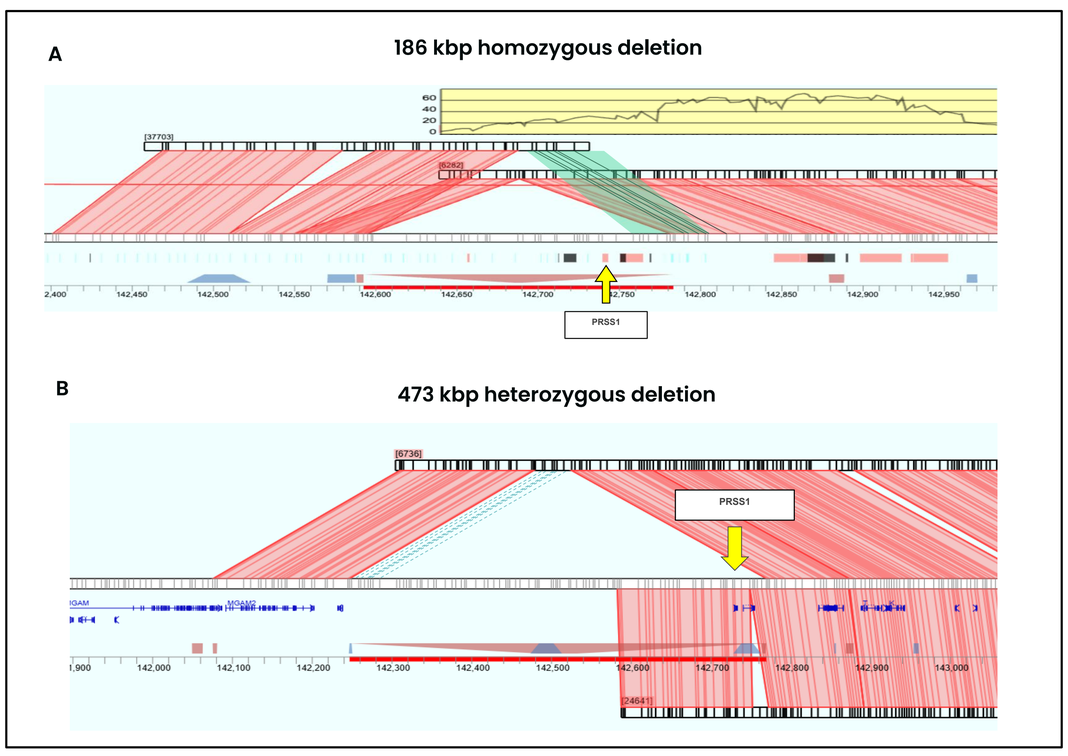


Figure 2. Orthogonal confirmation of PRSS1 deletions by EGM: A - 186kb homozygous deletion, B – 473kb heterozygous deletion

CONCLUSIONS

In this initial study, we established that the heterozygous as well as homozygous loss of PRSS1 found in T cells are also identified by EGM as large deletions (respectively, 473 kbp and 186 kbp). While we currently present orthogonal confirmation and greater resolution on SVs detected by NGS, ongoing efforts are aimed at both gene-specific and whole-genome curation employing a growing suite of tools offered for EGM data. Our early findings indicate that EGM reveals differences in the structural variant landscape between starting material and derived iPSCs that are not readily captured by sequencing alone, highlighting the value of combining sequence-level and long-range structural information for comprehensive genomic characterization.