Analyst: **Kumar Veerapen, PhD**
Report Date: **July 15th, 2020, updated on July 20th, 2020**
Replication title: *Replication of findings from a Recessive Genome-Wide Meta-Analysis of Type 2 Diabetes*
Application date: **May 1st, 2020**

SNPs tested:

SNPs (n=36 variants) requested and approved for replication

rsID	SNP Coordinate (GRCh37)	SNP Coordinate (GRCh38, liftover from UCSC Browser)	$AF_{FinnGen}$	Source
rs12031785	1:219686440-219686440	1:219513098-219513098	5.8e-1	Mark
rs1537818	1:39647038-39647038	1:39181366-39181366	7.2e-1	Mark
~~rs147325890~~	~~2:208939268-208939272~~	~~2:208074544-208074548~~	~~NA~~	~~Mark~~
~~rs9826367~~	~~3:12294202-12294202~~	~~3:12252703-12252703~~	~~NA~~	~~Proposal~~
rs114961124	5:162935998-162935998	5:163508992-163508992	5.7e-2	Mark
rs115018790	5:52088271-52088271	5:52792437-52792437	3.0e-2	Proposal
rs140453320	5:64485239-64485239	5:65189412-65189412	9.5e-3	Proposal
~~rs796708168~~	~~6:126689667-126689667~~	~~6:126368521-126368521~~	~~NA~~	~~Proposal~~
rs7773338	6:133100128-133100128	6:132778989-132778989	6.0e-1	Mark
rs11553430	6:32136771-32136771	6:32168994-32168994	2.0e-2	Proposal
rs35484705	6:32583051-32583051	6:32615274-32615274	5.6e-1	Proposal
~~rs4713572~~	~~6:32626952-32626952~~	~~6:32659175-32659175~~	~~NA~~	~~Proposal~~
rs2714337	6:7240577-7240577	6:7240344-7240344	4.0e-1	Proposal
rs972283	7:130466854-130466854	7:130782095-130782095	5.4e-1	Mark
rs143545473	7:28077270-28077270	7:28037651-28037651	1.5e-2	Mark
rs138760676	7:28107505-28107505	7:28067886-28067886	1.5e-2	Proposal
~~rs755900673~~	~~8:2008956-2008956~~	~~8:2060836-2060836~~		~~Proposal~~
rs11311906	8:2008956-2008957	8:2060836-2060837	3.2e-1	Mark
rs139998786	8:79662469-79662469	8:78750234-78750234	2.9e-2	Mark
rs12555274	9:22136440-22136440	9:22136441-22136441	2.8e-1	Mark
rs75518966	9:71045124-71045124	9:68430208-68430208	4.3e-3	Mark
rs7079711	10:114745788-114745788	10:112986029-112986029	2.2e-1	Proposal

rs74810181	10:115080503-115080503	10:113320744-113320744	2.4e-2	Proposal
rs33932777	10:12311465-12311465	10:12269466-12269466	5.2e-1	Proposal
rs12570111	10:12325058-12325058	10:12283059-12283059	4.3e-1	Proposal
rs7912748	10:12658743-12658743	10:12616744-12616744	6.1e-1	Proposal
rs2812533	10:71452285-71452285	10:69692529-69692529	8.1e-1	Proposal
rs4936409	11:117694392-117694392	11:117823677-117823677	3.7e-1	Mark
rs757110	11:17418477-17418477	11:17396930-17396930	5.3e-1	Proposal
rs231903	11:2729946-2729946	11:2708716-2708716	6.1e-1	Mark
~~rs150078842~~	~~11:72818294-72818294~~	~~11:73107249-73107249~~	~~4.1e-3~~	~~Proposal~~
rs2577960	15:42028354-42028354	15:41736156-41736156	6.0e-1	Mark
rs9929462	16:294210-294210	16:244211-244211	2.4e-1	Proposal
rs35944094	17:17400635-17400635	17:17497321-17497321	7.1e-1	Mark
rs3094515	17:36043653-36043653	17:37683650-37683650	4.3e-1	Proposal
rs35226705	19:46301456-46301456	19:45798198-45798198	4.6e-1	Proposal

- *Strikethrough markers were not captured in the FinnGen genotyping and imputation or did not pass QC prior to association analysis*

Endpoints tested:

```
T2D Type 2 diabetes, definitions combined T2D_WIDE | E4_DM2 ! E4_DM1
T2D_WIDE Type 2 diabetes, wide definition !PANCREATITIS E11
T2D_INCLAVO Type 2 diabetes, definitions combined, including avohilmo T2D_WIDE| E4_DM2
!E4_DM1
```

All *_Tx endpoints were filtered to include controls of more than 55 years old. All endpoints are binary.

Endpoint Analysed	Controls (N)	Cases (N)
T2D	182573	29193
T2D_Tx	111033	29193
T2D_WIDE	184778	17268
T2D_WIDE_Tx	110994	17268
T2D_INCLAVO	180722	28717
T2D_INCLAVO_Tx	109462	17268

SAIGE Single variant association test was run on the above binary traits in a recessive model using the FinnGen SAIGE pipeline.

Cromwell submission ID (submitted on July 12th, 2020):
466842b0-cb73-40ab-8abe-b25505a16914

Results were saved here:
gs://fg-cromwell/saige/466842b0-cb73-40ab-8abe-b25505a16914/call-test_combine/shard-*/**/*.gz*

Q-Q plots look fine with < 1.2 genomic inflation values

p-value threshold of < 0.001639785 was applied
Why? 0.05/N(SNPs)
Number of SNPs was 30-31 SNPs

Results processed in Kumar's Google VM: **/home/kumar/workDir/replications/T2D**
Shell script used to pull the subset of SNPs: ./test.sh
Output from files include: **1) all results (aforementioned SNPs above) (*replication.recessive.txt for recessive and *replication.additive.txt for additive), 2) SNPs that survive multiple testing significance threshold (*replication.recessive.sig.txt and *replication.additive.sig.txt for additive)**