

## Supplemental Information

**Title: A novel disease mechanism leading to the expression of a disallowed gene in the pancreatic beta-cell identified by non-coding, regulatory mutations controlling *HK1***

### Authors

Matthew N. Wakeling<sup>1\*</sup>, Nick D. L. Owens<sup>1\*</sup>, Jessica R. Hopkinson<sup>1</sup>, Matthew B. Johnson<sup>1</sup>, Jayne A.L. Houghton<sup>2</sup>, Antonia Dastamani<sup>3</sup>, Christine S. Flaxman<sup>1</sup>, Rebecca C. Wyatt<sup>1</sup>, Thomas I. Hewat<sup>1</sup>, Jasmin J. Hopkins<sup>1</sup>, Thomas W. Laver<sup>1</sup>, Rachel Van Heugten<sup>2</sup>, Michael N. Weedon<sup>1</sup>, Elisa De Franco<sup>1</sup>, Kashyap A. Patel<sup>1</sup>, Sian Ellard<sup>2</sup>, Noel G. Morgan<sup>1</sup>, Edmund Cheesman<sup>4</sup>, Indraneel Banerjee<sup>5,6</sup>, Andrew T. Hattersley<sup>1</sup>, Mark J. Dunne<sup>6</sup>, International Congenital Hyperinsulinism Consortium, Sarah J. Richardson<sup>1</sup> and, Sarah E. Flanagan<sup>1</sup>

### Affiliations

1. Institute of Biomedical and Clinical Science, University of Exeter Medical School, UK
2. Exeter Genomics Laboratory, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK
3. Endocrinology Department, Great Ormond Street Hospital for Children, London, UK
4. Department of Paediatric Pathology, Royal Manchester Children's Hospital, Oxford Road, Manchester, UK
5. Department of Paediatric Endocrinology, Royal Manchester Children's Hospital, Oxford Road, Manchester, UK
6. Faculty of Biology, Medicine and Health, The University of Manchester, Oxford Road, Manchester, UK

**Supplemental Table 1:** Summarised clinical characteristics for 162 probands diagnosed with congenital hyperinsulinism screened for mutations in the *HK1* regulatory region by whole-genome sequencing (n=135) or Sanger sequencing and ddPCR (n=27). Median values (IQR) provided.

	<b>Summary data</b>
<b>Female sex</b>	80 (49.1%)
<b>Birth weight SDS</b>	0.10 (-1.16 – 1.14)
<b>Age at onset (days)</b>	56 (7 – 168)
<b>Blood glucose at presentation (mmol/l)</b>	1.6 (1.2 – 2.1)
<b>Insulin at presentation (pmol/l)</b>	48 (25 – 100)
<b>Pancreatectomy performed</b>	34 (20.8%)

**Supplemental Table 2:** Summarised clinical characteristics for 17 individuals from 14 families with *HK1* mutations. Median values (IQR) provided.

	<b>Summary data</b>
<b>Sex (female)</b>	9 (53%)
<b>Birth weight (SDS)</b>	0.61 (0.10-1.84)
<b>Age at onset</b>	Birth (0-14 days)
<b>Blood glucose at presentation (mmol/l)</b>	1.6 (1.2-2.1)
<b>Insulin during hypoglycaemia (pmol/l)</b>	163 (95-326)
<b>Pancreatic resection performed</b>	5/17 (29%)

**Supplemental Table 3:** Primer sequences used for **a)** ddPCR to confirm the deletions identified by whole-genome sequencing in two individuals and an affected twin **b)** ddPCR to screen for deletions in the replication cohort within the 151bp region harbouring *de novo* single nucleotide variants **c)** Sanger sequencing of the *HK1* element

**a)**

Primer	Sequence (5' – 3')	Region amplified (Hg19)
HK1_A_Forward	tccctgcagtaatcagaatctgg	Chr10:71105908-71106051
HK1_A_Reverse	agcgccaatttatccctccg	
HK1_B_Forward	agggtctgcgtggctttatg	Chr10:71107602-71107732
HK1_B_Reverse	ggaaaccctaaaatacttcacccc	
HK1_C_Forward	ctcttacagtggcagggacg	Chr10:71109774-71109906
HK1_C_Reverse	aagccaatcacagccctctc	
HK1_D_Forward	gtgattggtacagaaagactcagtg	Chr10:71111625-71111726
HK1_D_Reverse	cagtatgaagtagctcctcgag	

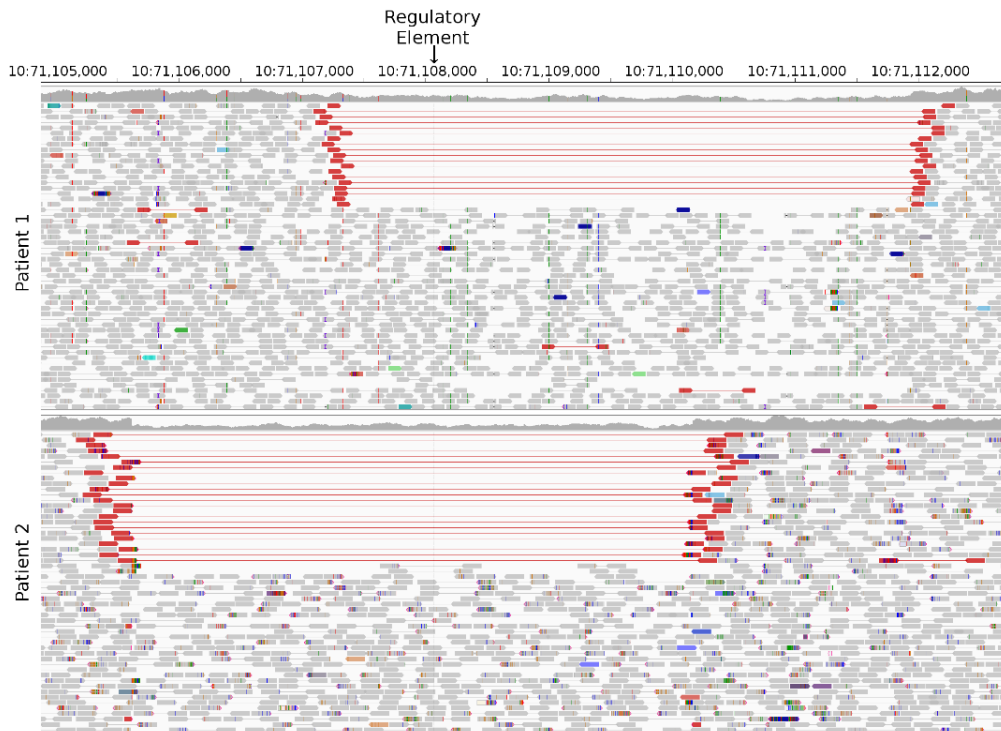
**b)**

Primer	Sequence (5'-3')	Region amplified (Hg19)
GCK_exon_1_Foward	TCCAATTGAGAAGCCTACTG	Chr7:44228423-44228617
GCK_exon_1_Reverse	TCAGATTCTGAGGCTCAAAC	
HK1_E_Forward	caggctggtactcgagacac	Chr10:71108653-71108835
HK1_E_Reverse	tgctgtgacctttctcac	
HK1_F_Forward	aggcagagttttgcttgcc	Chr10:71108620-71108804
HK1_F_Reverse	ccactgaagctgggagaacc	

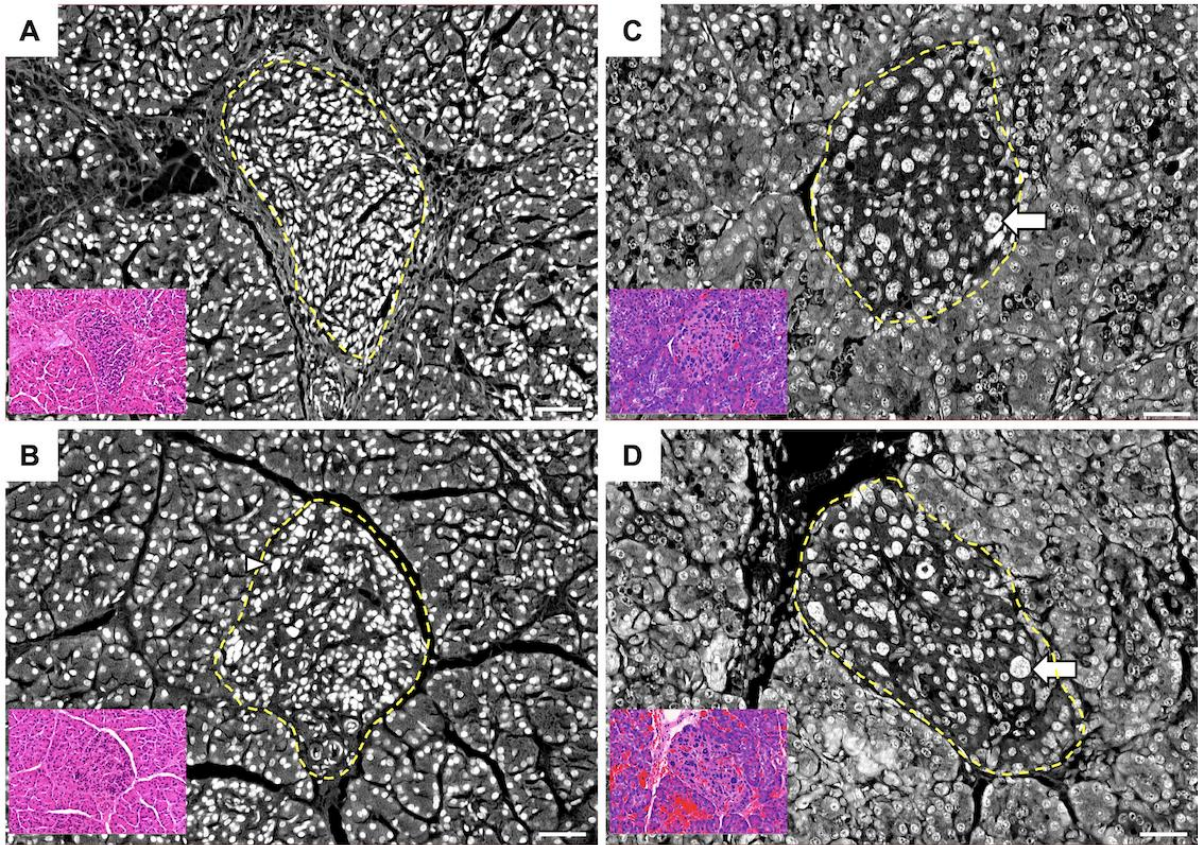
**c)**

	Sequence (5'-3')
<b>Forward</b>	AGCCTGGGCAACAGAAAC
<b>Reverse</b>	GCTACAAGCTCAGCCTCTTTC

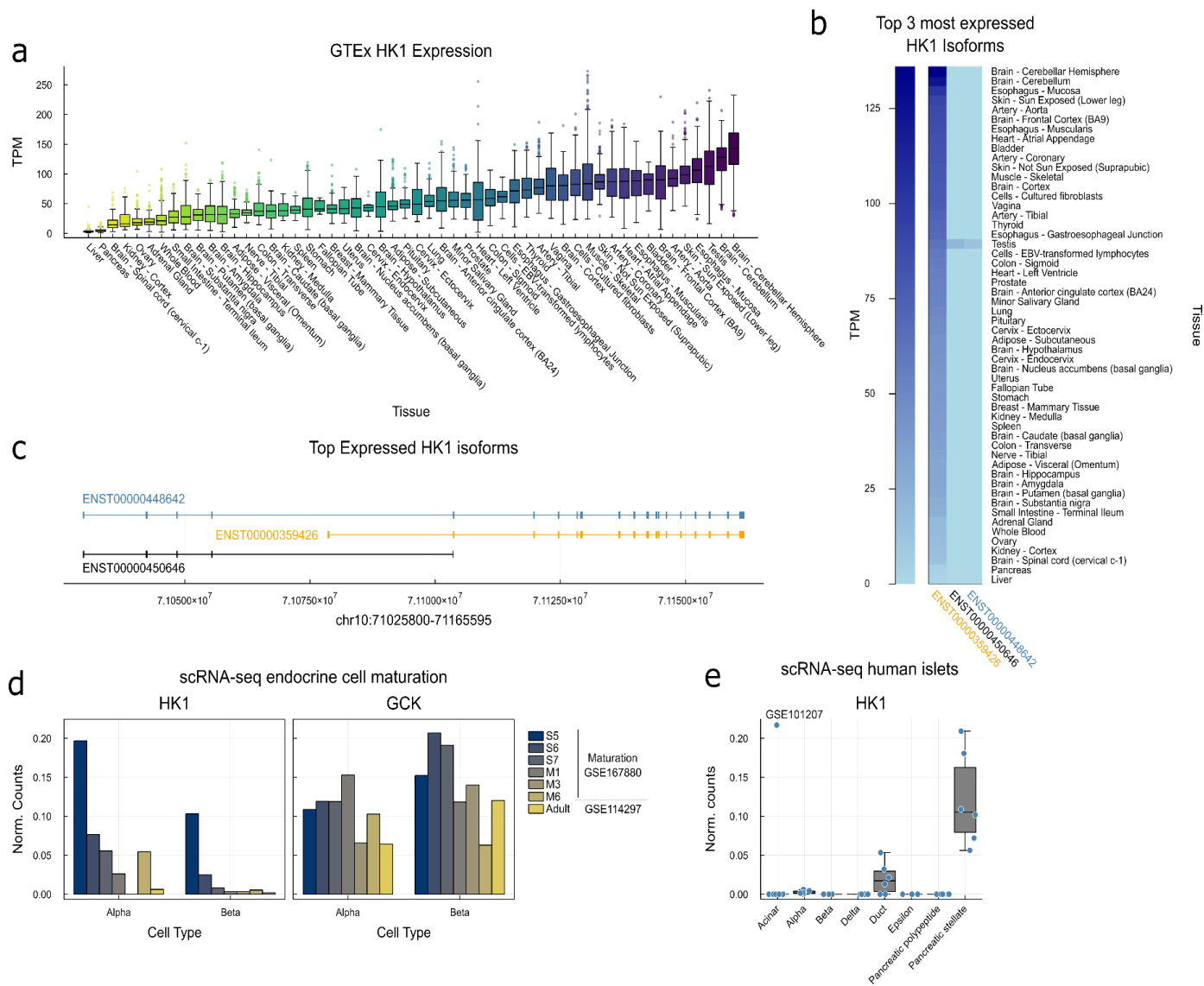
**Supplemental Table 4:** Table describing the primary and secondary candidate disrupted transcription factor motifs by each mutation.



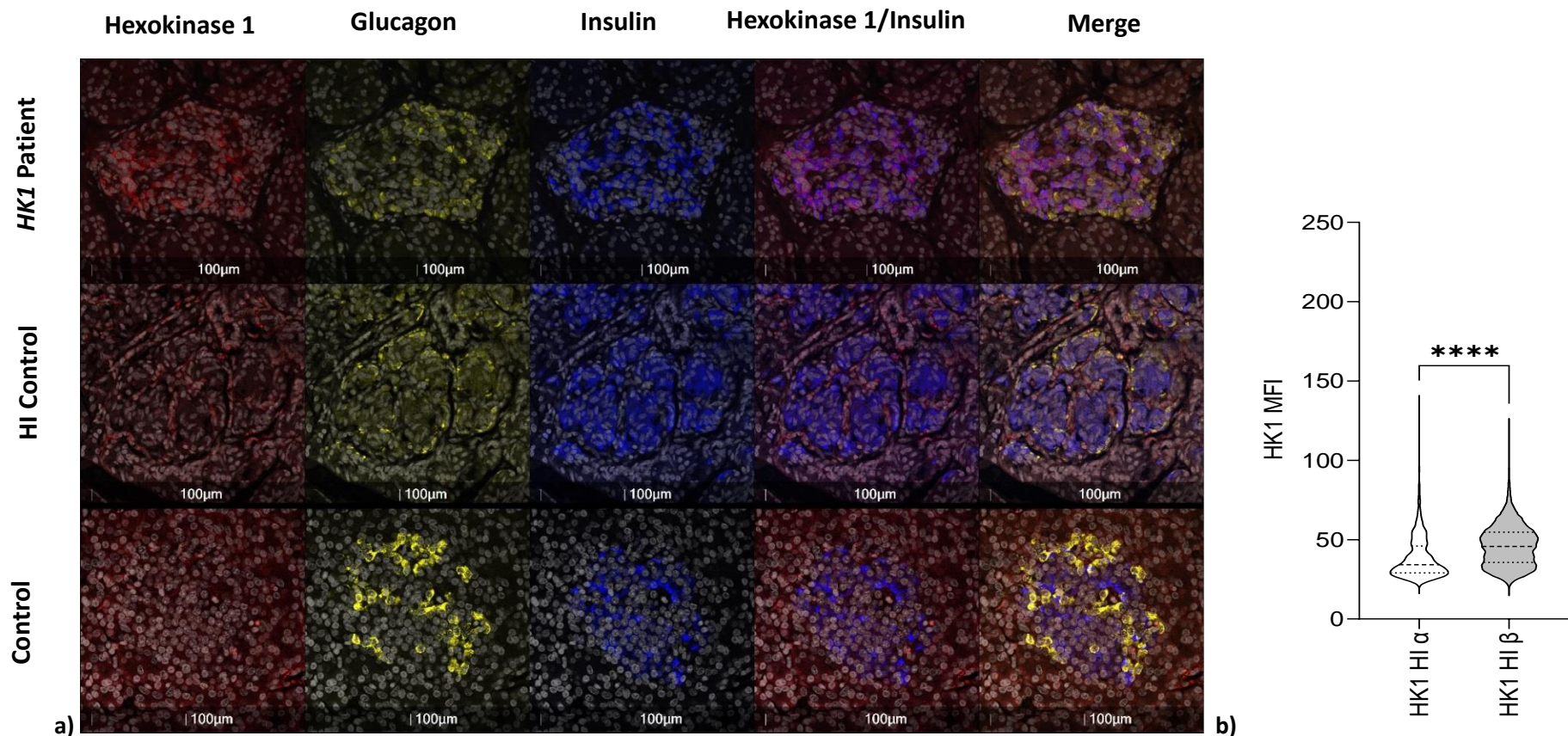
**Supplemental Figure 1: Integrative Genomics Viewer (IGV) screenshot of the ~4.5Kb deleted regions encompassing the regulatory element in intron 2 of *HK1* in two patients.** Both deletions are heterozygous, so reads are present in the deleted regions from the other haplotype, although the read depth is reduced. Reads highlighted in red have a long insert size. Both features are consistent with a deletion. The deleted region in one patient has detectable break points, as shown by the discontinuity in read depth and the reads clipped at the deletion boundary. In the second patient, the breakpoints do not result in read depth discontinuity or clipped reads as the deletion occurs between two matching repeating sequences.



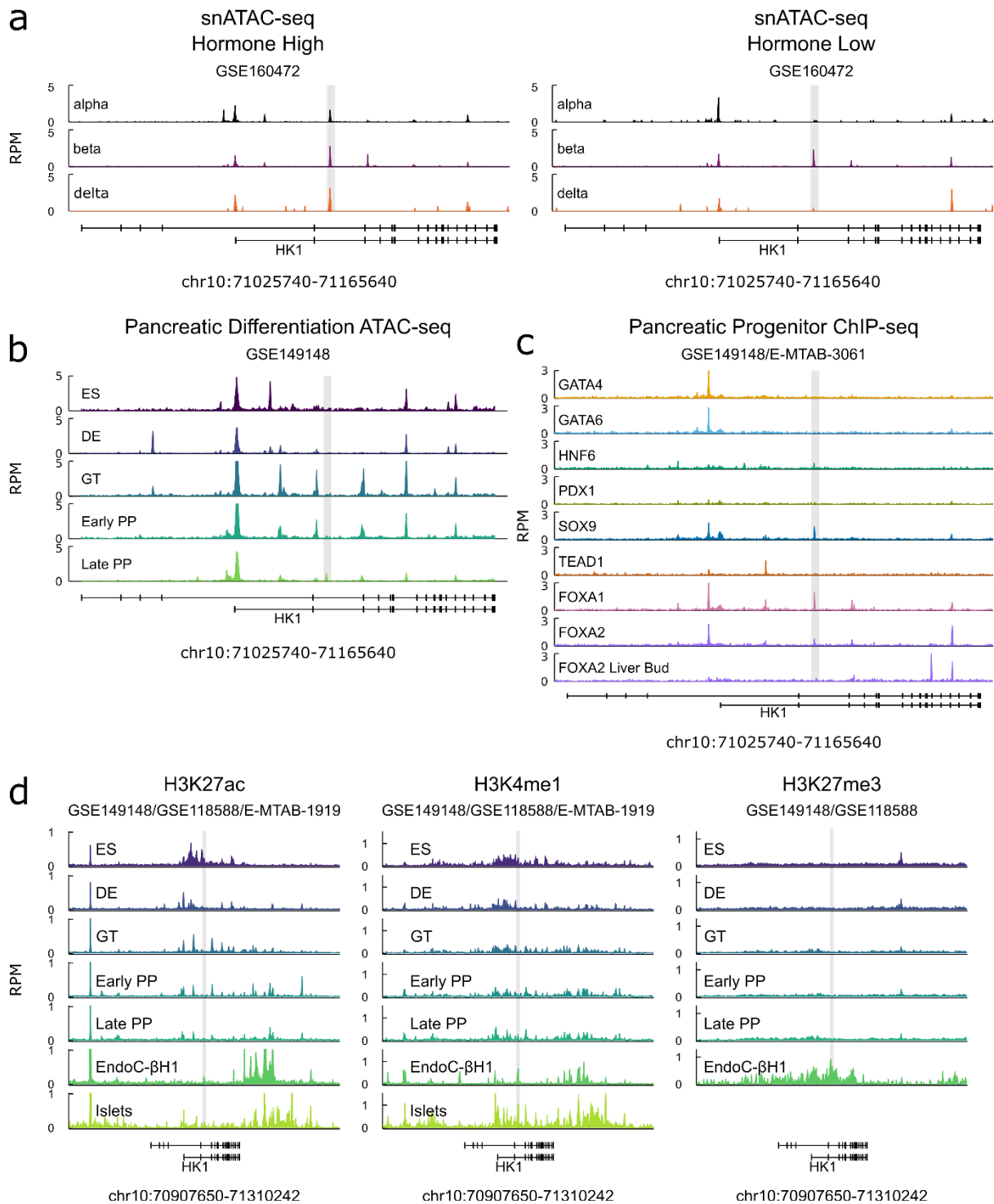
**Supplemental Figure 2. Histological appearance of islets from *HK1*-CHI and *ABCC8*-HI tissue.** Panels **A** and **B** illustrate islets from two *HK1* CHI patients in comparison to age-matched tissue from *ABCC8* CHI patients with diffuse pancreatic disease (**C and D**). The yellow dashed line delineates a single islet in each sample. Nuclear enlargement (B, arrowhead) and nucleomegaly (C, D arrows) was virtually undetected in control islets ( $n=12/72,145$  islet cells) but found in 0.8% of *HK1* islets ( $n=374/46,885$  cells) and 4.9% of *ABCC8* islet cells ( $n=70,929$ ). Inserts illustrate the corresponding haematoxylin and eosin images. Scale bar 50um



**Supplemental Figure 3: HK1 expression in GTEx** **a)** Boxplots describing gene level HK1 expression in Genotype-Tissue Expression GTEx project ordered by expression with least expression in liver and pancreas. Data shows sum of GTEx v8 isoform level data calculated with RSEM. The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 08/26/21. **b)** Isoform level HK1 expression in GTEx tissues as a), top 3 most expressed isoforms are given. Long HK1 isoform only expressed in testis, short isoform ubiquitously expressed across all tissues. **c)** Transcript models of top three HK1 isoforms given in b). **d)** Expression of HK1 and GCK for comparison in scRNA-seq across alpha and beta-cell maturation, all data quantified by (GSE167880)<sup>1</sup>, which quantified their own data over maturation and adult endocrine cells (GSE114297)<sup>2</sup>. **e)** HK1 expression in human islet cell types by scRNA-seq (GSE101207)<sup>3</sup>. Data points from independent human donors, boxplots describing median and IQR. HK1 expression only in duct and pancreatic stellate cells, expression absent in endocrine cell types.



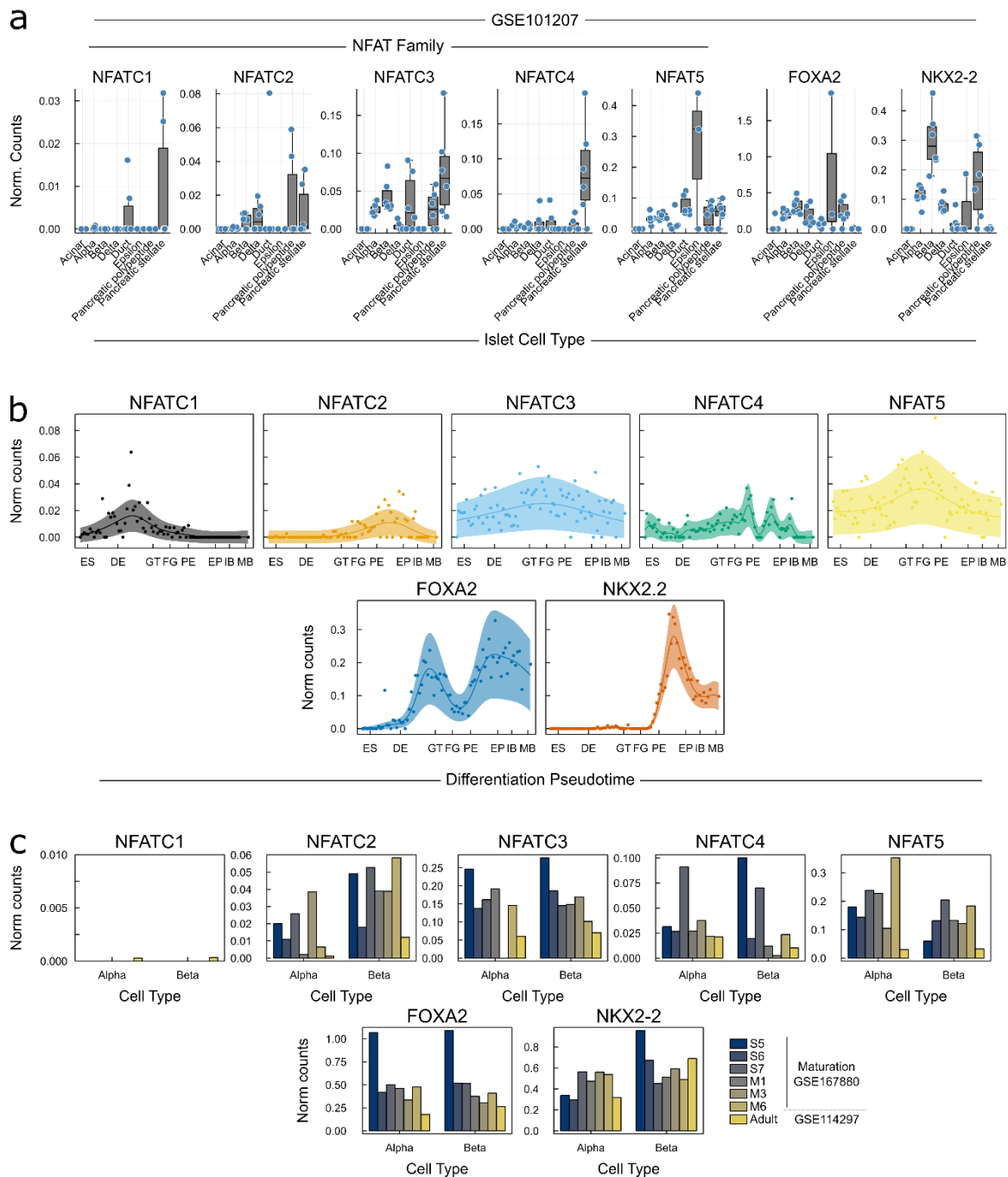
**Supplementary Figure 4: HK1 expression is present in the beta-cells of donors with *HK1*-CHI, but not *ABCC8*-Hyperinsulinism or Non-HI controls.** **a)** Staining of HK1 (red); glucagon (yellow); insulin (blue) and DAPI (white) in pancreatic tissue resected from one patient with a *HK1* mutation. HK1 co-localises with insulin (purple staining), but not with glucagon. Staining in tissue resected from an individual with *ABCC8*-Hyperinsulinism and a Non-HI control shows no expression of HK1 in either alpha (glucagon-positive) or beta (insulin-positive cells). Scale bar – 100µm. **b)** HK1 expression (Median fluorescence intensity (MFI)) is higher in the *HK1* HI donor beta-cells (n=2 donors; 17015 beta-cells), when compared to alpha-cells (n=3034). Two-tailed Mann-Whitney (U=16323914; \*\*\*\*p<0.0001). This suggests the impact of the *HK1* mutation is beta-cell specific.



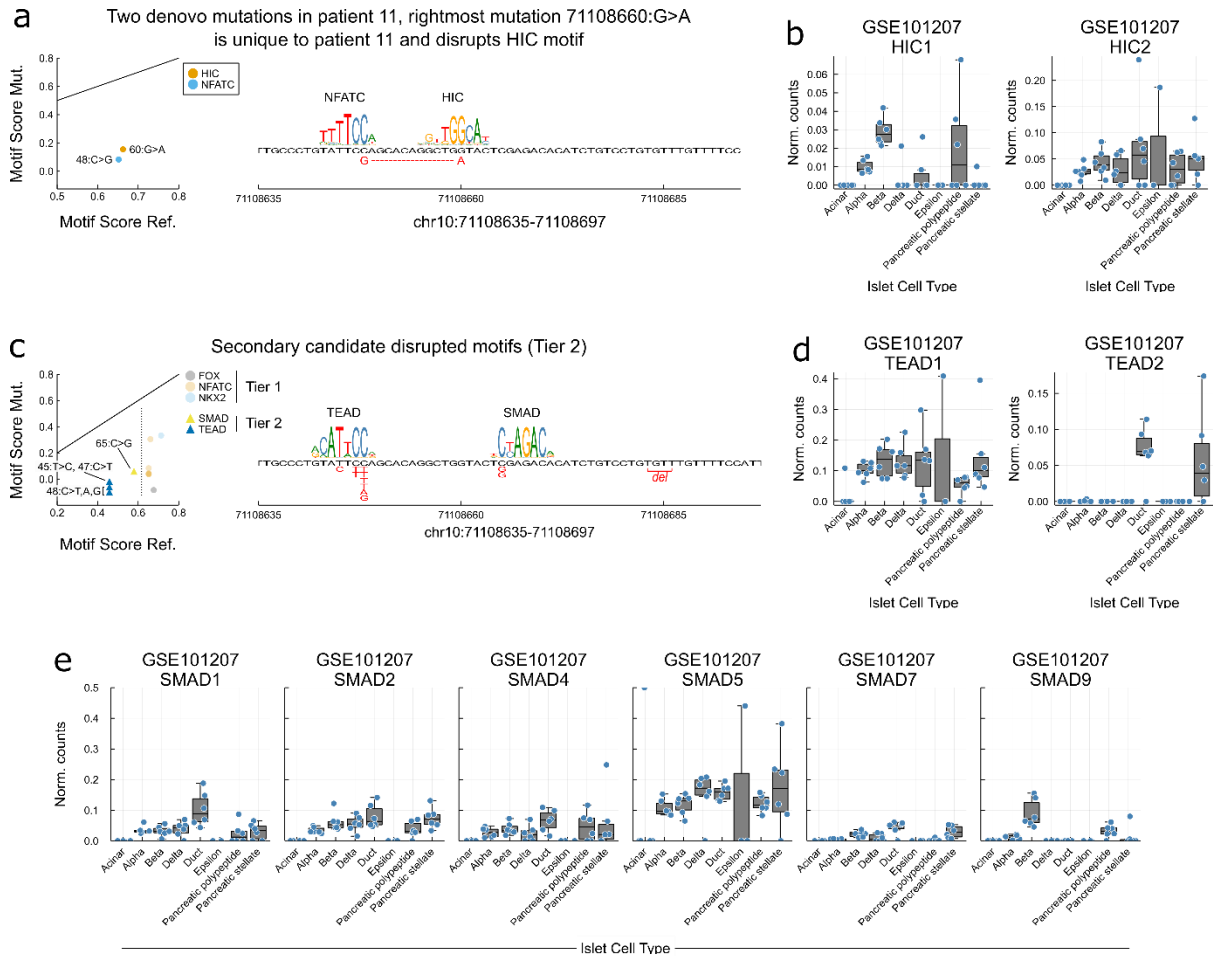
**Supplemental Figure 5: Chromatin accessibility, Transcription Factor binding and Epigenetic modifications at the *HK1* locus over pancreatic differentiation** a) snATAC-seq data in human islets showing mean chromatin accessibility over nuclei assigned to alpha, beta and delta cell clusters in hormone high and hormone low conditions (GSE160472)<sup>4</sup>. Critical region containing mutations encompassed within grey box. Alpha, beta and delta cells each show accessibility in hormone high state, whereas only beta-cells retain accessibility in hormone low cells. Hormone high and low states characterised by level of accessibility over endocrine cell secreted hormone promoter, alpha – GCG, beta –



INS, delta – SST. Data show in reads per million (RPM). **b)** Chromatin accessibility of pancreatic differentiation assessed by ATAC-seq (GSE149148)<sup>5</sup>, wider locus view of same data shown in **Figure 3c**. Critical region containing mutations encompassed within grey box. Stages: ES – embryonic stem cell, DE – definitive endoderm, Early/Late PP – pancreatic progenitor. **c)** Transcription factor binding in pancreatic progenitors (GSE149148)<sup>5</sup> and in Carnegie Stage 16-18 liver bud for FOXA2 (E-MTAB-3061)<sup>6</sup>, data shown as a) and b). Liver bud data reveals that whilst FOXA2 binds critical region in pancreatic progenitors and human islets (**Figure 3b**), it does not bind in liver bud, suggesting that critical region encodes pancreas specific regulation of *HK1*. **d)** Epigenetic modifications over *HK1* locus, shown are active marks H3K27ac and H3K4me1 and repressive mark H3K27me3 over in vitro pancreatic cell differentiation (GSE149148)<sup>5</sup> stages as in b); in vitro beta-cell line EndoC-BH1 (GSE118588)<sup>5</sup> and human islets (E-MTAB-1919)<sup>7</sup>. Grey box encompasses critical region containing mutations. Data reveals that H3K27ac broadly marks short isoform promoter (**Supplemental Figure 3**) at ES cells stage is reduced by DE stage, whilst focal mark over regulatory region is present in EndoC-BH1 and human islets. Similarly, H3K4me1 shows string focal enrichment over regulatory region in EndoC-BH1 and islets. Finally, repressive mark H3K27me3 is broadly absent over the locus over the course of pancreatic cell differentiation but broadly marks region in mature Endoc-BH1 cells.



**Supplemental Figure 6: Expression of family members of disrupted transcription factor binding motifs in critical region: NFAT family, NKX2-2 and FOXA2: a)** In single-cell RNA-seq in human islets (GSE101207)<sup>3</sup>, boxplots of mean normalised counts from independent islet donors in cells assigned to islet cell type clusters (data shown as in **Supplemental Figure 3e**). **b)** Over the course of pancreatic differentiation from embryonic stem cells to maturing beta-cells, expression data over a beta-cell differentiation pseudotime<sup>8</sup>, Gaussian process regression median and 95% CI shown. ES – embryonic stem cell, DE – definitive endoderm, GT – gut tube, FG – foregut, PE – pancreatic endoderm, EP – endocrine precursors, IB – immature beta-cells, MB – maturing beta-cells. **c)** scRNA-seq across alpha and beta-cell maturation, all data quantified by (GSE167880)<sup>1</sup>, which quantified their own data over maturation and adult endocrine cells (GSE114297)<sup>2</sup>.



### Supplemental Figure 7: Additional transcription factor motifs disrupted by mutations

**a-b)** Sanger sequencing identified 2 *de novo* variants 12 base pairs apart

(g.71,108,648C>G and g.71,108,660G>A) in one patient. It is not known whether these variants are *in cis* or *in trans*. g.71,108,648 is affected by 3 different *de novo* substitutions in 6 individuals and resides within a predicted NFATC binding motif. Two further *de novo* mutations identified in 3 patients are also predicted to impact on NFATC binding (**Figure 3**). In contrast, further variants affecting g.71,108,660 have not been identified, this suggests that the g.71,108,648C>G is causative of the disease in this patient.

Nevertheless, it is of note that g.71,108,660G>A is predicted to affect a HIC family motif **a)** and human islet scRNA-seq (GSE101207)<sup>3</sup> determines HIC2 is expressed in beta-

cells **b)** Interestingly, HIC2 is a transactivator of SIRT1<sup>9</sup> and the loss of SIRT1 impairs glucose sensing in beta-cells in mice<sup>10</sup>.

**c-e)** Tier 2, secondary candidate motifs for disruption have lower normalised motif scores than Tier 1 matches. **c)** Two additional motif families are implicated TEAD and SMAD. TEAD shares partial consensus with NFAT (TTCCA) and is alternative candidate to NFAT. TEAD1 but not TEAD2 is expressed in beta-cells **e)** and TEAD1 plays a critical role in pancreatic progenitors<sup>6</sup>, however it should be noted that TEAD1 does not bind the critical region in pancreatic progenitors when the region is bound by FOXA2 (**Figure 3b**). The SMAD family are signal transducers of TGF-beta signalling and play an important role in beta-cell development, function, and proliferation<sup>11</sup>.

## References

1. Balboa, D. *et al.* Functional, metabolic and transcriptional maturation of stem cell derived beta cells. *bioRxiv*, 2021.03.31.437748 (2021).
2. Sui, L. *et al.* Reduced replication fork speed promotes pancreatic endocrine differentiation and controls graft size. *JCI Insight* **6**(2021).
3. Fang, Z. *et al.* Single-Cell Heterogeneity Analysis and CRISPR Screen Identify Key beta-Cell-Specific Disease Genes. *Cell Rep* **26**, 3132-3144 e7 (2019).
4. Chiou, J. *et al.* Single-cell chromatin accessibility identifies pancreatic islet cell type- and state-specific regulatory programs of diabetes risk. *Nat Genet* **53**, 455-466 (2021).
5. Geusz, R.J. *et al.* Pancreatic progenitor epigenome maps prioritize type 2 diabetes risk genes with roles in development. *Elife* **10**(2021).
6. Cebola, I. *et al.* TEAD and YAP regulate the enhancer network of human embryonic pancreatic progenitors. *Nat Cell Biol* **17**, 615-626 (2015).
7. Pasquali, L. *et al.* Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet* **46**, 136-143 (2014).
8. Weng, C. *et al.* Single-cell lineage analysis reveals extensive multimodal transcriptional control during directed beta-cell differentiation. *Nat Metab* **2**, 1443-1458 (2020).
9. Song, J.Y. *et al.* HIC2, a new transcription activator of SIRT1. *FEBS Lett* **593**, 1763-1776 (2019).
10. Luu, L. *et al.* The loss of Sirt1 in mouse pancreatic beta cells impairs insulin secretion by disrupting glucose sensing. *Diabetologia* **56**, 2010-20 (2013).
11. Lee, J.H., Lee, J.H. & Rane, S.G. TGF-beta Signaling in Pancreatic Islet beta Cell Development and Function. *Endocrinology* **162**(2021).

## Supplemental Note

### The International Congenital Hyperinsulinism Consortium

Ivo Barić<sup>1</sup>, Liat de Vries<sup>2,3</sup>, Samar S. Hassan<sup>4</sup>, Khadija Nuzhat Humayun<sup>5</sup>, Floris Levy-Khademi<sup>6</sup>, Catarina Limbert<sup>7</sup>, Birgit Rami-Merhar<sup>8</sup>, Verónica Mericq<sup>9</sup>, Kristen Neville<sup>10</sup>, Yasmine Ouarezki<sup>11</sup>, Ana Tangari<sup>12</sup>, Charles Verge<sup>10</sup>, Esko Wiltshire<sup>13</sup>

1. Division for Genetics and Metabolic Diseases, Department of Pediatrics University Hospital Center Zagreb and University of Zagreb, School of Medicine, Kišpatičeva 12, 10000 Zagreb, Croatia
2. The Jesse Z and Sara Lea Shafer Institute for Endocrinology and Diabetes, National Center for Childhood Diabetes, Schneider Children's Medical Center of Israel, Petach Tikva, Israel
3. Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
4. Department of Pediatric Endocrinology, Gaafar Ibn Auf Pediatric Tertiary Hospital, Khartoum, Sudan.
5. Department of Paediatrics and Child Health, Aga Khan University, Karachi, Pakistan
6. Division of Pediatric Endocrinology, Shaare Zedek Medical Center, Faculty of medicine, Hebrew University, Jerusalem, Israel
7. Unit for Paediatric Endocrinology and Diabetes, Centro Hospitalar e Universitário de Lisboa Central, Hospital Dona Estefania, Lisbon, Portugal
8. Medical University of Vienna, Department of Pediatric and Adolescent Medicine, Vienna, Austria
9. Department of Pediatrics, Institute of Maternal and Child Research, Clinica Las Condes, Santiago, Chile
10. Endocrinology, Sydney Children's Hospital, Randwick, Australia and School of Women's and Children's Health, UNSW, Australia
11. EPH Hassan Badi, Algiers, Algeria
12. Saredo, Sanatorio Güemes, Pediatric Endocrinology Unit, Buenos Aires, Argentina
13. Department of Paediatrics and Child Health, University of Otago Wellington, Wellington, New Zealand