Stem Cell Reviews and Reports
https://doi.org/10.1007/s12015-026-11091-y

RESEARCH

Gene Expression at the Pluripotency Stage Predicts Pancreatic Endocrine Differentiation in iPSC Clones

Valentina Zamarian[1] · Laura Monaco[1,2] · Manuela Marras[1] · Chiara Ceriani[1,2] · Silvia Pellegrini[1] · Lorenzo Piemonti[1,2] · Valeria Sordi[1]

Received: 15 January 2026 / Accepted: 16 February 2026

Abstract

Background Induced pluripotent stem cell (iPSC)-derived β-like cells hold great promise for cell replacement therapy in type 1 diabetes. However, the reprogramming process generates iPSC clones with variable differentiation capacity, hindering the selection of optimal cell lines. This study aimed to identify an early-stage transcriptional signature capable of predicting the β cell differentiation potential of donor-matched iPSC clones.

Methods Eleven iPSC clones derived from a single donor were differentiated to the definitive endoderm (DE) stage; six were further driven toward pancreatic progenitors (PP) and insulin-producing cells. Differentiation efficiency was evaluated by flow cytometry and qPCR at iPSC, DE, PP, and β cell stages. At the pluripotent stage, expression profiling of 770 genes related to pluripotency and trilineage specification was performed to identify predictive molecular markers.

Results Transcriptomic analysis segregated the clones into two groups (Gr1 and Gr2) with significantly different differentiation outcomes. Gr2 clones exhibited superior DE efficiency ($Cxcr4^+$: 90.1 ± 5.6% vs. 79.8 ± 3.6%; $P = 0.027$) and higher expression of PP markers ($Pdx1^+$, $Nkx6.1^+$, and double-positive cells; $P \leq 0.05$). At the β cell stage, Gr2 clones showed increased frequencies of $Pdx1^+/Ins^+$ and $Nkx6.1^+/Ins^+$ cells ($P \leq 0.05$), along with enhanced glucose-stimulated insulin secretion. A set of 73 differentially expressed genes, enriched in pathways related to naïve/primed pluripotency, endoderm commitment, and metabolism, was identified. From this, a ten-gene signature validated by qPCR strongly correlated with pancreatic marker expression at all stages.

Conclusions An early gene expression signature at the pluripotent stage predicts the pancreatic endocrine differentiation potential of iPSC clones. This molecular screening approach may enable rapid preselection of high-performing clones, thereby accelerating the development of personalized stem cell–based therapies for diabetes.

Summary Cellular reprogramming is a fundamental tool in regenerative medicine but often produces iPSC clones with heterogeneous differentiation potential. Identifying the most suitable clones typically requires time-consuming assays and prolonged in vitro testing. This study presents a streamlined transcriptomic approach to predict, at the pluripotent stage, the differentiation efficiency of iPSC clones into pancreatic endoderm and insulin-producing cells, enabling early selection of high-performing lines for the development of diabetes cell therapy.

Keywords Induced pluripotent stem cells (iPSCs) · β cell differentiation · Gene expression signature · Pluripotency · Type 1 diabetes · Lineage commitment

Introduction

Diabetes is characterized by insufficient pancreatic insulin production, requiring innovative therapeutic strategies to restore physiological glycaemic control. The differentiation of pluripotent stem cells into insulin-producing cells offers a potential curative approach by replacing damaged or dysfunctional pancreatic tissue. Among pluripotent stem cell sources, induced pluripotent stem cells (iPSCs),

✉ Valeria Sordi
sordi.valeria@hsr.it

[1] Diabetes Research Institute, IRCCS San Raffaele Scientific Institute, Milan, Italy

[2] Università Vita-Salute San Raffaele, Milan, Italy

Published online: 04 March 2026

reprogrammed from adult somatic cells, represent a major advancement in personalized regenerative medicine [1, 2]. Notably, patient-derived iPSCs enable autologous β cell differentiation strategies relevant to type 1 diabetes -like [3–6].

Despite their transformative potential, iPSC reprogramming typically generates numerous clones with variable biological behaviours. Individual clones may differ in levels of pluripotency and lineage-specific differentiation capacity, requiring rigorous molecular and functional assessment [7, 8]. This variability is particularly impactful in β cell differentiation, where even donor-matched iPSC clones often exhibit divergent efficiencies and functional outcomes. Previous studies have attributed such heterogeneity to residual epigenetic memory, incomplete reprogramming, or intrinsic differences in transcriptional and metabolic states [9–11]. Consequently, the yield and functional quality of iPSC-derived β-like cells, including glucose-responsive insulin secretion, can vary considerably even among genetically identical clones [12].

Traditionally, identifying the most suitable iPSC clone requires prolonged culture, clonal expansion, and multiple rounds of phenotypic and functional screening. These procedures are labour-intensive, costly, and often span several months [13, 14], underscoring the need for strategies that allow earlier prediction of differentiation competence.

Transcriptomic profiling offers a promising solution to assess intrinsic differentiation potential directly at the pluripotent stage. Although iPSCs derived from the same donor are genetically matched, they often exhibit considerable transcriptional heterogeneity [10, 15], which may influence lineage commitment and functional maturation.

In this study, we propose a transcriptomic strategy for early identification of iPSC clones with high pancreatic endocrine differentiation potential. By comparing gene expression profiles across multiple donor-matched iPSC clones, we aimed to define early-stage molecular markers that predict downstream β cell differentiation outcomes. This approach may enable the preselection of high-performing iPSC clones and accelerate the development of patient-specific stem cell therapies for diabetes.

Materials and Methods

IPSC Reprogramming and Characterization

A total of eleven human iPSC clones (designated DRI1 clones #1, #3, #4, #5, #7, #9, #10, #11, #14, #16 and #17) were generated by reprogramming of $CD34^{+}$-enriched peripheral blood cells from a healthy donor using Sendai virus-based technology (CytoTune-iPS Sendai Reprogramming Kit, ThermoFisher, Waltham, MA, USA). Human biological samples were collected after obtaining written informed consent from the donor, in accordance with the Declaration of Helsinki and institutional guidelines. The study was approved by the local Ethics Committee (protocol *MetabolicStemCells*). Donors authorized the use of anonymized samples and data for research purposes and for publication.

Clones were manually selected and cultured in six well plate (Corning Incorporated, Costar) pre-coated with Vitronectin (0.5 µg/cm^2; ThermoScientific). Cells were maintained at 37 °C, 5% CO_2 in Essential 8 Basal Medium (Gibco) supplemented with 1% of Penicillin/Streptomycin (Lonza). Cells were passed using 0.5 mM EDTA (Ultrapure, ThermoScientific) upon reaching ~ 80% confluence.

All iPSC clones were karyotyped at ISENET (Milan, Italy) using a 16-hour treatment with 0.2 µg/mL colchicine, followed by Q-banding and array-based comparative genomic hybridization (aCGH) using a 600 K probe panel (median spacing: 41 kb; Agilent Technologies, Santa Clara, CA, USA).

Pluripotency was evaluated by immunofluorescence staining for OCT4 (Novus Biologicals) and SSEA4 (Abcam), and by flow cytometry analysis of OCT4 expression (BD, Bioscience). In addition, expression of the core pluripotency genes *POU5F1 (OCT4)*, *SOX2*, and *NANOG* was assessed by RT-PCR.

Immunofluorescence Staining

IPSCs were cultured on chamber slides and fixed with 4% paraformaldehyde (PFA) for 10 min at room temperature (RT). Cells were then washed three times with PBS and permeabilized using Perm Buffer (0.05% Tween 20, 0.2% Triton X-100 in PBS) for 10 min at RT. Following permeabilization, cells were incubated for 1 h at RT in Blocking ASE Buffer (2% BSA, 0.1% H2O2, 10mM Glycine in PBS). Primary antibodies (see Table S1) were applied overnight at 4 C°. The following day, cells were washed three times with PBS and incubated with the appropriate secondary antibodies (Table S1) for 1 h at RT. Nuclear staining was performed using staining Hoechst 33,342 (Invitrogen; 1:1000) for 10 min at RT, followed by three PBS washes. Coverslips were mounted using mounting medium, and images were acquired using an Olympus fluoVIEW FV3000RS confocal microscope.

Cytofluorimetric Analysis

Cells were harvested at defined differentiation stages using either 1× Trypsin (Lonza) or Accutase (Merck) for 5 min at 37 °C. Cell viability was assessed using the Fixable Near-IR Dead Cell Stain Kit (Life Technology), according to the manufacturer's protocol. For extracellular staining,

cells were washed with FACS buffer (PBS supplemented with 0.2% BSA) and incubated with primary antibodies for 15 min at RT. Cells were subsequently fixed with Cytofix/Cytoperm (BD Biosciences) for 20 min at 4 °C. For the intracellular staining, fixed cells were washed, permeabilized with BD Phosflow Perm Buffer III (BD Biosciences) for 30 min at 4° C in the dark and then incubated with the appropriate intracellular antibodies (see Table S1) for 30 min at 4 °C in the dark. All washes steps were performed at 1200 rpm for 5 min. Samples were analyzed on a FACS Canto cytometer (BD) and data were analyzed using FlowJo 10.8 Software (BD Life Sciences).

RNA Extraction

Total RNA was extracted using the mirVana miRNA Isolation Kit (AM1561, Invitrogen) according to the manufacturer's protocol. RNA concentrations were quantified by spectrophotometry (Epoch microplate reader, Gen5 software; BioTek, Winooski, VT, USA). Genomic DNA contamination was removed using the TURBO DNase Kit (Ambion). Complementary DNA (cDNA) was synthesized from purified RNA using the SuperScript IV Reverse Transcriptase Kit (Invitrogen), according to the manufacturer's instructions.

NanoString Analysis

Gene expression profiling of 770 stem cell-related genes was performed using the nCounter® Stem Cell Characterization Panel (NanoString Technologies, Seattle, WA, USA). For each sample, 50 ng of total RNA was hybridized and processed with the nCounter Prep Station, and analyzed using the nCounter Digital Analyzer (NanoString) with the 280 Field of View (FOV) setting, according to the manufacturer's protocol. Quality control and data normalization were conducted using the ROSALIND® platform (NanoString Technologies). Housekeeping genes for normalization were selected using the geNorm algorithm implemented in the *NormqPCR* R package. Differential expression analysis was conducted with false discovery rate (FDR) correction using the Benjamini–Hochberg method. Clustering of differentially expressed genes (DEGs) was performed for heatmap visualization using the Partitioning Around Medoids (PAM) algorithm in the fpc R package, taking into account pathway topology and gene interactions. Gene ontology and pathway enrichment analyses were carried out using a hypergeometric test across multiple reference databases, including InterPro, NCBI, MSigDB, Reactome, and WikiPathways. Principal component analysis (PCA) and unsupervised hierarchical clustering were additionally performed using nSolver™ 4.0 software (NanoString Technologies).

Real Time PCR on Selected DEGs

Quantitative real-time PCR (RT-qPCR) was performed using TaqMan Gene Expression Assays (Applied Biosystems) for the selected differentially expressed genes (DEGs):

- *upregulated in Gr2*: GBX2 (Hs00230965_m1), ZFHX3 (Hs00199344_m1), UTF1 (Hs00864535_s1), CNTFR (Hs00181798_m1), CPT1A (Hs00912671_m1).
- *downregulated in Gr2*: PDGFRB (Hs01019589_m1), TNNT2 (Hs00943911_m1), AFP (Hs01040598_m1), IDO1 (Hs00984148_m1), and ACBD7 (Hs00744676_s1).

GAPDH (Hs99999905_m1) was used as an endogenous control for normalization. Relative gene expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method. Reactions were run on a 7900 Real-Time PCR System (Applied Biosystems).

An overall z-score was calculated for each clone by standardizing the expression level of each selected DEG according to the formula $z = (x - \mu)/\sigma$, where x represents the gene expression value in a given sample, μ the mean expression across all samples, and σ the corresponding standard deviation; individual gene z-scores were then combined to generate a composite score reflecting the relative enrichment of the β cell–associated gene signature.

IPSC Differentiation into Definitive Endoderm and Insulin Producing Cells

Differentiation of iPSC clones toward pancreatic β-like cells was carried out using a multi-stage differentiation protocol. For each of the six iPSC clones included in the functional analyses, the differentiation protocol was independently repeated three times. All culture media were supplemented with 1% penicillin–streptomycin (Lonza) and replaced daily unless otherwise specified.

Stage 1. Definitive Endoderm (DE). Cells were seeded at a density of 2.1 × 10^5 cells/cm² and induced toward definitive endoderm for 3 days using the STEMdiff Definitive Endoderm Kit (STEMCELL Technologies), according to the manufacturer's instructions.

Following DE induction, cells were further differentiated into mature β-like cells using a seven-stage protocol [16], with minor modifications.

Stage 2. Pancreatic Progenitors (PP). DE cells were cultured for 3 days in MCDB131 medium (Life Technologies) supplemented with GlutaMAX 100× (Life Technologies), glucose (10 mM; Sigma-Aldrich), $NaHCO_3$ (1.5 g/L; Sigma-Aldrich), BSA (0.5%; Sigma-Aldrich), vitamin C (0.25 mM; Sigma-Aldrich), and FGF-7 (50 ng/mL; Genscript). On days 7 and 8, the medium was further supplemented with

SANT1 (0.625 mM; Sigma-Aldrich), retinoic acid (RA; 2.5 mM; Sigma-Aldrich), LDN-193,189 (0.25 mM; Selleckchem), and TPB (0.5 mM; Santa Cruz). From days 9 to 12, the medium was supplemented with hEGF (100 ng/mL; Peprotech), Activin A (10 ng/mL; Qkine), Y-27,632 (10 µM; Selleckchem), and nicotinamide (10 mM; Sigma-Aldrich).

Stage 3. Endocrine Commitment. Cells were incubated for 4 days in MCDB131 medium supplemented with GlutaMAX (100×), glucose (10 mM), $NaHCO_3$ (1.5 g/L), BSA (2%), ITS-X (100×; Thermo Fisher Scientific), $ZnSO_4$ (10 µM; Sigma-Aldrich), heparin (10 µg/mL; Sigma-Aldrich), betacellulin (100 ng/µL; Peprotech), Y-27,632 (10 mM), SANT1 (0.625 mM), RA (2.5 mM), LDN-193,189 (0.25 mM), GC1 (1.5 mM; Tocris), GSiXX (0.15 mM; Millipore), and ALK5inhII (15 mM; Selleckchem).

Stage 4. β-like Cell Maturation. Cells were cultured for 7 days in medium supplemented with LDN-193,189 (0.2 mM), GC1 (2 mM), GSiXX (0.2 mM), and ALK5inhII (20 mM). From day 24 onward, cells were maintained in CMRL 1066 medium (Corning) supplemented with GlutaMAX (100×), ITS-X (100×), sodium pyruvate (0.5 mM; Lonza), $ZnSO_4$ (10 µM), heparin (10 µg/mL), lipid concentrate (2000×; Invitrogen), trace elements A and B (2000× each; Cellgro), and BSA (2%), with the addition of N-acetylcysteine (1 mM; Sigma-Aldrich), triiodothyronine (T3; 10 nM; Sigma-Aldrich), and ZM-447,439 (0.5 µM; Selleckchem).

Differentiation efficiency was assessed at each stage by flow cytometry and quantitative PCR analysis of stage-specific markers.

Glucose-Stimulated Insulin Secretion

The functional response of mature β-like cells to glucose stimulation was evaluated using a dynamic perifusion system (BioRep® Perifusion V2.0.0, Biorep Technologies Inc., Miami Lakes, FL, USA). β-like cell clusters were loaded into a column containing Bio-Gel P-4 Gel (Bio-Rad, Hercules, CA, USA), maintained at 37 °C, and sequentially stimulated with 11 mM glucose followed by 30 mM KCl. Eluates were collected throughout the perifusion process, and human insulin concentrations were quantified using a commercial ELISA kit (Mercodia, Winston-Salem, NC, USA) according to the manufacturer's instructions. Optical density was measured with a microplate reader (Model 680, Bio-Rad). The stimulation index was calculated as the ratio of insulin secreted at high glucose over insulin secreted at low glucose.

Statistical Analysis

Statistical analyses were conducted using GraphPad Prism software (GraphPad Software, San Diego, CA, USA). Data were evaluated for normality, and appropriate tests were selected accordingly. For comparisons between two groups, either an unpaired Student's *t*-test (parametric) or the Mann–Whitney U test (non-parametric) was applied. Pearson correlation analysis was used to assess associations between continuous variables. Statistical significance was defined as $P \leq 0.05$.

Results

Generation and Characterization of iPSC Clones

A total of twenty iPSC clones, referred to as DRI1-iPSC, were generated by reprogramming CD34-enriched peripheral blood cells from a healthy donor, using Sendai virus technology. Of these eleven clones (#1, #3, #4, #5, #7, #9, #10, #11, #14, #16, #17) were stabilized (Fig. 1A) and their pluripotency was evaluated. All clones expressed key pluripotency markers, as confirmed by immunofluorescence for Oct4 and SSEA-4 (Fig. 1B). Flow cytometry revealed high Oct4 expression across clones, with an average of 99.45 ± 0.89% (range: 96.93–99.93%) (Fig. 1C). RT-PCR analysis confirmed expression of the core pluripotency genes *POU5F1 (OCT4)*, *SOX2*, and *NANOG* in all clones (Fig. 1D). Cytogenetic analysis demonstrated normal karyotypes across all clones, indicating preserved genomic stability (Fig. 1E). Collectively, these findings confirmed that the 11 selected DRI1-iPSC clones displayed comparable pluripotency profiles and maintained genomic integrity at the undifferentiated stage.

Heterogeneity among iPSC Clones During Differentiation into Insulin Producing Cells

All DRI1-iPSC clones exhibited a high and consistent Oct4 expression at the pluripotent stage, with an average of 99.45 ± 0.89% (range = 96.93–99.93%). However, marked interclonal variability emerged during subsequent differentiation stages. At the definitive endoderm (DE) stage, the proportion of $Cxcr4^+$ cells varied substantially among clones (mean 77.75 ± 18.58%; range = 30.79–94.65%), accompanied by heterogeneous residual expression of Oct4 (21.03 ± 18.81%; range = 3.80–67.65%). This variability persisted at the pancreatic progenitor (PP) stage, with Pdx1 expression averaging 61.31 ± 24.52% (range = 29.05–87.55%) and Nkx6.1 expression 38.14 ± 20.56% (range = 9.35–76.85%). At the insulin-producing β-like cell (iβ) stage, heterogeneity remained evident. The proportion of $Pdx1^+$ cells was 37.80 ± 10.07% (range = 21.47–49.63%) and $Nkx6.1^+$ cells averaged 53.52 ± 14.81% (range = 26.08–66.90%). Minor differences were observed in the proportion

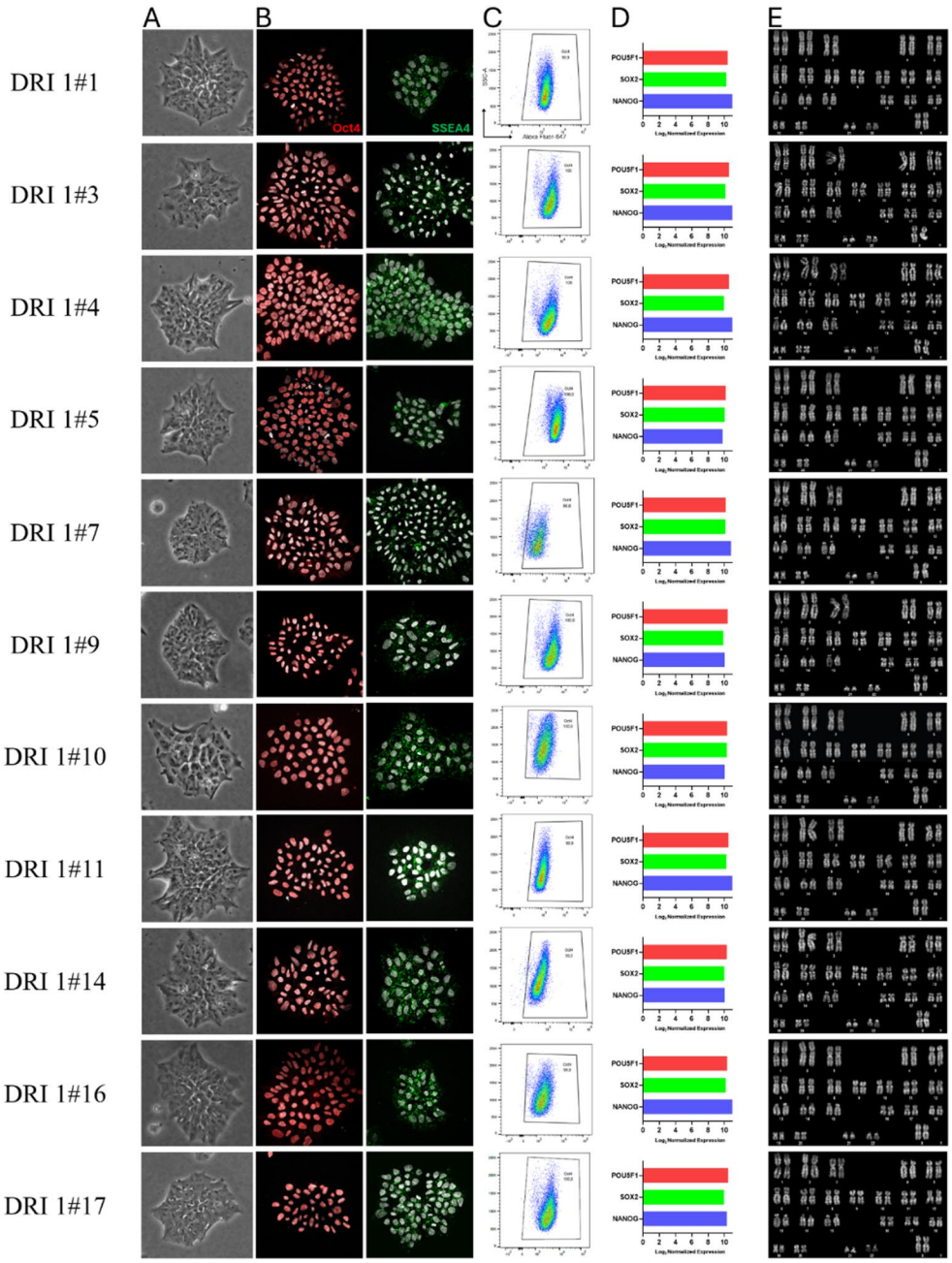


Fig. 1 Characterization of DRI1 iPSC clones. (**A**)Representative morphology of the 11 selected DRI1 iPSC clones at the undifferentiated stage (10× magnification). (**B**) Immunofluorescence staining for pluripotency markers Oct4 (red) and SSEA-4 (green); nuclei are counterstained with Hoechst (white) (60× magnification). (**C**) Flow cytometry analysis showing high Oct4 expression levels across all clones. (**D**) Normalized mRNA expression of pluripotency-associated genes POU5F1 (OCT4), SOX2, and NANOG by RT-PCR. (**E**) Karyograms of the 11 DRI1 iPSC clones

of insulin (Ins) positive cells at this stage (29.79±3.54%; range = 25.02–34.40%). These findings indicate that, despite comparable pluripotency and genomic stability at baseline, DRI1-iPSC clones differ markedly in their ability to progress through the DE, PP, and iβ differentiation stages (Fig. 2).

Transcriptional Variability among iPSC Clones at the Pluripotent Stage

To explore the molecular basis of the observed interclonal variability, we performed transcriptomic profiling of 770 genes in 11 DRI1-iPSC clones using a stem cell–focused panel covering pathways related to pluripotency, lineage specification, metabolism, and signalling regulation. As expected, the analysis clearly discriminated iPSC sample from DE samples, confirming both the panel's sensitivity and the robustness of our approach (Supplementary Fig. 1A). A total of 249 genes were differentially expressed between the iPSC and DE stages (LogFC > 1.5 or < −1.5; FDR-adjusted $P < 0.05$), including 181 upregulated and 68 downregulated genes (Supplementary File 1). Canonical pluripotency markers such as POU5F1 (OCT4) and NANOG were downregulated upon differentiation, whereas

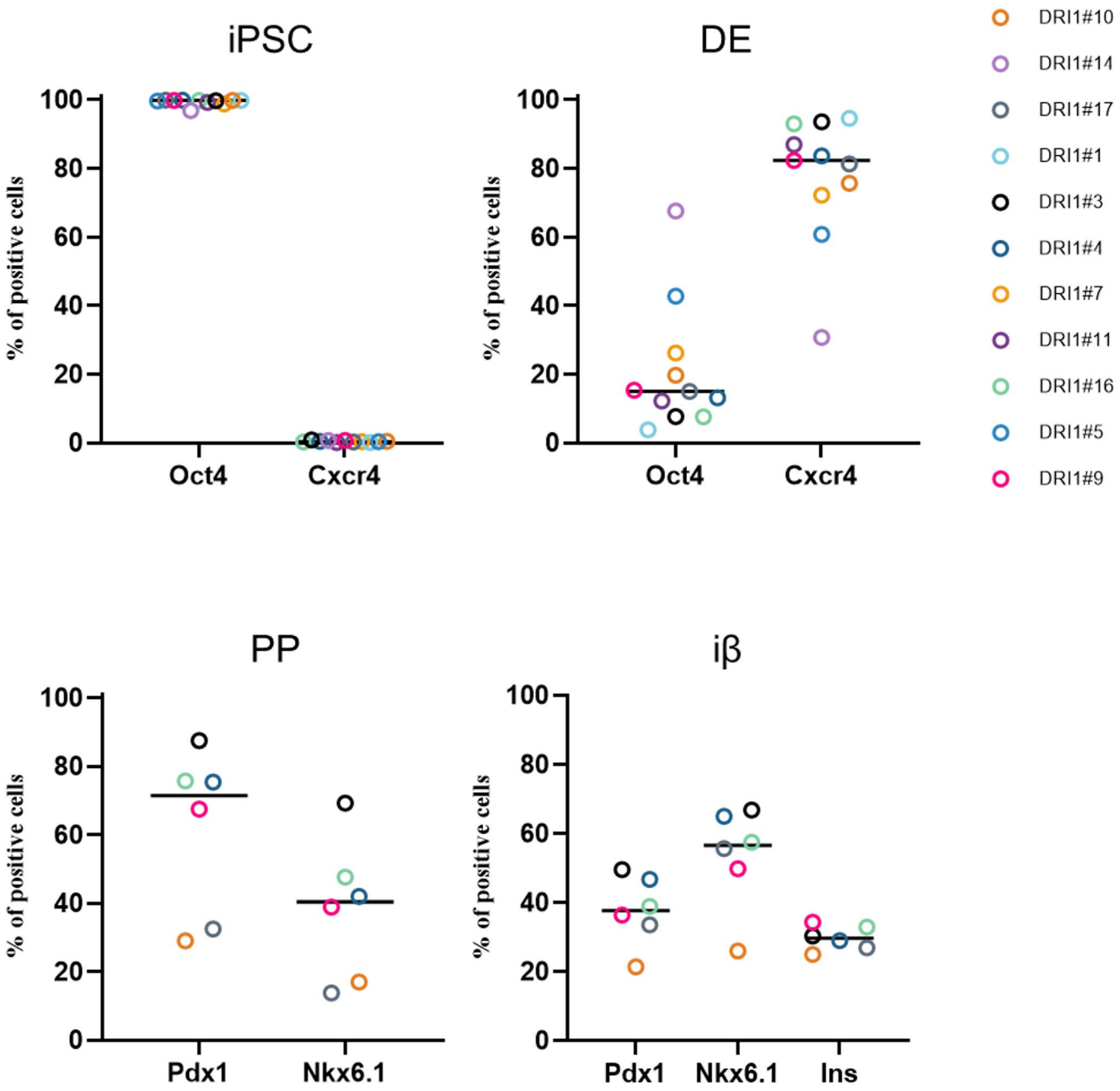
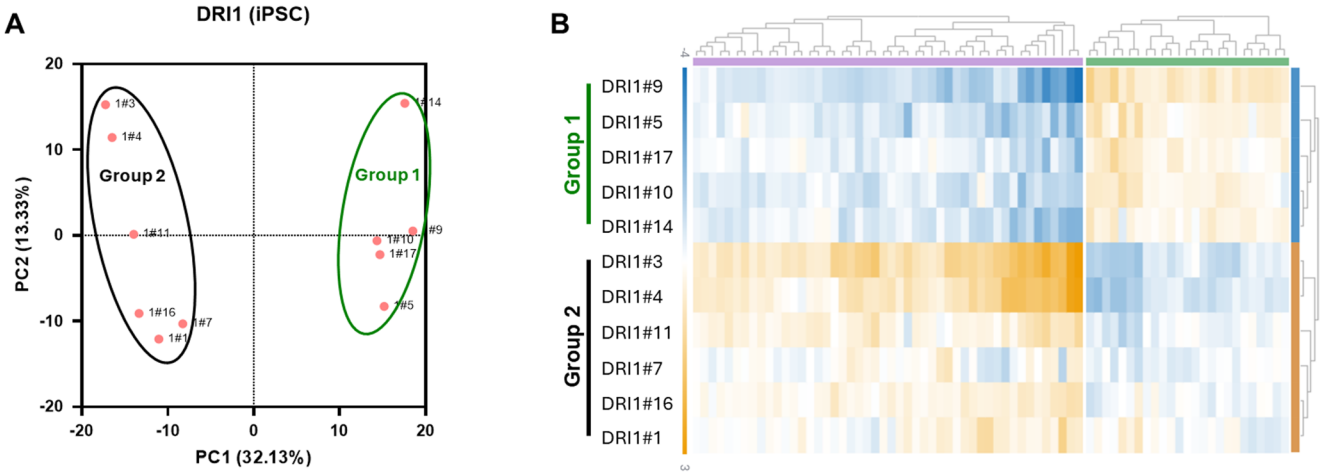


Fig. 2 Flow cytometry analysis of DRI1 iPSC clones at key stages of differentiation. Shown are the percentages of cells expressing stage-specific markers at the pluripotent ($Oct4^+$), definitive endoderm ($Cxcr4^+$), pancreatic progenitor ($Pdx1^+$, $Nkx6.1^+$), and insulin-producing cell (Ins^+, $Pdx1^+/Ins^+$, $Nkx6.1^+/Ins^+$) stages

endoderm-associated genes including FOXA2 and GATA6 were upregulated. Pathway enrichment analysis revealed significant associations with endoderm development, naïve-to-primed pluripotency transitions, and signaling pathways such as MET/EMT and TGF-β (Supplementary Fig. 1B). Beyond these expected stage-specific differences, principal component analysis (PCA) revealed notable transcriptional heterogeneity among the 11 clones already at the pluripotent stage (Fig. 3A). Specifically, clones #5, #9, #10, #14, and #17 (group 1, Gr1) exhibited a distinct gene expression profile compared to clones #1, #3, #4, #7, #11, and #16 (group 2, Gr2). Differential expression analysis between Gr1 and Gr2 identified 73 significantly modulated genes (25 upregulated, 48 downregulated; LogFC > 1.5 or < −1.5; FDR-adjusted $P < 0.05$) (Fig. 3B; Supplementary File 2). These genes were primarily linked to naïve/primed pluripotency states, early endodermal commitment, and metabolic pathways such as glutamine, amino acid, and fatty acid metabolism (Fig. 3C). To validate these findings, the 10 most differentially expressed genes (5 upregulated, 5



C

Term ID	Number of Genes in Term	Undirected Enrichment Score
Naïve State	29	8.6063
Primed State	10	5.6579
Endodermal Lineage	42	5.4947
Rho/ROCK Signaling	10	5.3451
Glutamine Metabolism	10	5.2714
Hippo Signaling	53	4.8057
Amino Acid Metabolism	42	4.7414
Mesodermal Lineage	67	4.6888
Fatty Acid Metabolism	28	4.6299
TGF-beta Signaling	59	4.4875

Fig. 3 Early transcriptional heterogeneity among iPSC clones. (**A**) Principal component analysis (PCA) of gene expression profiles from 11 DRI1-iPSC clones at the pluripotent stage, revealing segregation into two transcriptionally distinct groups (Gr1 and Gr2). (**B**) Heatmap of 73 differentially expressed genes (DEGs) distinguishing Gr1 and Gr2 clones (LogFC > 1.5 or < –1.5; FDR-adjusted P < 0.05). (**C**) Gene ontology (GO) enrichment analysis of DEGs, highlighting pathways related to naïve and primed pluripotency states, endodermal lineage commitment, and metabolic processes

downregulated) were assessed by qPCR across all 11 DRI1-iPSC clones. Expression patterns were consistent with transcriptomic data and clearly separated Gr1 from Gr2 clones (Fig. 4). Specifically, *GBX2*, *ZFHX3*, *UTF1*, *CNTFR*, and *CPT1A* were significantly upregulated in Gr2, while *PDGFRB*, *TNNT2*, *AFP*, *IDO1*, and *ACBD7* were downregulated ($P < 0.05$ for all comparisons). Taken together, these results demonstrate that substantial transcriptional heterogeneity exists among donor-matched iPSC clones at the pluripotent stage. These intrinsic molecular differences likely underpin the divergent differentiation trajectories observed and support the rationale for early-stage transcriptomic screening to predict lineage-specific competence.

Distinct Differentiation Efficiency and Functional Maturation Between iPSC Clone Groups

To determine whether transcriptional differences observed at the pluripotent stage translated into divergent differentiation outcomes, six DRI1-iPSC clones were selected based on the PCA-defined grouping: clones #9, #10, and #17 from Gr1, and clones #3, #4, and #16 from Gr2. These clones were subjected to a standardized, stepwise differentiation protocol toward mature insulin-producing cells over 30 days, with progression monitored using stage-specific markers (Fig. 5A).

At the iPSC stage (day 0), all clones were uniformly positive for Oct4 (Gr1: 99.91 ± 0.03%; Gr2: 99.86 ± 0.07%) and negative for Cxcr4, confirming comparable baseline pluripotency. Divergence emerged at the DE stage (day 4), where Gr2 clones showed a higher proportion of $Cxcr4^+$ cells (90.13 ± 5.57%) compared to Gr1 (79.81 ± 3.57%; $p = 0.027$), along with a lower frequency of residual $Oct4^+$ cells (Gr2: 9.49 ± 3.16%; Gr1: 16.72 ± 2.60%; $p = 0.05$).

Differences became more pronounced at the pancreatic progenitor (PP) stage (day 15): Gr2 clones exhibited significantly higher percentages of $Pdx1^+$ cells (79.60 ± 6.89% vs. 43.04 ± 21.30%; $p = 0.05$), $Nkx6.1^+$ cells (53.00 ± 14.44% vs. 23.29 ± 13.65%; $p = 0.03$), and $Pdx1^+/Nkx6.1^+$ double-positive cells (49.22 ± 16.98% vs. 19.62 ± 12.93%;

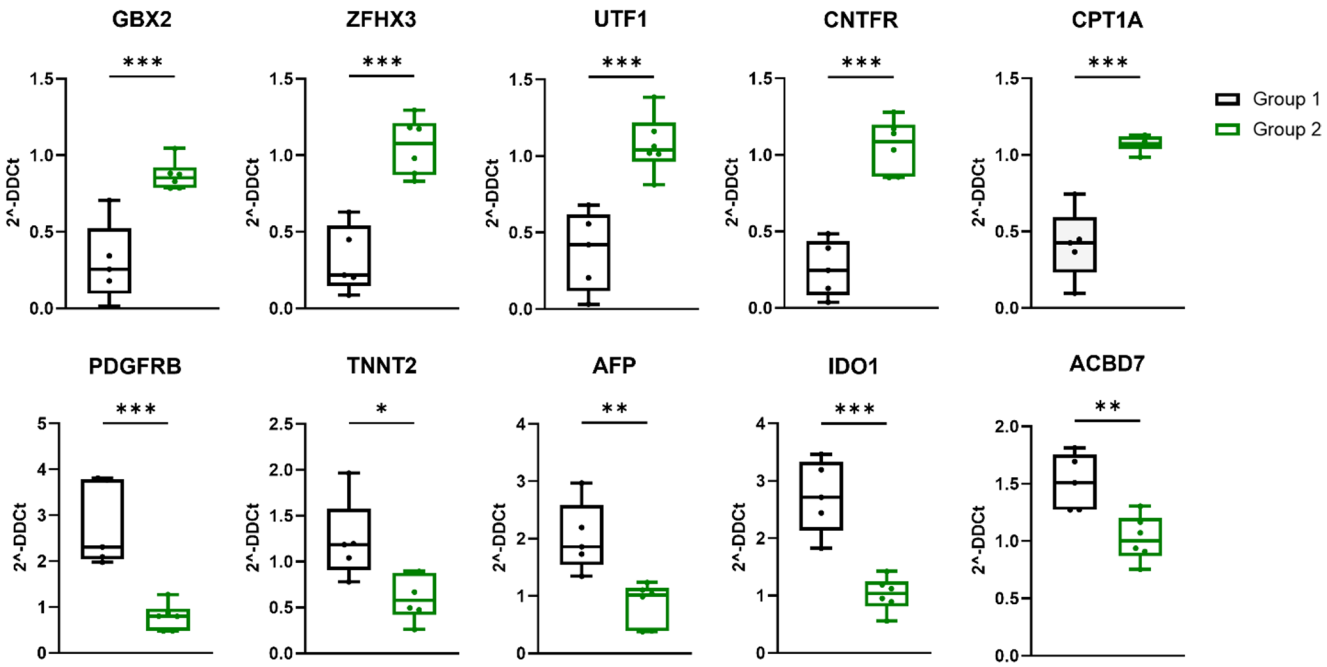


Fig. 4 Validation of group-specific transcriptional differences by RT-qPCR. Box plots showing the expression levels of the top 10 differentially expressed genes (DEGs) distinguishing Gr1 and Gr2 iPSC clones at the pluripotent stage. Statistical significance was defined as $P<0.05$ (*), $P<0.01$ (**) and $P<0.001$ (***). Lines within boxes indicate medians, and whiskers represent variability outside the upper and lower quartiles. Five genes were upregulated in Gr2 compared with Gr1 (GBX2, ZFHX3, UTF1, CNTFR, CPT1A), whereas five were downregulated (PDGFRB, TNNT2, AFP, IDO1, ACBD7)

$p=0.05$). At the iβ stage (day 30), Gr2 clones maintained this advantage, with significantly greater frequencies of $Pdx1^+$ (44.11 ± 5.54% vs. 30.50 ± 7.95%; $p=0.03$), $Nkx6.1^+$ (63.16 ± 4.93% vs. 43.88 ± 15.69%; $p=0.05$), and double-positive $Pdx1^+/Ins^+$ (21.48 ± 2.47% vs. 16.49 ± 2.03%; $p=0.03$) and $Nkx6.1^+/Ins^+$ (21.36 ± 1.82% vs. 17.54 ± 2.69%; $p=0.05$) cells.

These phenotypic differences were supported by qPCR, which confirmed significantly higher expression levels of *PDX1* and *NKX6.1* in Gr2 clones at both the PP stage ($P=0.009$ and 0.05, respectively) and iβ stage ($P=0.02$ for both) (Fig. 5B). Functionally, β-like cells derived from Gr2 clones demonstrated superior glucose responsiveness. In dynamic perifusion assays, insulin secretion in response to high glucose was significantly higher in Gr2-derived cells (7.13 ± 1.38 pg/mL) than in those from Gr1 (4.07 ± 0.95 pg/mL; $P=0.04$). A similar trend was observed upon KCl-induced depolarization (26.16 ± 5.23 pg/mL vs. 12.11 ± 4.30 pg/mL; $P=0.02$). Total insulin output, measured as area under the curve (AUC), was approximately 2.2-fold greater in Gr2 compared to Gr1 (379.6 ± 121.6 vs. 174.6 ± 121.6) (Fig. 5C).

Thus, clones belonging to Gr2 not only progressed more efficiently through the pancreatic lineage but also produced β-like cells with superior glucose-responsive insulin secretion, indicating that early molecular profiles can predict downstream differentiation performance.

Early Gene-Expression Signature Predicts β Cell Differentiation Performance

To assess whether transcriptional variability among iPSC clones could predict downstream differentiation outcomes, we integrated the expression levels of the 10 validated differentially expressed genes (DEGs) into a composite gene-expression score. For each clone, a standardized *z*-score was calculated, providing a unified metric of relative expression across the gene set (Fig. 6A). This gene-expression signature was then correlated with phenotypic readouts of β cell development, including stage-specific marker expression and insulin secretion. Pearson correlation analysis revealed strong positive associations ($r>0.5$) between the z-score and key developmental markers at multiple stages (Fig. 6B). Specifically, higher z-scores correlated with increased Cxcr4 expression at the DE stage, elevated Pdx1 expression at the pancreatic progenitor (PP) stage, and enhanced Nkx6.1 expression at the insulin-producing cell (iβ) stage. Importantly, the z-score also showed positive correlations with the proportion of mature β-like cells co-expressing Pdx1, Nkx6.1, and insulin, as well as with functional insulin secretion during dynamic perifusion assays. Clones with higher transcriptomic scores consistently exhibited greater glucose responsiveness and insulin output. This supports the utility of early-stage transcriptomic profiling as a strategy

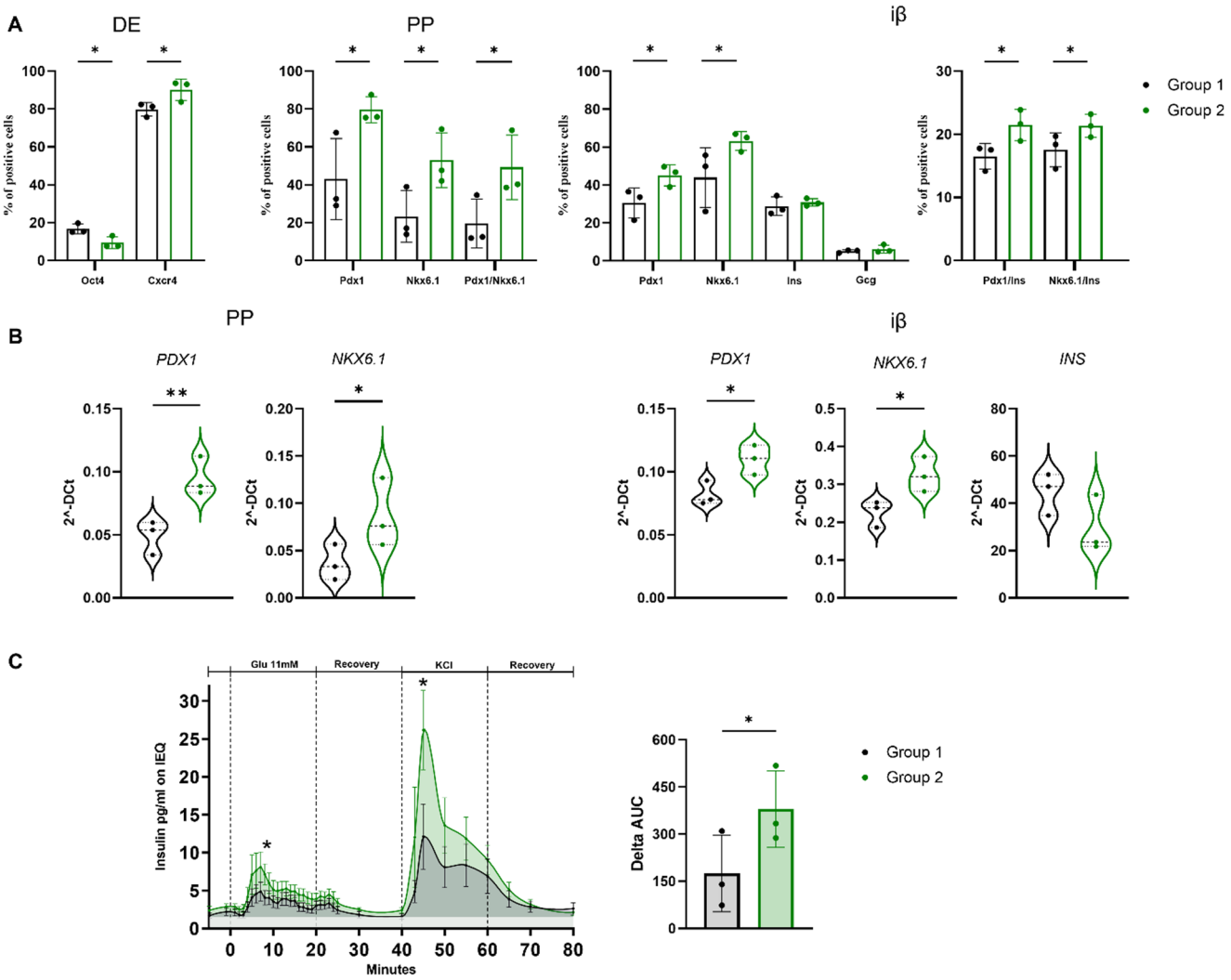


Fig. 5 Divergent differentiation efficiency and functional maturation of Gr1 and Gr2 iPSC clones. (**A**) Flow cytometry analysis of stage-specific markers during differentiation of six selected DRI1-iPSC clones (three from Gr1 and three from Gr2, each differentiated 3 times) at the definitive endoderm (DE), pancreatic progenitor (PP), and insulin-producing β-like cell (iβ) stages. (**B**) RT-qPCR quantification of PDX1, NKX6.1, and INS expression at the PP and iβ stages in three independent differentiation runs (**C**) Dynamic perifusion analysis of glucose-responsive insulin secretion in terminally differentiated β-like cells. Cell clusters were loaded into perifusion columns containing Bio-Gel P-4 Gel and maintained at 37 °C. Samples were perifused with basal medium (0.5 mM glucose), stimulated with 11 mM glucose, and subsequently depolarized with 30 mM KCl. Insulin concentrations were measured in the effluent at multiple time points (left). The delta area under the curve (AUC) relative to basal secretion is shown on the right

to preselect high-performing iPSC clones prior to initiating time- and resource-intensive differentiation protocols.

Discussion

The efficient generation of functional β-like cells from induced pluripotent stem cells (iPSCs) holds transformative promise for cell-based therapy in type 1 diabetes. However, clonal heterogeneity remains a significant barrier to clinical translation, as donor-matched iPSC clones often exhibit divergent capacities for lineage progression. Here, we demonstrate that such variability is already evident at the pluripotent stage and can be anticipated through early transcriptional profiling.

Our findings show that iPSC clones segregate into two transcriptionally distinct groups (Gr1 and Gr2) prior to differentiation, and that these groupings predict downstream performance across all key stages: definitive endoderm (DE), pancreatic progenitor (PP), and insulin-producing cell stages. Clones in Gr2 consistently outperformed those in Gr1 in yielding higher proportions of $Pdx1^{+}$, $Nkx6.1^{+}$, and insulin-positive cells, and exhibiting enhanced glucose-stimulated insulin secretion. These observations align with previous reports attributing interclonal variability to residual epigenetic memory, stochastic reprogramming events, or

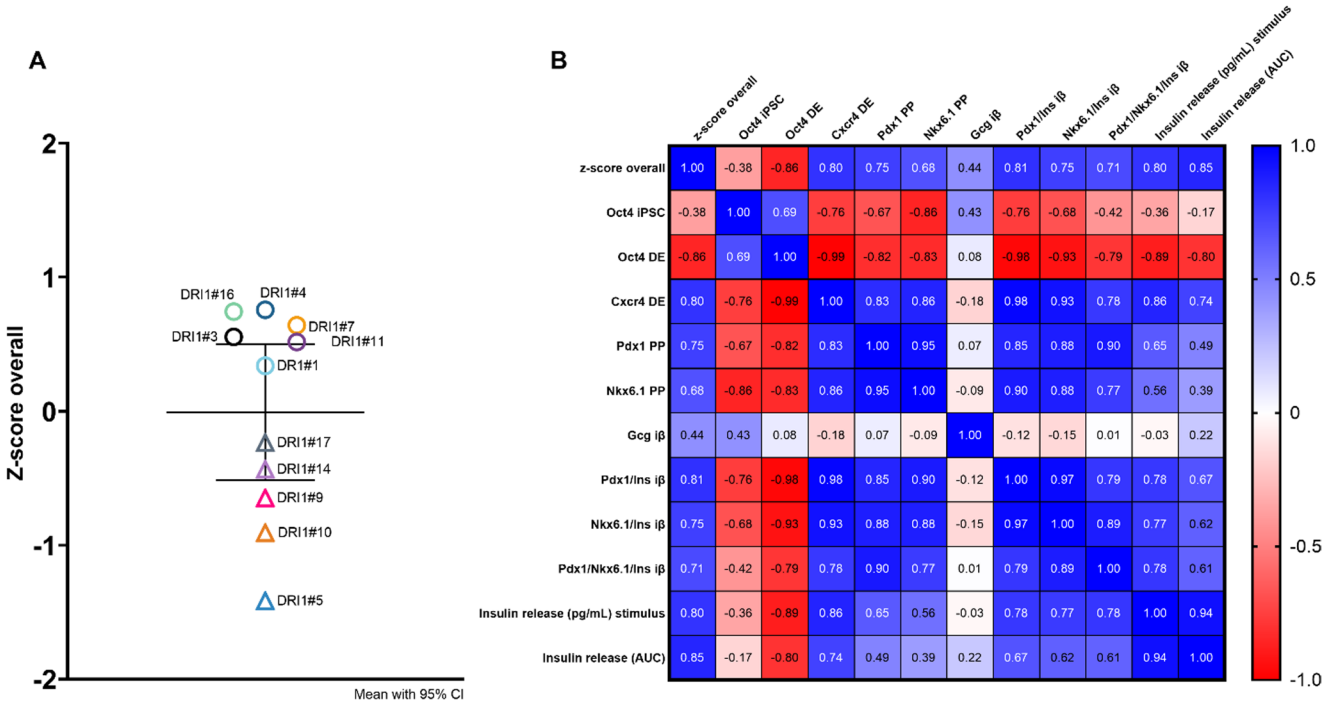


Fig. 6 Early gene-expression signature correlates with β-like cell differentiation outcomes. (**A**) Z-score derived from the expression of the 10 selected DEGs in 11 DRI1 iPSC clones at the pluripotent stage. Triangles and circles represent clones belonging to Gr1 and Gr2, respectively. (**B**) Pearson correlation matrix showing associations between the z-score, β cell marker expression measured by flow cytometry, and insulin secretion during dynamic perifusion

metabolic priming [9, 15]. A major contribution of this study is the identification of a ten-gene transcriptional signature, comprising five upregulated and five downregulated genes in Gr2 versus Gr1 that captures this intrinsic variability. This gene set, validated through qPCR, includes regulators of pluripotency and germ layer specification such as UTF1 and GBX2, as well as genes associated with alternative lineage biases such as PDGFRB and AFP [7, 16]. The presence of metabolic genes such as CPT1A, whose expression correlates with differentiation fitness, aligns with growing evidence for a role of metabolic state in dictating stem cell fate [1]. The integration of these genes into a composite *z*-score yielded a robust quantitative metric that positively correlated with stage-specific differentiation markers and functional β cell output. This supports the concept that intrinsic molecular states at the pluripotent level can prime iPSCs for specific developmental trajectories and functional endpoints. While transcriptomic profiling has been used to evaluate differentiation potential toward cardiomyocytes or neural lineages [15, 17], our work provides one of the first examples of a reproducible, predictive signature for pancreatic endocrine differentiation.

In recent years, artificial intelligence–based approaches, particularly those combining machine learning or deep learning with label-free imaging, have emerged as powerful tools for anticipating cell fate decisions in stem cell systems. Several studies have demonstrated that subtle morphological features captured at early stages encode latent information related to lineage commitment, even before canonical molecular markers become detectable [18, 19]. This concept is further supported by work showing that cell morphology alone can forecast subsequent differentiation outcomes, highlighting a tight coupling between cell shape, dynamic behavior, and the underlying molecular state [20, 21]. Within this framework, imaging-based predictive strategies offer important advantages, including non-invasive monitoring, scalability, and compatibility with high-throughput workflows; however, they typically provide limited mechanistic insight into the molecular determinants of fate decisions. In parallel with the present transcriptomic study, we contributed to a complementary work employing label-free imaging and deep learning to estimate iPSC differentiation potential at early stages [22]. While the imaging-based approach enables rapid and automated clone stratification, the gene expression strategy presented here offers higher-resolution molecular insight into lineage priming and functional competence. An integrated platform combining both modalities may therefore further enhance predictive accuracy and streamline iPSC clone selection.

This study has limitations. Functional differentiation was evaluated in a limited number of clones from a single donor, and broader validation across genetically diverse

backgrounds, reprogramming technologies, and differentiation protocols will be required to generalize these findings. Accordingly, the present study should be interpreted as a proof-of-concept demonstrating the feasibility of early transcriptomic stratification of iPSC clones, rather than as a definitive or broadly validated predictive framework. At the same time, the donor-matched experimental design partially mitigates genetic variability, allowing a focused assessment of intrinsic clonal heterogeneity. Incorporating additional molecular layers, such as chromatin accessibility or DNA methylation, could further refine predictive models and reveal mechanistic drivers of lineage commitment.

In conclusion, this study establishes proof of concept that early transcriptomic profiling can reliably identify iPSC clones with high propensity for β cell differentiation. In this context, patient-derived iPSCs represent a particularly valuable platform for the development of autologous β cell–based strategies and for investigating immune–endocrine interactions relevant to type 1 diabetes. Such an approach could significantly accelerate the development of iPSC-based therapies for diabetes by reducing reliance on time-consuming and resource-intensive empirical screeing.

Data Availability

All data are available in the main text or the supplementary materials.

Contributions

Writing– review & editing - LP.; Writing– original draft – VZ; VS; Methodology – VZ; VS; Investigation – VZ; LM, MM, CC, SP; Data curation - VZ; VS.; Conceptualization - VZ; VS; LP.; Supervision – VS; LP; Funding acquisition -LP.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12015-026-11091-y.

Author contributions Writing– review & editing - LP.; Writing– original draft – VZ; VS; Methodology – VZ; VS; Investigation – VZ; LM, MM, CC, SP; Data curation - VZ; VS.; Conceptualization - VZ; VS; LP.; Supervision – VS; LP; Funding acquisition -LP;

Funding This research was supported by the Hub Life Science – Terapia Avanzata (LSH–TA) project (PNC-E3-2022-23683269; CUP: C43C22002210001), funded by the Italian Ministry of Health under the Piano Nazionale Complementare – Ecosistema Innovativo della Salute (unique investment code: PNC-E.3), awarded to VS and LP. Additional support was provided by the fundraising campaign "SOStegno 70 – Insieme ai ragazzi diabetici" Associazione Onlus through the project "Beta is Better". The authors gratefully thank Paola Macchieraldo, Antonio Mincione, Elena Riva, Antonio Civita, Andrea Marchesi, and Michele Mainardi for their support of the fundraising campaign "Un brutto t1po". The fundraising campaign also supported LM's PhD fellowship.

Data Availability All data are available in the main text or the supplementary materials.

Declarations

Consent for Publication Donors authorized the use of anonymized samples and data for research purposes and for publication.

Consent to Participate Not Applicable.

Human Ethics Approval Human biological samples were collected after obtaining written informed consent from the donor, in accordance with the Declaration of Helsinki and institutional guidelines. The study was approved by the local Ethics Committee (Comitato Etico Territoriale Lombardia 1-CETL1; protocol *MetabolicStemCells*).

Clinical Trial Number Not applicable.

Competing Interests The authors declare no competing interests.

References

1. Shi, Y., Inoue, H., Wu, J. C., & Yamanaka, S. (2017). Induced pluripotent stem cell technology: A decade of progress. *Nature Reviews Drug Discovery, 16*(2), 115–130. https://doi.org/10.1038/nrd.2016.245
2. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019
3. Sordi, V., Pellegrini, S., & Piemonti, L. (2017). Immunological Issues After Stem Cell-Based β Cell Replacement. *Current Diabetes Reports*, *17*(9). https://doi.org/10.1007/s11892-017-0901-4
4. Maxwell, K. G., & Millman, J. R. (2021). Applications of *iPSC*-derived beta cells from patients with diabetes. *Cell Reports Medicine, 2*(4), Article 100238. https://doi.org/10.1016/J.XCRM.2021.100238
5. Wang, S., Du, Y., Zhang, B., Meng, G., Liu, Z., Liew, S. Y., Liang, R., Zhang, Z., Cai, X., Wu, S., Gao, W., Zhuang, D., Zou, J., Huang, H., Wang, M., Wang, X., Liang, T., Liu, T., Gu, J., … Shen, Z. (2024). Transplantation of chemically induced pluripotent stem-cell-derived islets under abdominal anterior rectus sheath in a type 1 diabetes patient. *Cell*. https://doi.org/10.1016/j.cell.2024.09.004
6. Wu, J., Li, T., Guo, M., Ji, J., Meng, X., Fu, T., … Fu, Z. (2024, December 1). Treating a type 2 diabetic patient with impaired

pancreatic islet function by personalized endoderm stem cell-derived islet tissue. *Cell Discovery*. Springer Nature. https://doi.org/10.1038/s41421-024-00662-3.

7. Hartjes, K. A., Li, X., Martinez-Fernandez, A., Roemmich, A. J., Larsen, B. T., Terzic, A., & Nelson, T. J. (2014). Transgene-free disease-specific induced pluripotent stem cells from patients with type 1 and type 2 diabetes. *Stem Cells*, *32*(9). https://doi.org/10.1002/stem.1734
8. Kudva, Y. C., Ohmine, S., Greder, L. V., Dutton, J. R., Armstrong, A., De Lamo, J. G., Khan, Y. K., Thatava, T., Hasegawa, M., Fusaki, N., Slack, J. M., & Ikeda, Y. (2012). Transgene-free disease-specific induced pluripotent stem cells from patients with type 1 and type 2 diabetes. *Stem Cells Translational Medicine, 1*(6), 451–61. https://doi.org/10.5966/sctm.2011-0044
9. Efrat, S. (2021). Epigenetic memory: Lessons from *iPS* cells derived from human β cells. *Frontiers in Endocrinology*. https://doi.org/10.3389/fendo.2020.614234
10. Fantuzzi, F., Toivonen, S., Schiavo, A. A., Chae, H., Tariq, M., Sawatani, T., Pachera, N., Cai, Y., Vinci, C., Virgilio, E., Ladriere, L., Suleiman, M., Marchetti, P., Jonas, J.-C., Gilon, P., Eizirik, D., Igoillo-Esteve, M., & Cnop, M. (2022). In depth functional characterization of human induced pluripotent stem cell-derived beta cells in vitro and in vivo. *Frontiers in Cell and Developmental Biology*. https://doi.org/10.3389/fcell.2022.967765
11. Arroyave, F., Uscátegui, Y., & Lizcano, F. (2024). From iPSCs to pancreatic β cells: Unveiling molecular pathways and enhancements with vitamin C and retinoic acid in diabetes research. *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms25179654
12. Southard, S. M., Kotipatruni, R. P., & Rust, W. L. (2018). Generation and selection of pluripotent stem cells for robust differentiation to insulin-secreting cells capable of reversing diabetes in rodents. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0203126
13. Song, H. W., Solomon, J. N., Masri, F., Mack, A., Durand, N., Cameau, E., Dianat, N., Hunter, A., Oh, S., Schoen, B., Marsh, M., Bravery, C., Sumen, C., Clarke, D., Bharti, K., Allickson, J. G., & Lakshmipathy, U. (2024). Bioprocessing considerations for generation of *iPSCs* intended for clinical application: Perspectives from the ISCT Emerging Regenerative Medicine Technology working group. *Cytotherapy*. https://doi.org/10.1016/J.JCYT.2024.05.024
14. Nath, S. C., Menendez, L., & Friedrich Ben-Nun, I. (2023, December 1). Overcoming the Variability of iPSCs in the Manufacturing of Cell-Based Therapies. *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms242316929
15. Cuomo, A. S. E., Seaton, D. D., McCarthy, D. J., Martinez, I., Bonder, M. J., Garcia-Bernardo, J., & Stegle, O. (2020). Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14457-z
16. Barsby, T., Ibrahim, H., Lithovius, V., Montaser, H., Balboa, D., Vähäkangas, E., Chandra, V., Saarimäki-Vire, J., & Otonkoski, T. (2022). Differentiating functional human islet-like aggregates from pluripotent stem cells. *STAR Protocols*. https://doi.org/10.1016/j.xpro.2022.101711
17. Banovich, N. E., Li, Y. I., Raj, A., Ward, M. C., Greenside, P., Calderon, D., Tung, P. Y., Burnett, J. E., Myrthil, M., Thomas, S. M., Burrows, C. K., Romero, I. G., Pavlovic, B. J., Kundaje, A., Pritchard, J. K., & Gilad, Y. (2018). Impact of regulatory variation across human iPSCs and differentiated cells. *Genome Research, 28*(1), 122–131. https://doi.org/10.1101/gr.224436.117
18. Soelistyo, C. J., Vallardi, G., Charras, G., & Lowe, A. R. (2022). Learning biophysical determinants of cell fate with deep neural networks. *Nature Machine Intelligence, 4*(7), 636–644. https://doi.org/10.1038/s42256-022-00503-6
19. Cunha, I., Latron, E., Bauer, S., Sage, D., & Griffié, J. (2024, October 1). Machine learning in microscopy–insights, opportunities and challenges. *Journal of Cell Science*. Company of Biologists Ltd. https://doi.org/10.1242/jcs.262095
20. Palma, A., Theis, F. J., & Lotfollahi, M. (2025). Predicting cell morphological responses to perturbations using generative modeling. *Nature Communications, 16*(1), Article 505. https://doi.org/10.1038/s41467-024-55707-8
21. Cunha, I., Panconi, L., Bauer, S., Gestin, M., Latron, E., Sahai, E., … Griffié,J. (n.d.). AI4CellFate: Interpretable Early Cell Fate Prediction with Generative AI. https://doi.org/10.1101/2025.05.12.653464.
22. Schöb, F. J., Binder, A., Zamarian, V., Sordi, V., Scholz, H., Malthe-Sørenssen, A., & Dysthe, D. K. (2025, May 27). Deep Learning for Predicting Stem Cell Efficiency for use in Beta Cell Differentiation. *bioRxiv*. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2025.05.22.652867