Supplementary methods

Exons targeted. The assay was designed to target exons from genes shown in Table S1. For genes which are expressed in multiple isoforms, we targeted those exons contained in transcripts expressed in the pancreas and in which mutations cause monogenic diabetes; the appropriate RefSeq identifier(s) for these transcripts is shown in Table S1. Mutations in some genes may result in either a MODY or neonatal diabetes phenotype depending upon either zygosity, mutation or age at presentation. A greater priority for high, even coverage was assigned to the 20 genes indicated by an asterisk for whom a Sanger sequencing test is available in our laboratory.

gene name	transcript(s) targeted
**ABCC8*	U63421 and L78208
BLK	NM_001715.2
CEL	NM_001807.3
**EIF2AK3*	AF110146.1
**FOXP3*	NM_014009.2
GATA4	NM_002052.3
**GATA6*	NM_005257.3
**GCK*	NM_000162.2
**GLIS3*	NM_001042413.1
** HNF1A*	NM_000545.3
**HNF1B*	NM_000458.1
** HNF4A*	LRG_483
IER3IP1	NM_016097.3
**INS*	NM_000207.2
**IPF1 (PDX1)*	NM_000209.1
**KCNJ11*	NM_000525.3
KLF11	NM_003597.4
**LMNA*	NM_005572.2
m.3243 region	NC_012920.1
**NEUROD1*	NM_002500.2
**NEUROG3*	NM_020999.2
PAX4	NM_006193.2
**PPARG*	NM_005169.3
**PTF1A*	NM_178161.2
**RFX6*	NM_173560.3
**SLC19A2*	NM_006996.2
**SLC2A2 (GLUT2)*	NM_000340.1
WFS1	NM_006005.3
ZFP57	NM_001109809.2

Table S1. Target genes and associated transcripts.

Bait tiling and replication scheme. Base-by-base coverage was calculated for a cohort of nine samples, unrelated to the study reported here, captured using the Agilent SureSelect Human All Exon v1 (38 Mb) system. In this system, all baits are equimolar and designed at a tiling density of one (i.e. baits are non-overlapping), giving a baseline value for the inherent efficiency of capture of target sequences by 120-mer RNA baits in solution. Bases in target genes to be included in the monogenic diabetes panel, all of which were represented in the 38 Mb All Exon system, were scored by coverage depth and each base assigned to a group according to Table S2, below.

average coverage depth	bait group	tiling density	replication on array
>100x	A	2	1
50-100x	B	2	2
20-50x	C	2	4
10-20x	D	4	9
<10x	E	5	9

Table S2. Bait tiling and replication scheme.

Baits across all target regions were then re-designed using the Agilent eArray tool, at a tiling density appropriate to the predominant depth of capture of the target region in the All Exon v1 system. For synthesis of the capture library, bait replication was adjusted as shown in order to make maximum use of the 57,610 features available on the array; to ensure high coverage of the 20 genes currently screened by Sanger sequencing in our laboratory (see Methods), replication was further increased by a factor of 4 for all relevant baits. For capture of the appropriate region of the mitochondrial genome, a single novel bait was designed to target a 120-nt region including the m.3243 position and was included on the synthesis array at the lowest replication. Bait re-balancing resulted in improved uniformity of coverage, compared to that observed in exome samples (Figure S1).