






RESEARCH ARTICLE

# Type 2 Diabetes Risk Allele Loci in the Qatari Population

Sarah L. O’Beirne<sup>1,2</sup>, Jacqueline Salit<sup>1</sup>, Juan L. Rodriguez-Flores<sup>1</sup>, Michelle R. Staudt<sup>1</sup>, Charbel Abi Khalil<sup>3\*</sup>, Khalid A. Fakhro<sup>3,4</sup>, Amal Robay<sup>3</sup>, Monica D. Ramstetter<sup>1,5</sup>, Iman K. Al-Azwani<sup>3</sup>, Joel A. Malek<sup>3</sup>, Mahmoud Zirie<sup>6</sup>, Amin Jayyousi<sup>6</sup>, Ramin Badii<sup>7</sup>, Ajayeb Al-Nabet Al-Marri<sup>7</sup>, Maria J. Chiuchio<sup>1</sup>, Alya Al-Shakaki<sup>3</sup>, Omar Chidiac<sup>3</sup>, Maey Gharbiah<sup>3</sup>, Abdulbari Bener<sup>8</sup>, Dora Stadler<sup>9</sup>, Neil R. Hackett<sup>1</sup>, Jason G. Mezey<sup>1,5</sup>, Ronald G. Crystal<sup>1,2</sup>

**1** Department of Genetic Medicine, Weill Cornell Medical College, New York, New York, United States of America, **2** Division of Pulmonary and Critical Care Medicine, Department of Medicine, Weill Cornell Medical College, New York, New York, United States of America, **3** Department of Genetic Medicine, Weill Cornell Medical College-Qatar, Doha, Qatar, **4** Division of Translational Medicine, Sidra Medical Research Centre, Doha, Qatar, **5** Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, New York, United States of America, **6** Department of Medicine, Hamad Medical Corporation, Doha, Qatar, **7** Laboratory Medicine and Pathology, Hamad Medical Corporation, Doha, Qatar, **8** Department of Medical Statistics and Epidemiology, Hamad Medical Corporation, Doha, Qatar, **9** Department of Medicine, Weill Cornell Medical College-Qatar, Doha, Qatar

 These authors contributed equally to this work.

\* [geneticmedicine@med.cornell.edu](mailto:geneticmedicine@med.cornell.edu)



 OPEN ACCESS

**Citation:** O’Beirne SL, Salit J, Rodriguez-Flores JL, Staudt MR, Abi Khalil C, Fakhro KA, et al. (2016) Type 2 Diabetes Risk Allele Loci in the Qatari Population. *PLoS ONE* 11(7): e0156834. doi:10.1371/journal.pone.0156834

**Editor:** Giuseppe Novelli, Tor Vergata University of Rome, ITALY

**Received:** March 22, 2016

**Accepted:** May 22, 2016

**Published:** July 6, 2016

**Copyright:** © 2016 O’Beirne et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All sequence read data was submitted to the Short Read Archive (SRA) section of the NCBI SRA database, and is available through the following link: <http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?study=SRP060765>.

**Funding:** These studies were supported, in part, by The Qatar Foundation and the Weill Cornell Medical College in Qatar and Qatar National Research Foundation NPRP 09-741-3-193, 09-740-3-192 and NPRP 5-436-3-116. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

### Background

The prevalence of type 2 diabetes (T2D) is increasing in the Middle East. However, the genetic risk factors for T2D in the Middle Eastern populations are not known, as the majority of studies of genetic risk for T2D are in Europeans and Asians.

### Methods

All subjects were  $\geq 3$  generation Qataris. Cases with T2D ( $n = 1,124$ ) and controls ( $n = 590$ ) were randomly recruited and assigned to the 3 known Qatari genetic subpopulations [Bedouin (Q1), Persian/South Asian (Q2) and African (Q3)]. Subjects underwent genotyping for 37 single nucleotide polymorphisms (SNPs) in 29 genes known to be associated with T2D in Europeans and/or Asian populations, and an additional 27 tag SNPs related to these susceptibility loci. Pre-study power analysis suggested that with the known incidence of T2D in adult Qataris (22%), the study population size would be sufficient to detect significant differences if the SNPs were risk factors among Qataris, assuming that the odds ratio (OR) for T2D SNPs in Qatari’s is greater than or equal to the SNP with highest known OR in other populations.

### Results

Haplotype analysis demonstrated that Qatari haplotypes in the region of known T2D risk alleles in Q1 and Q2 genetic subpopulations were similar to European haplotypes. After Benjamini-Hochberg adjustment for multiple testing, only two SNPs (rs7903146 and rs4506565), both associated with transcription factor 7-like 2 (TCF7L2), achieved statistical significance in

**Competing Interests:** The authors have declared that no competing interests exist.

the whole study population. When T2D subjects and control subjects were assigned to the known 3 Qatari subpopulations, and analyzed individually and with the Q1 and Q2 genetic subpopulations combined, one of these SNPs (rs4506565) was also significant in the admixed group. No other SNPs associated with T2D in all Qataris or individual genetic subpopulations.

## Conclusions

With the caveats of the power analysis, the European/Asian T2D SNPs do not contribute significantly to the high prevalence of T2D in the Qatari population, suggesting that the genetic risks for T2D are likely different in Qataris compared to Europeans and Asians.

## Introduction

Susceptibility to type 2 diabetes (T2D), a disorder characterized by chronic hyperglycemia in the context of inadequate insulin secretion and peripheral tissue insulin resistance, is influenced by inheritance with superimposed dietary and lifestyle factors [1, 2]. Genome-wide association studies (GWAS) and exome sequencing have identified a large number of genes linked to susceptibility to T2D, helping to understand the pathogenesis of T2D and the genes that influence pancreatic beta-cell function/insulin secretion and insulin resistance [3, 4].

The majority of genetic studies of T2D have been carried out in populations of European descent, populations where the prevalence of T2D in adults is 2.4 to 14.7% [3–7]. In contrast, other than studies of candidate genes, little attention has been focused on the genetic risk factors associated with Middle Eastern populations, despite the growing epidemic of T2D in the Middle East, with an estimated prevalence of 6 to 23.9% in adults [7].

In the context of the high prevalence of T2D in the Middle East, we initiated a study of genetic risk factors of T2D among Qataris, a highly consanguineous population located at the migration crossroads of Africa and Eurasia with an adult prevalence of T2D of 22% [7]. The Qatari population is comprised of 3 genetic subgroups, Bedouin (Q1), Persian/South Asian (Q2) and African (Q3), all of which have a similar, high prevalence of T2D [8, 9]. To assess whether the European and/or Asian risk alleles explained, in part, the high prevalence of T2D in the Qatari population, we evaluated 1,124 Qataris with T2D and 590 Qatari controls with 37 SNPs representing 29 genes for which there are definitive links to an increased risk for T2D in European and Asian populations [3, 6, 10]. Based on analysis of Qatari haplotypes in regions flanking known T2D risk single nucleotide polymorphisms (SNPs), demonstrating similarity of the Q1 and Q2 Qatari subpopulations to European haplotypes, our power analysis predicted that, with this high prevalence of T2D, if these variants contribute to the high prevalence of T2D in Qataris, the European and Asian T2D at risk SNPs should be easily detectable in this Qatari population. Contrary to our hypothesis, we found that risk allele frequencies for almost all of these T2D SNPs were similar in Qatari and European populations, with only 2 of 37 SNPs associated with T2D in the European or Asian populations significantly associated with T2D in a high prevalence Qatari population, suggesting that the genetic risks for T2D in Qataris differ from those of the European and Asian populations.

## Methods

### Ethics Statement

Subjects were recruited from Hamad Medical Corporation (HMC) clinics and written informed consent obtained under protocols approved by the Joint Institutional Review Boards (JIRB) of Hamad Medical College and Weill Cornell Medical College Qatar (WCMC-Q).

## Study Population

A total of 1,714 subjects (1,124 cases with T2D and 590 controls) were enrolled. T2D was diagnosed based on the American Diabetes Association (ADA) criteria, consisting of fasting plasma glucose  $\geq 126$  mg/dL, 2 hr plasma glucose  $\geq 200$  mg/dL during an oral glucose tolerance test, and/or HbA1C  $\geq 6.5\%$  [11].

A questionnaire, physical examination and medical record review were performed and blood collected on each subject. Demographic data including age, date of birth and gender, personal and family history of diabetes was recorded along with self-reported history of gestational diabetes, duration of diabetes and medication use including duration of insulin therapy if applicable. Physical characteristics such as height and weight were collected and body mass index (BMI) calculated. Laboratory tests including HbA1C, and blood glucose levels were performed.

All subjects were over the age of 30 and were a minimum three generations Qatari. Cases were excluded if any of the following were present: history of type 1 diabetes, maturity onset diabetes of the young (MODY), maternally inherited diabetes and deafness syndrome (MIDD), a first degree relative with type 1 diabetes, or secondary diabetes.

## Haplotype Analysis

To test if previously described European and Asian T2D SNPs were potentially relevant to the Qatari population, European, Asian, African, and Qatari haplotypes were compared in the region of the chosen 37 T2D risk alleles [3, 6, 10] and an additional 27 tag SNPs with high minor allele frequency adjacent to these loci (S1 Table). The haplotypes of 100 Qataris who had undergone deep sequencing (detailed below), were inferred and divided into intervals of 2,000 SNPs. With the 1,000 Genomes Project Phase 1 as a reference, haplotype blocks in the regions flanking these 64 SNPs in Europeans, Africans and Asians were compared to Qatari subpopulation haplotype blocks for each of the three major genetic subgroups in the Qatari population: Bedouin (Q1), Persian/South Asian (Q2) and African (Q3) [8, 9]. The population with most similar haplotype was determined using SupportMix [12], by inferring the percentage of Qatari haplotypes assigned to European, Asian, African, or Admixed for each subpopulation.

## Qatari Genetic Subpopulation Genotyping

DNA was extracted from blood using the QIAamp DNA Blood Maxi Kit (Qiagen Sciences Inc, Germantown, MD). The 1,714 subjects were classified into the three genetic subgroups described in the Qatari population [8, 9] using a TaqMan SNP Genotyping Assay (Life Technologies, Carlsbad, CA) for a previously described panel of 48 ancestry informative SNPs [8, 9]. Average genotype call rate was 96% and analyzed in STRUCTURE with  $K = 3$ . Q1, Q2 or Q3 population was assigned if highest proportion was  $>65\%$ ; otherwise they were classed as “admixed” (S1 Fig).

## T2D SNP Genotyping

Allele frequencies were determined for T2D cases ( $n = 1,124$ ) and controls ( $n = 590$ ) for 21 SNPs associated with T2D by GWAS in various populations representing 16 genes, and an additional 27 tag SNPs adjacent to these loci (Table 1 and S1 Table). A subgroup ( $n = 626$  cases, and  $n = 326$  controls) of the total study population underwent genotyping for an additional 16 T2D associated SNPs, representing 15 genes (Table 1 and S1 Table). All genotyping was performed using TaqMan SNP Genotyping Assays (Life Technologies, Carlsbad, CA) as

Table 1. Type 2 Diabetes Risk Allele SNPs.

Category	Gene	SNP	All Qatari			Q1 + Q2 <sup>1</sup>			Q1 <sup>1</sup>			Q2 <sup>1</sup>			Q3 <sup>1</sup>			Admixed		
			p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>
Beta-cell	CAMK1D/ CDC123	rs12779790	5.0x10 <sup>-1</sup>	1.13	0.909– 1.401	2.5x10 <sup>-1</sup>	1.26	0.992– 1.603	1.33	0.972– 1.812	1.16	0.792– 1.685	0.45	0.174– 1.176	9.5x10 <sup>-1</sup>	0.85	0.442– 1.631	9.5x10 <sup>-1</sup>	0.85	0.442– 1.631
dysfunction	CDKAL1	rs10946398	4.3x10 <sup>-1</sup>	1.13	0.931– 1.364	5.5x10 <sup>-1</sup>	1.12	0.905– 1.394	1.01	0.764– 1.340	1.32	0.935– 1.859	1.50	0.713– 3.158	9.5x10 <sup>-1</sup>	1.23	0.704– 2.158	9.5x10 <sup>-1</sup>	1.23	0.704– 2.158
		rs7756992	8.0x10 <sup>-1</sup>	1.06	0.840– 1.327	5.8x10 <sup>-1</sup>	1.13	0.865– 1.469	1.01	0.718– 1.420	1.45	0.919– 2.278	1.14	0.485– 2.672	9.5x10 <sup>-1</sup>	0.71	0.355– 1.411	9.5x10 <sup>-1</sup>	0.71	0.355– 1.411
	CDKN2A-B	rs564398	7.6x10 <sup>-1</sup>	0.94	0.772– 1.134	5.6x10 <sup>-1</sup>	0.90	0.727– 1.107	0.87	0.651– 1.151	0.94	0.689– 1.293	1.33	0.418– 4.224	9.5x10 <sup>-1</sup>	1.08	0.599– 1.961	9.5x10 <sup>-1</sup>	1.08	0.599– 1.961
		rs10811661	1.9x10 <sup>-1</sup>	0.81	0.651– 1.002	3.3x10 <sup>-1</sup>	0.82	0.645– 1.039	0.82	0.595– 1.125	0.82	0.567– 1.176	0.18	0.049– 0.683	9.5x10 <sup>-1</sup>	0.98	0.531– 1.804	9.5x10 <sup>-1</sup>	0.98	0.531– 1.804
	G6PC2	rs560887	9.7x10 <sup>-1</sup>	1.01	0.752– 1.305	6.7x10 <sup>-1</sup>	0.90	0.819– 1.515	0.89	0.767– 1.640	1.11	0.510– 1.596	0.79	0.358– 4.474	9.5x10 <sup>-1</sup>	1.67	0.246– 1.440	9.5x10 <sup>-1</sup>	1.67	0.246– 1.440
	HHEX/IDE/ KIF11	rs1111875	1.8x10 <sup>-1</sup>	0.82	0.685– 0.979	1.8x10 <sup>-1</sup>	0.79	0.649– 0.968	0.94	0.719– 1.232	0.63	0.465– 0.859	1.29	0.614– 2.725	9.5x10 <sup>-1</sup>	0.75	0.443– 1.269	9.5x10 <sup>-1</sup>	0.75	0.443– 1.269
	HNF1A	rs7957197	2.7x10 <sup>-1</sup>	0.80	0.611– 1.058	4.0x10 <sup>-1</sup>	0.81	0.589– 1.101	0.83	0.550– 1.240	0.85	0.505– 1.427	1.72	0.488– 6.040	9.5x10 <sup>-1</sup>	0.67	0.294– 1.516	9.5x10 <sup>-1</sup>	0.67	0.294– 1.516
	HNF1B (TCF2)	rs4430796	1.8x10 <sup>-1</sup>	0.80	0.649– 0.991	1.8x10 <sup>-1</sup>	0.75	0.595– 0.953	0.71	0.521– 0.962	0.83	0.567– 1.223	1.21	0.422– 3.445	9.5x10 <sup>-1</sup>	0.84	0.430– 1.639	9.5x10 <sup>-1</sup>	0.84	0.430– 1.639
	JAZF1	rs864745	1.4x10 <sup>-1</sup>	0.81	0.686– 0.957	2.4x10 <sup>-1</sup>	0.83	0.691– 0.999	0.96	0.755– 1.230	0.68	0.514– 0.909	0.69	0.308– 1.546	9.5x10 <sup>-1</sup>	0.92	0.560– 1.508	9.5x10 <sup>-1</sup>	0.92	0.560– 1.508
	KCNU11	rs5215	8.1x10 <sup>-1</sup>	1.04	0.853– 1.258	7.5x10 <sup>-1</sup>	1.04	0.841– 1.293	1.07	0.794– 1.447	1.04	0.762– 1.431	0.60	0.173– 2.061	9.5x10 <sup>-1</sup>	1.05	0.615– 1.779	9.5x10 <sup>-1</sup>	1.05	0.615– 1.779
	MADD	rs7944584	8.0x10 <sup>-1</sup>	0.94	0.724– 1.224	7.5x10 <sup>-1</sup>	1.06	0.792– 1.421	1.37	0.934– 2.008	0.75	0.463– 1.223	0.23	0.045– 1.220	5.9x10 <sup>-1</sup>	0.52	0.249– 1.089	5.9x10 <sup>-1</sup>	0.52	0.249– 1.089
	NOTCH2	rs10923931	7.9x10 <sup>-1</sup>	1.08	0.825– 1.399	5.8x10 <sup>-1</sup>	1.16	0.846– 1.597	1.01	0.846– 1.563	1.35	0.847– 2.143	0.96	0.453– 2.050	9.5x10 <sup>-1</sup>	1.06	0.506– 2.222	9.5x10 <sup>-1</sup>	1.06	0.506– 2.222
	THADA	rs7578597	7.6x10 <sup>-1</sup>	0.90	0.667– 1.207	5.6x10 <sup>-1</sup>	0.83	0.570– 1.197	0.87	0.465– 1.620	0.84	0.527– 1.325	1.50	0.693– 3.244	9.5x10 <sup>-1</sup>	0.95	0.462– 1.952	9.5x10 <sup>-1</sup>	0.95	0.462– 1.952
Reduced insulin secretion	ADRA2A	rs10885122	2.4x10 <sup>-1</sup>	0.82	0.642– 1.036	4.0x10 <sup>-1</sup>	0.82	0.620– 1.089	0.73	0.514– 1.042	1.00	0.609– 1.640	0.96	0.415– 2.232	9.5x10 <sup>-1</sup>	0.68	0.345– 1.327	9.5x10 <sup>-1</sup>	0.68	0.345– 1.327
	FADS1	rs174550	1.9x10 <sup>-1</sup>	1.26	0.998– 1.585	1.8x10 <sup>-1</sup>	1.33	1.026– 1.710	1.28	0.935– 1.757	1.44	0.918– 2.262	0.94	0.214– 4.097	9.5x10 <sup>-1</sup>	1.11	0.558– 2.220	9.5x10 <sup>-1</sup>	1.11	0.558– 2.220
	GCK	rs1799884	9.4x10 <sup>-1</sup>	1.02	0.802– 1.287	7.1x10 <sup>-1</sup>	0.93	0.715– 1.214	1.04	0.742– 1.444	0.74	0.461– 1.182	1.06	0.358– 3.150	9.5x10 <sup>-1</sup>	1.35	0.642– 2.840	9.5x10 <sup>-1</sup>	1.35	0.642– 2.840
		rs4607517	9.8x10 <sup>-1</sup>	1.00	0.786– 1.265	6.4x10 <sup>-1</sup>	0.91	0.695– 1.179	1.00	0.720– 1.395	0.72	0.450– 1.151	1.61	0.434– 5.992	9.5x10 <sup>-1</sup>	1.32	0.616– 2.808	9.5x10 <sup>-1</sup>	1.32	0.616– 2.808
	GLIS3	rs7034200	2.7x10 <sup>-1</sup>	0.83	0.668– 1.042	5.8x10 <sup>-1</sup>	0.89	0.689– 1.141	0.81	0.585– 1.116	1.11	0.726– 1.697	0.59	0.209– 1.691	6.3x10 <sup>-1</sup>	0.58	0.303– 1.107	6.3x10 <sup>-1</sup>	0.58	0.303– 1.107
	LGR5/ TSPAN8	rs7961581	6.5x10 <sup>-1</sup>	1.07	0.913– 1.258	5.8x10 <sup>-1</sup>	1.09	0.909– 1.296	1.19	0.943– 1.507	0.97	0.740– 1.283	1.33	0.614– 2.876	9.5x10 <sup>-1</sup>	1.02	0.633– 1.632	9.5x10 <sup>-1</sup>	1.02	0.633– 1.632
	PROX1	rs340874	8.1x10 <sup>-1</sup>	1.04	0.843– 1.281	5.8x10 <sup>-1</sup>	1.11	0.877– 1.414	1.23	0.901– 1.669	1.11	0.736– 1.665	0.71	0.233– 2.164	9.5x10 <sup>-1</sup>	0.82	0.470– 1.416	9.5x10 <sup>-1</sup>	0.82	0.470– 1.416
	SLC2A2 (GLUT2)	rs11920090	3.3x10 <sup>-1</sup>	1.24	0.922– 1.657	3.5x10 <sup>-1</sup>	1.31	0.933– 1.829	1.40	0.909– 2.165	1.11	0.651– 1.891	2.41	0.719– 8.059	9.3x10 <sup>-1</sup>	0.54	0.220– 1.348	9.3x10 <sup>-1</sup>	0.54	0.220– 1.348
	SLC30A8	rs11558471	2.7x10 <sup>-1</sup>	0.78	0.581– 1.058	6.8x10 <sup>-1</sup>	0.89	0.637– 1.252	1.02	0.643– 1.632	0.80	0.470– 1.367	0.50	0.133– 1.902	5.9x10 <sup>-1</sup>	0.43	0.171– 1.096	5.9x10 <sup>-1</sup>	0.43	0.171– 1.096
		rs13266634	9.6x10 <sup>-1</sup>	1.02	0.738– 1.396	6.9x10 <sup>-1</sup>	1.11	0.769– 1.614	1.07	0.590– 1.942	1.20	0.737– 1.936	0.21	0.044– 1.005	9.5x10 <sup>-1</sup>	1.06	0.488– 2.298	9.5x10 <sup>-1</sup>	1.06	0.488– 2.298

(Continued)

Table 1. (Continued)

Category	Gene	SNP	All Qatari			Q1 + Q2 <sup>1</sup>			Q1 <sup>1</sup>			Q2 <sup>1</sup>			Q3 <sup>1</sup>			Admixed		
			p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>	p value <sup>2</sup>	OR <sup>3</sup>	CI <sup>4</sup>
	TCF7L2	rs4506565	3.7x10 <sup>-2</sup>	1.33	1.121–1.586	1.8x10 <sup>-1</sup>	1.24	1.023–1.507	5.5x10 <sup>-1</sup>	1.18	0.916–1.515	4.8x10 <sup>-1</sup>	1.34	0.986–1.820	9.7x10 <sup>-1</sup>	1.35	0.637–2.875	4.5x10 <sup>-2</sup>	2.43	1.409–4.195
		rs7901695	1.8x10 <sup>-1</sup>	1.20	1.016–1.414	3.6x10 <sup>-1</sup>	1.15	0.958–1.389	7.6x10 <sup>-1</sup>	1.11	0.866–1.412	6.8x10 <sup>-1</sup>	1.22	0.908–1.620	9.7x10 <sup>-1</sup>	0.95	0.473–1.916	1.2x10 <sup>-1</sup>	1.90	1.155–3.131
		rs7903146	2.9x10 <sup>-2</sup>	1.36	1.146–1.618	1.8x10 <sup>-1</sup>	1.28	1.057–1.551	5.5x10 <sup>-1</sup>	1.20	0.932–1.537	3.4x10 <sup>-1</sup>	1.40	1.034–1.890	9.7x10 <sup>-1</sup>	1.18	0.503–2.779	7.4x10 <sup>-2</sup>	2.17	1.291–3.649
Pancreatic	HHEX	rs5015480	1.9x10 <sup>-1</sup>	0.86	0.731–1.007	1.8x10 <sup>-1</sup>	0.82	0.685–0.981	8.0x10 <sup>-1</sup>	0.92	0.728–1.167	2.1x10 <sup>-1</sup>	0.69	0.521–0.915	9.7x10 <sup>-1</sup>	1.07	0.541–2.116	9.5x10 <sup>-1</sup>	1.90	1.155–3.131
development	IGF2BP2	rs4402960	4.8x10 <sup>-1</sup>	1.11	0.933–1.312	2.6x10 <sup>-1</sup>	1.20	0.985–1.451	5.0x10 <sup>-1</sup>	1.32	1.025–1.711	9.4x10 <sup>-1</sup>	1.04	0.769–1.414	9.7x10 <sup>-1</sup>	0.99	0.482–2.030	9.5x10 <sup>-1</sup>	0.83	0.523–1.328
Insulin	GCKR	rs780094	1.8x10 <sup>-1</sup>	1.25	0.647–0.993	4.0x10 <sup>-1</sup>	1.18	0.671–1.078	5.5x10 <sup>-1</sup>	1.32	0.563–1.024	9.1x10 <sup>-1</sup>	0.91	0.725–1.655	9.7x10 <sup>-1</sup>	2.22	0.156–1.316	9.5x10 <sup>-1</sup>	1.19	0.424–1.657
resistance	KLF14	rs972283	9.4x10 <sup>-1</sup>	1.02	0.808–1.278	6.8x10 <sup>-1</sup>	0.92	0.705–1.193	9.0x10 <sup>-1</sup>	0.93	0.666–1.298	8.6x10 <sup>-1</sup>	0.87	0.549–1.368	9.7x10 <sup>-1</sup>	1.77	0.629–5.005	9.5x10 <sup>-1</sup>	1.26	0.639–2.469
	PPARG	rs1801282	8.3x10 <sup>-1</sup>	0.96	0.685–1.341	7.3x10 <sup>-1</sup>	0.93	0.640–1.340	9.9x10 <sup>-1</sup>	1.04	0.604–1.774	7.8x10 <sup>-1</sup>	0.78	0.488–1.360	9.7x10 <sup>-1</sup>	2.41	0.255–22.710	9.5x10 <sup>-1</sup>	0.86	0.331–2.223
Possible insulin resistance	ADAMTS9	rs4607103	7.9x10 <sup>-1</sup>	1.05	0.892–1.243	5.8x10 <sup>-1</sup>	1.08	0.899–1.306	9.0x10 <sup>-1</sup>	1.05	0.827–1.344	7.9x10 <sup>-1</sup>	1.14	0.845–1.525	9.7x10 <sup>-1</sup>	1.12	0.601–2.100	9.5x10 <sup>-1</sup>	1.09	0.831–2.313
Obesity	FTO <sup>5</sup>	rs11642841	2.4x10 <sup>-1</sup>	1.16	0.981–1.363	2.5x10 <sup>-1</sup>	1.18	0.989–1.417	6.3x10 <sup>-1</sup>	1.14	0.897–1.436	4.8x10 <sup>-1</sup>	1.28	0.962–1.695	9.7x10 <sup>-1</sup>	0.61	0.260–1.447	9.5x10 <sup>-1</sup>	1.30	0.790–2.136
		rs8050136	8.1x10 <sup>-1</sup>	1.04	0.881–1.218	7.5x10 <sup>-1</sup>	1.04	0.864–1.240	9.2x10 <sup>-1</sup>	0.96	0.751–1.219	9.7x10 <sup>-1</sup>	1.02	0.776–1.345	9.7x10 <sup>-1</sup>	1.25	0.649–2.390	9.5x10 <sup>-1</sup>	1.11	0.688–1.790
		rs9939609	8.1x10 <sup>-1</sup>	1.03	0.878–1.216	6.9x10 <sup>-1</sup>	0.95	0.789–1.139	9.6x10 <sup>-1</sup>	0.97	0.761–1.244	8.6x10 <sup>-1</sup>	1.09	0.824–1.443	9.7x10 <sup>-1</sup>	0.95	0.502–1.814	9.5x10 <sup>-1</sup>	1.10	0.683–1.760
Circadian rhythm	CRY2	rs11605924	5.0x10 <sup>-1</sup>	0.89	0.720–1.098	4.8x10 <sup>-1</sup>	0.87	0.682–1.101	5.5x10 <sup>-1</sup>	0.73	0.536–0.993	7.8x10 <sup>-1</sup>	1.25	0.834–1.873	9.7x10 <sup>-1</sup>	1.27	0.465–3.484	9.5x10 <sup>-1</sup>	1.11	0.605–2.018

<sup>1</sup> Q1 = Bedouin, Q2 = Persian/South Asian, Q3 = African, Admixed = Structure cut-off <0.65 in all sub-populations (k = 3).

<sup>2</sup> Benjamini-Hochberg corrected p-value.

<sup>3</sup> Estimated odds ratio (OR).

<sup>4</sup> Lower and upper bounds of 95% confidence interval (CI) for odds ratio.

<sup>5</sup> No BMI covariate used for FTO gene.

doi:10.1371/journal.pone.0156834.t001

per manufacturer's protocols. The 29 genes were classified into those associated with insulin resistance or insulin secretion [3, 6, 10] (S1 Table).

## Deep Sequencing of Qatari Genomes

One hundred Q1, Q2, Q3 Qatari genomes were sequenced to a median depth of 37x (minimum 30x) using Illumina reads on a HiSeq 2000 (Illumina Inc., San Diego, CA), as previously described [13]. Reads were mapped to the 1,000 Genomes Project version of the hg19/GRCh37 human reference genome using BWA 0.5.9 [14]. Sufficient paired-end 100 bp reads were generated in order to produce a median of 112 GB of sequence data passing filters and aligned to the hg19/GRCh37 human reference genome with a median insert size of 301 bp, where at least 85% of bases with quality score  $\geq 30$  (Q30) passed filtering steps and were aligned. Among non-N bases in the reference genome, at least 98% were covered by at least one base in all 100 genomes. Mapped reads were prepared for genotyping using the "Best practices for variant detection v3" GATK (<http://www.broadinstitute.org/gatk/>) pipeline, including removal of PCR-duplicate reads, realignment across known indels, and base quality score recalibration [15]. The number of exome bases covered at  $\geq 1x$  and  $\geq 10x$  was determined using SAMtools [16].

## Statistical Analysis

The PLINK software package (v1.07; <http://pngu.mgh.harvard.edu/purcell/plink/>) was used for statistical analyses [17]. First a pairwise similarity was calculated between all subjects using the 48 ancestry informative genotypes [8, 9], then a multi-dimensional scaling (MDS) algorithm was used to generate ethnic covariates for the case vs control logistic regression. To establish the set of MDS variables that achieved optimal separation between the sub-populations of Qataris we used the Akaike Information Criterion (AIC) as a means for model selection. Age, gender, BMI and T2D phenotypes were assigned as fixed effects and multi-dimensional scaling factors were sequentially added to the model. We found that the AIC scores converged once five or more multi-dimensional factors were introduced to the model.

The logistic regression analysis was performed using PLINK with and without BMI for each SNP individually under the additive genetic model. Other covariates included age, gender, and the first 5 MDS ethnic covariates as described above. SupportMix, a machine learning method for admixture analysis [12], was utilized to determine the population categorization (African, European or Asian) of Qatari haplotypes in regions flanking tag SNPs.

## Power Calculation

Genetic power calculations for 37 SNPs previously linked to T2D were conducted using the Genetic Power Calculator (GPC) [18]. The GPC analysis for case-control study of discrete traits takes as input the high risk allele frequency (based on TaqMan genotyping of Qataris), the disease prevalence (22% in Qatar), the genotype relative risk for heterozygotes [approximately equal to the odds ratio (OR)], the genotype relative risk for homozygotes (approximately the square of the OR), the D-prime linkage disequilibrium between the marker and risk variants (assumed to be 1.0), the marker allele frequency (risk allele frequency), number of cases, and the control:case ratio. The OR for the known T2D associated risk alleles ranged from 1.01 to 2.17 in other populations [19, 20], and prior studies of ethnic variance in OR suggests that the directionality is expected to be the same in Qataris [21, 22]. In the analysis, two scenarios were considered, one where the OR in Qatar is the same as in population of the reported SNP, and one where the OR in Qatar is equal to the highest reported OR among the 37 SNPs (2.71). In the first scenario, the n of cases needed for 80% power was calculated and



compared to our study, and in the second scenario the power given our sample size was calculated (S2 Table). Additionally, for each SNP, the minimum OR needed to detect an association with >80% power, given the sample size was calculated using a range of OR between 1.01 and 2.17, in 0.01 step intervals. The lowest OR with >80% power was reported (S2 Table). The minimum OR ranged from 1.18–1.44, with a mean of 1.25. The low end of this range (1.18), matched the mean OR of the reported SNPs (also 1.18), hence it was assumed that the OR could be 1.18 or higher in Qatar, as high as 2.17.

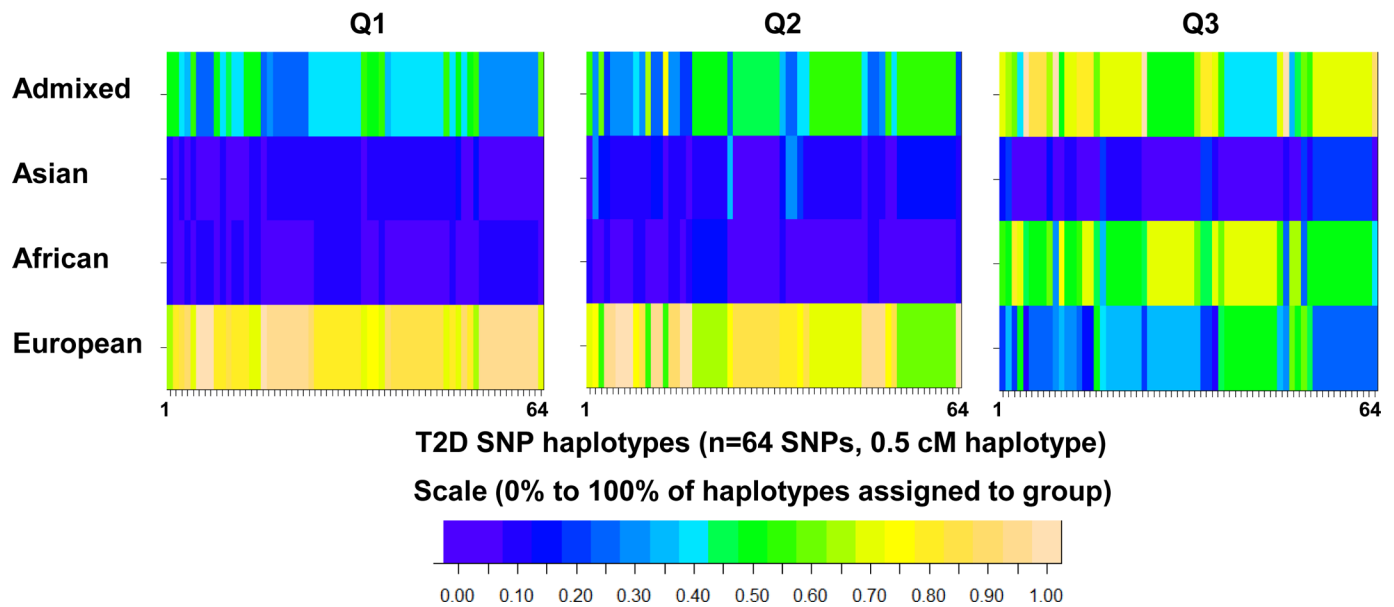
## Results

A total of 1,714 subjects were recruited from diabetes clinics and other departments in Hamad Medical Corporation in Doha, Qatar (S3 Table). All were self-reported third generation native Qatari over the age of thirty. Of these, 1,124 were classified as cases of T2D based on ADA criteria for fasting plasma glucose, oral glucose tolerance test or HbA1C level, without evidence of type 1 diabetes, MODY or MIDD as detailed in Methods. Forty percent (235/590) of control subjects had at least one first degree relative with TD2. Blood glucose and/or HbA1C data was available for 77% (455/590) of control subjects, providing information about their diabetes “at risk” status. Based on this, 53% (181/336) of control subjects were at risk with a HbA1C of 5.7–6.4%, and 25% (71/284) were at risk based on a random blood glucose level >5.6 mmol/L. One-hundred and sixty-six control subjects had both blood glucose and HbA1C levels measured, and of these, 29 (17%) were at risk by both criteria. The majority of the subjects were female both for cases with T2D (677/1,124; 60%) and controls (377/590; 64%;  $p = 0.14$ ). T2D cases were significantly older with a mean age of  $55 \pm 10$  yr vs  $46 \pm 9$  yr in controls ( $p < 10^{-10}$ ). BMI was significantly higher in T2D cases ( $34 \pm 7$  kg/m<sup>2</sup>) compared to controls ( $32 \pm 7$  kg/m<sup>2</sup>;  $p = 1.4 \times 10^{-6}$ ). HbA1C levels were also significantly higher in T2D cases ( $8.3 \pm 1.9$ ) compared to controls ( $5.6 \pm 0.4$ ;  $p < 10^{-10}$ ).

For all subjects, DNA was extracted from blood and Qatari subpopulation was determined by the use of 48 ancestry informative SNPs [8, 9]. Forty-five % of subjects were assigned as Q1 (Bedouin), 34% as Q2 (Persian, South Asian), and 8% as Q3 (African), leaving 13% classified as admixed (S1 Fig). The proportion of each subpopulation in T2D cases and controls was similar ( $p = 9.4 \times 10^{-1}$ ).

We then chose 37 index SNPs known to be associated with risk for T2D based on GWAS studies performed in Europeans and South Asians [3, 6, 10] (Table 1). Of the 37 SNPs selected for genotyping in Qataris, the majority were associated with T2D in previous GWAS, with many of the susceptibility loci replicated in subsequent independent studies and meta-analyses. As the genetic risks for T2D in Qataris are unknown, SNPs were chosen from GWAS performed in both European and Asian populations. An additional 27 tag SNPs with high minor allele frequency adjacent to these loci were also chosen to obtain more information on Qatari haplotypes (S1 Table). Qatari haplotypes in the region of these SNPs were compared to the haplotypes of other populations to ensure that these SNPs were potentially relevant to Qataris prior to genotyping a large number of subjects. Using the 1000 Genomes Project Phase 1 as a reference, haplotype blocks in the regions flanking these 64 SNPs in Europeans, Africans and Asians were compared to Qatari subpopulation haplotype blocks. This analysis revealed that the T2D haplotypes of Qataris from Q1 and Q2 subpopulations were similar to European T2D haplotypes in (Fig 1). In contrast, the majority of T2D haplotypes in the Q3 subpopulation were categorized as African.

Given the high prevalence of T2D in Qatar, and the similarities in T2D haplotypes between Qataris and Europeans, we genotyped 1,714 subjects (1,124 cases and 590 controls) for these established T2D risk alleles. Following quality control analysis, the risk allele frequency (RAF)



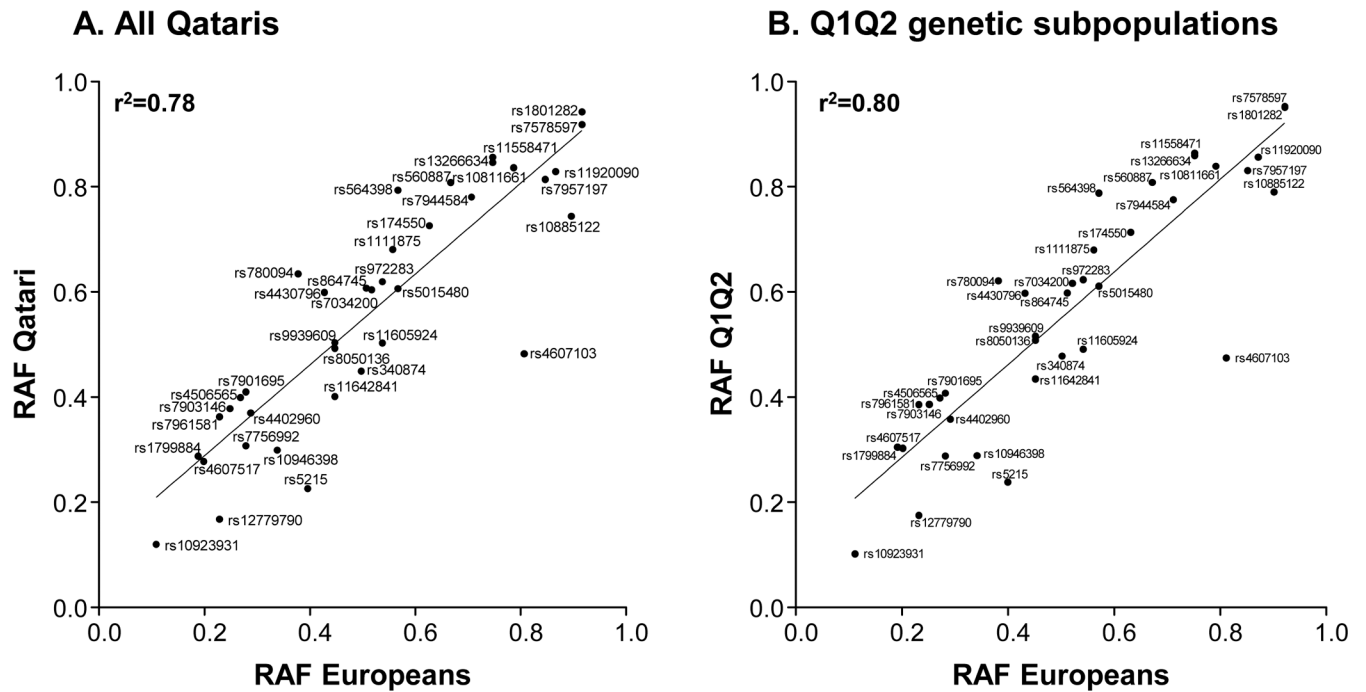
**Fig 1. Comparison of Qatari haplotypes in the regions flanking known single nucleotide polymorphisms (SNPs) associated with type 2 diabetes (T2D) or nearby tag SNPs to the relevant haplotypes of European, Asian, African, or Admixed 1000 Genomes populations.** Admixture deconvolution was used to determine if Qatari haplotypes flanking 64 ‘T2D SNPs’ matched the populations where these SNPs were discovered. Haplotypes were inferred for 100 deeply sequenced Qatari genomes (n = 60 Bedouin Q1, n = 20 Persian Q2, n = 20 African Q3), and divided into 2000 SNP (0.5 cM) intervals. For each interval, SupportMix [12] was used to infer the population with the most similar haplotype, using 1000 Genomes Project Phase 1 as a reference (Admixed = ASW, PUR, CLM, MXL; Asian = CHB, CHS, JPT; African = LWK, YRI; European = TSI, IBS, CEU, FIN, GBR). For each interval containing a T2D risk SNP (n = 64, S1 Table), the percentage of Qatari haplotypes assigned to European, Asian, African, or Admixed was determined. Shown is a heatmap of the results, where each column represents a SNP, and each row represents a 1000 Genomes group. Scaled from blue (0%) to tan (100%), colors represent the % of haplotypes for the SNP that are most similar to each Qatari population (Q1, Q2, or Q3).

doi:10.1371/journal.pone.0156834.g001

of each SNP was determined. Then, the RAF of 37 T2D associated SNPs in Qataris were compared to those of Europeans (obtained from HapMap; Fig 2A) [23]. There was good overall correlation between Qatari and European RAF (Pearson correlation coefficient 0.89,  $r^2 = 0.78$ ,  $p < 10^{-4}$ ), although the RAF was higher in Qataris than Europeans for 70% of SNPs (26/37). Of the 37 SNPs with known risk alleles, the SNP with the highest RAF in all Qataris was rs1801282 in PPAR $\gamma$  (peroxisome proliferator-activated receptor gamma) with a RAF of 0.94. SNP rs7903146 in transcription factor 7-like 2 (TCF7L2), the SNP most strongly associated with T2D in other populations [24], had a RAF in all Qataris of 0.37, higher than the RAF observed in Europeans (0.25). Since the haplotype analysis revealed that the Q1 and Q2 genetic subpopulations were most similar to Europeans in the region of T2D associated SNPs, the RAF frequencies in these subpopulations were combined and compared to Europeans. The combined Q1 and Q2 RAF were in closer agreement with European RAF for T2D associated SNPs (Pearson correlation coefficient 0.89,  $r^2 = 0.8$ ,  $p < 10^{-4}$ , Fig 2B) than were all Qataris. When the distribution of risk alleles in 952 Qatari cases and controls (of all genetic subpopulations) who had undergone genotyping for all 37 T2D associated SNPs was examined, 47% of all subjects had  $\geq 40$  T2D associated alleles (Fig 3). There was no significant difference between the number of risk alleles in cases compared to controls ( $p > 0.2$ ).

Assuming the OR in Qatar is equal to the highest reported OR (2.17) among the 37 SNPs with known risk alleles, given our sample size, there was  $>80\%$  power to detect an association for a given SNP, with a type I error rate of 0.05 (S2 Table). The difference in allele frequencies between T2D cases and controls was then determined (Table 1). Age, BMI and gender were





**Fig 2. Population risk allele frequencies.** Population risk allele frequencies (RAF) in **A.** all Qataris, and **B.** Q1 and Q2 genetic subpopulations combined, compared to European RAF (obtained from Hapmap [23]) for 37 SNPs previously associated with T2D, with the straight line indicating the regression line of best fit of the data.

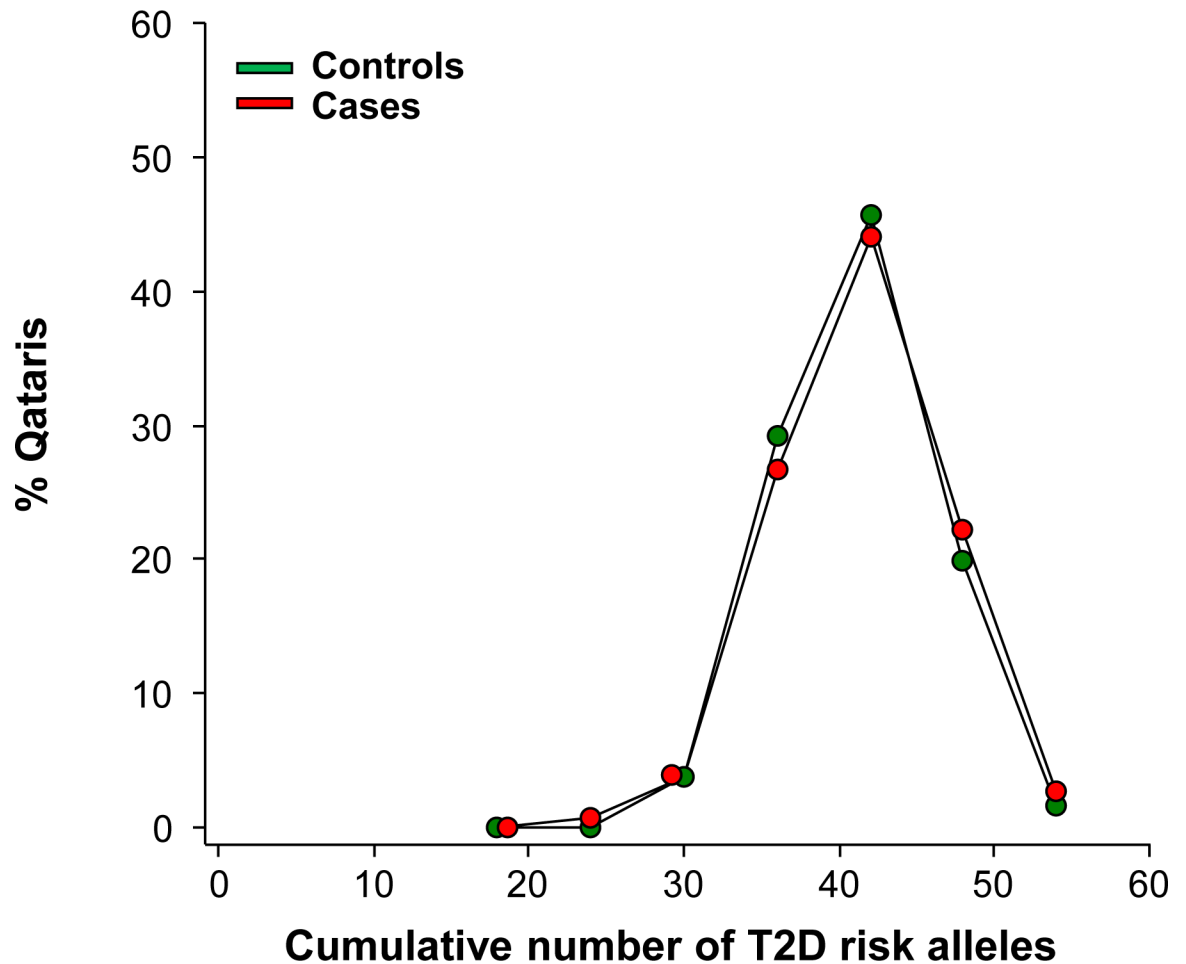
doi:10.1371/journal.pone.0156834.g002

introduced as covariates in the analysis since they differed significantly between cases and controls (S3 Table), as well as the first 5 MDS ethnic covariates. As the Q1 and Q2 genetic subpopulations were most similar to Europeans, in addition to analyzing each subpopulation individually, these subpopulations were also combined and analyzed. However, despite the high RAF observed in Qataris, after Benjamini-Hochberg (BH) adjustment for multiple testing, only two SNPs, both in transcription factor 7-like 2 (TCF7L2) achieved statistical significance in the whole study population, rs7903146 ( $p = 0.028$ ) and rs4506565 ( $p = 0.036$ ). When T2D subjects and control subjects were assigned to the Qatari subpopulations, one of these SNPs (rs4506565) was also significant in the admixed group ( $p = 0.045$ ; Table 1).

Twenty-seven tag SNPs adjacent to the known T2D associated SNPs were also genotyped in all Qataris, and in a subgroup of subjects, genotyping was performed for the additional 16 SNPs in 15 genes associated with T2D in Europeans and South Asians (S1 Table). For these SNPs, a similar analysis was conducted comparing allele frequencies between cases and controls for all Qataris and individual genetic subpopulations, including Q1 and Q2 combined. Significance was not observed in any additional SNPs among the Qatari population as a whole, although the SNP rs4074718 in TCF7L2 ( $p < 0.05$ ) was significant in the subpopulation defined as admixed (S1 Table). For all other SNPs confidence intervals for ORs gave no indication that these SNPs contribute to diabetes risk in Qataris (S1 Table).

## Discussion

Qatar has one of the highest incidences of T2D in the world [7]. Despite sharing similar haplotypes and risk allele frequency (RAF) with Europeans, the T2D SNPs commonly associated with T2D in Europeans and Asians do not help explain the very high prevalence of T2D in



**Fig 3. Cumulative distribution of 37 Type 2 Diabetes (T2D) risk allele counts.** Cumulative distribution of 37 T2D risk allele counts in Qataris (n = 952) cases (red, n = 626), and controls (green, n = 326). Two-tailed, unequal variance t-test, p = 0.26.

doi:10.1371/journal.pone.0156834.g003

Qataris. This observation emphasizes the importance of population-specific genetic variation in the pathogenesis of common disorders such as T2D.

### T2D Risk Alleles in Europeans and Asians

To date, GWAS have identified over 90 loci associated with T2D [3–6, 10], helping to understand the pathogenesis of T2D, identifying novel metabolic pathways and potential targets for therapy [5]. The SNP that is most strongly associated with T2D is rs7903146 in TCF7L2, a transcription factor related to insulin and proglucagon genes [3, 24–26]. The association between TCF7L2 and T2D has been reproduced in various ethnic groups [24], including the Qataris in this study. However, despite the robust association between TCF7L2 and T2D, the causal variant has not yet been identified [27]. Other loci strongly associated with T2D include hematopoietically expressed homeobox (HHEX), peroxisome proliferator-activated receptor gamma (PPARG), solute carrier family 30 (zinc transporter) member 8 (SCL30A8), CDK5 regulatory subunit associated protein 1-like 1 (CDKAL1), and fat mass and obesity-associated protein (FTO) [3–6, 10]. Most of these GWAS have been carried out in European and Asian populations, and given the significant differences in genetic architecture in different populations [28, 29], the loci identified may not be transferable to other populations.

## T2D Risk Alleles in Middle Eastern Populations

There have been no prior GWAS of T2D risk alleles in Middle Eastern populations, but there have been several candidate genes studies, albeit with inconsistent results. While it has been reported that many of the known European risk alleles are associated with T2D in Lebanese Arabs [30], the Saudi Arabian population [31], and in Tunisians and Moroccans with cumulative effects [32], other studies have not demonstrated an association between the European T2D risk SNPs in these populations. For example, though a link between SNPs in TCF7L2 and T2D has been reported in Moroccans [32], only a marginal association has been found in Arabs [33], and several European SNPs were not associated with T2D in Tunisians [34]. Similarly, only 2 of 23 loci associated with BMI in other populations have been linked to obesity in Qataris [35].

Consistent with these observations, despite demonstrating that T2D haplotypes of Q1 and Q2 Qataris are more similar to European than African and Asian haplotypes, we found that only two SNPs in TCF7L2, the susceptibility gene most strongly associated with T2D in Europeans, were associated with T2D in all Qataris. This may be explained by variations in patterns of linkage disequilibrium (LD) between SNPs and susceptibility loci, whereby tag SNPs may be transferable between populations as demonstrated by the similarity between Q1 and Q2 Qatari and European T2D haplotypes, while the haplotype blocks may still vary sufficiently so that the causal variants are not in LD with tag SNPs [36, 37].

Diabetes results from complex interactions between genetic, environmental and behavioral factors [1, 2]. Most of the T2D susceptibility genes identified by GWAS to date relate to beta-cell dysfunction rather than insulin resistance [3–6, 10]. In contrast to many of the subjects included in prior GWAS studies, Qataris have been subject to rapidly changing environmental factors including Westernization of diet, and an increasingly sedentary lifestyle, with consequent high rates of obesity [38, 39]. There may be population specific environmental factors associated with disease risk or protection [40], and perhaps insulin resistance pathways may play a greater role in the development of T2D in these subjects than other populations.

## Population Risk Allele Frequencies

There was significant correlation between the observed RAF in Qataris and Europeans, consistent with previous reports of a directionally similar association to T2D between populations [21, 41, 42]. Despite this finding, only 2 of 37 known T2D risk alleles were associated with T2D in Qataris, suggesting they are not responsible for the condition in this population. This finding highlights the importance of population-specific variation in the pathogenesis of common condition such as T2D, and the need for genetic studies in diverse populations.

GWAS identifies common variants with a low OR [43], and it is believed that multiple SNPs in an individual are required for the pathogenesis of complex conditions such as T2D [44]. The risk allele frequency in Qataris for the 37 SNPs associated with T2D ranged from 0.12 to 0.94 (mean 0.55), indicating that the risk variants were at a high frequency within both cases and controls. In fact, 47% of subjects carried  $\geq 40$  risk alleles. Additionally, the risk variants were individually associated with a low OR (maximum OR 1.36, TCF7L2 rs7903146), with each SNP conferring only a small risk of T2D. As a large number of the subjects carry multiple risk alleles without developing T2D, this low penetrance suggests that each variant had only a mild effect, and that other factors such as diet and lifestyle are required for development of T2D.

## Supporting Information

**S1 Fig. Scatter plot of multidimensional scaling (MDS) illustrating subject distribution by Qatari genetic subpopulation.**

(PDF)

**S1 References. References Used in Supplemental Data.**  
(PDF)

**S1 Table. Type 2 Diabetes Risk Allele SNPs.**  
(PDF)

**S2 Table. Statistical Power to Detect Associations with Type 2 Diabetes for Known Risk Alleles (with references).**  
(PDF)

**S3 Table. Demographics of Type 2 Diabetic Cases and Controls.**  
(PDF)

## Acknowledgments

We thank N. Mohamed for editorial assistance. These studies were supported, in part, by The Qatar Foundation and the Weill Cornell Medical College in Qatar and Qatar National Research Foundation NPRP 09-741-3-193, 09-740-3-192 and NPRP 5-436-3-116.

## Author Contributions

Conceived and designed the experiments: SLO MRS NRH RGC. Performed the experiments: CAK AR MJC AA-S OC MG. Analyzed the data: SLO JS JLR-F MRS KAF MDR NH JGM. Contributed reagents/materials/analysis tools: CAK JAM JM MZ AJ RB IKA-A AAA-M MG AB DS JGM RGC. Wrote the paper: SLO JS JLR-F JGM RGC. Patient recruitment, consent and clinical phenotyping: MZ AJ RB AAA-M AA-S OC DS.

## References

1. Stumvoll M, Goldstein BJ, van Haeften TW. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*. 2005; 365: 1333–1346 PMID: [15823385](#)
2. McCarthy MI. Genomics, type 2 diabetes, and obesity. *N Engl J Med*. 2010; 363: 2339–2350 doi: [10.1056/NEJMra0906948](#) PMID: [21142536](#)
3. Bonnefond A, Froguel P. Rare and common genetic events in type 2 diabetes: what should biologists know? *Cell metabolism*. 2015; 21: 357–368 doi: [10.1016/j.cmet.2014.12.020](#) PMID: [25640731](#)
4. Grarup N, Sandholt CH, Hansen T, Pedersen O. Genetic susceptibility to type 2 diabetes and obesity: from genome-wide association studies to rare variants and beyond. *Diabetologia*. 2014; 57: 1528–1541 doi: [10.1007/s00125-014-3270-4](#) PMID: [24859358](#)
5. Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann N Y Acad Sci*. 2010; 1212: 59–77 doi: [10.1111/j.1749-6632.2010.05838.x](#) PMID: [21091714](#)
6. Sanghera DK, Blackett PR. Type 2 Diabetes Genetics: Beyond GWAS. *J Diabetes Metab*. 2012; 3:
7. Foundation ID. IDF Diabetes Atlas Sixth Edition. Available: <http://www.diabetesatlas.org/2013>. Accessed 2016 Jan 1.
8. Hunter-Zinck H, Musharoff S, Salit J, Al-Ali KA, Chouchane L, Gohar A, et al. Population genetic structure of the people of Qatar. *American journal of human genetics*. 2010; 87: 17–25 doi: [10.1016/j.ajhg.2010.05.018](#) PMID: [20579625](#)
9. Rodriguez-Flores JL, Fuller J, Hackett NR, Salit J, Malek JA, Al-Dous E, et al. Exome sequencing of only seven Qataris identifies potentially deleterious variants in the Qatari population. *PLoS one*. 2012; 7: e47614 doi: [10.1371/journal.pone.0047614](#) PMID: [23139751](#)
10. Prasad RB, Groop L. Genetics of type 2 diabetes-pitfalls and possibilities. *Genes*. 2015; 6: 87–123 doi: [10.3390/genes6010087](#) PMID: [25774817](#)
11. American Diabetes A. Standards of medical care in diabetes—2013. *Diabetes care*. 2013; 36 Suppl 1: S11–66 doi: [10.2337/dc13-S011](#) PMID: [23264422](#)
12. Omberg L, Salit J, Hackett N, Fuller J, Matthew R, Chouchane L, et al. Inferring genome-wide patterns of admixture in Qataris using fifty-five ancestral populations. *BMC genetics*. 2012; 13: 49 doi: [10.1186/1471-2156-13-49](#) PMID: [22734698](#)

13. Rodriguez-Flores JL, Fakhro K, Agosto-Perez F, Ramstetter MD, Arbiza L, Vincent TL, et al. Indigenous Arabs are descendants of the earliest split from ancient Eurasian populations. *Genome Res.* 2016; 26: 151–162 doi: [10.1101/gr.191478.115](https://doi.org/10.1101/gr.191478.115) PMID: [26728717](https://pubmed.ncbi.nlm.nih.gov/26728717/)
14. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010; 26: 589–595 doi: [10.1093/bioinformatics/btp698](https://doi.org/10.1093/bioinformatics/btp698) PMID: [20080505](https://pubmed.ncbi.nlm.nih.gov/20080505/)
15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20: 1297–1303 doi: [10.1101/gr.107524.110](https://doi.org/10.1101/gr.107524.110) PMID: [20644199](https://pubmed.ncbi.nlm.nih.gov/20644199/)
16. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25: 2078–2079 doi: [10.1093/bioinformatics/btp352](https://doi.org/10.1093/bioinformatics/btp352) PMID: [19505943](https://pubmed.ncbi.nlm.nih.gov/19505943/)
17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics.* 2007; 81: 559–575 PMID: [17701901](https://pubmed.ncbi.nlm.nih.gov/17701901/)
18. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics.* 2003; 19: 149–150 PMID: [12499305](https://pubmed.ncbi.nlm.nih.gov/12499305/)
19. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature genetics.* 2010; 42: 105–116 doi: [10.1038/ng.520](https://doi.org/10.1038/ng.520) PMID: [20081858](https://pubmed.ncbi.nlm.nih.gov/20081858/)
20. Xu J, Wang J, Chen B. SLC30A8 (ZnT8) variations and type 2 diabetes in the Chinese Han population. *Genetics and molecular research: GMR.* 2012; 11: 1592–1598 doi: [10.4238/2012.May.24.1](https://doi.org/10.4238/2012.May.24.1) PMID: [22653633](https://pubmed.ncbi.nlm.nih.gov/22653633/)
21. Replication DIG, Meta-analysis C, Asian Genetic Epidemiology Network Type 2 Diabetes C, South Asian Type 2 Diabetes C, Mexican American Type 2 Diabetes C, Type 2 Diabetes Genetic Exploration by Nex-generation sequencing in multi-Ethnic Samples C, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics.* 2014; 46: 234–244 doi: [10.1038/ng.2897](https://doi.org/10.1038/ng.2897) PMID: [24509480](https://pubmed.ncbi.nlm.nih.gov/24509480/)
22. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nature genetics.* 2004; 36: 1312–1318 PMID: [15543147](https://pubmed.ncbi.nlm.nih.gov/15543147/)
23. International HapMap C. The International HapMap Project. *Nature.* 2003; 426: 789–796 PMID: [14685227](https://pubmed.ncbi.nlm.nih.gov/14685227/)
24. Cauchi S, El Achhab Y, Choquet H, Dina C, Kremler F, Weitgasser R, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med (Berl).* 2007; 85: 777–782
25. Zhou Y, Park SY, Su J, Bailey K, Ottosson-Laakso E, Shcherbina L, et al. TCF7L2 is a master regulator of insulin production and processing. *Hum Mol Genet.* 2014; 23: 6419–6431 doi: [10.1093/hmg/ddu359](https://doi.org/10.1093/hmg/ddu359) PMID: [25015099](https://pubmed.ncbi.nlm.nih.gov/25015099/)
26. Shao W, Wang D, Chiang YT, Ip W, Zhu L, Xu F, et al. The Wnt signaling pathway effector TCF7L2 controls gut and brain proglucagon gene expression and glucose homeostasis. *Diabetes.* 2013; 62: 789–800 doi: [10.2337/db12-0365](https://doi.org/10.2337/db12-0365) PMID: [22966074](https://pubmed.ncbi.nlm.nih.gov/22966074/)
27. Hara K, Shojima N, Hosoe J, Kadowaki T. Genetic architecture of type 2 diabetes. *Biochem Biophys Res Commun.* 2014; 452: 213–220 doi: [10.1016/j.bbrc.2014.08.012](https://doi.org/10.1016/j.bbrc.2014.08.012) PMID: [25111817](https://pubmed.ncbi.nlm.nih.gov/25111817/)
28. Agarwala V, Flannick J, Sunyaev S, Go TDC, Altshuler D. Evaluating empirical bounds on complex disease genetic architecture. *Nat Genet.* 2013; 45: 1418–1427 doi: [10.1038/ng.2804](https://doi.org/10.1038/ng.2804) PMID: [24141362](https://pubmed.ncbi.nlm.nih.gov/24141362/)
29. Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant... or not? *Hum Mol Genet.* 2002; 11: 2417–2423 PMID: [12351577](https://pubmed.ncbi.nlm.nih.gov/12351577/)
30. Almawi WY, Nemr R, Keleshian SH, Ehtay A, Saldanha FL, AlDoseri FA, et al. A replication study of 19 GWAS-validated type 2 diabetes at-risk variants in the Lebanese population. *Diabetes Res Clin Pract.* 2013; 102: 117–122 doi: [10.1016/j.diabres.2013.09.001](https://doi.org/10.1016/j.diabres.2013.09.001) PMID: [24145053](https://pubmed.ncbi.nlm.nih.gov/24145053/)
31. Al-Daghri NM, Alkharfy KM, Alokail MS, Alenad AM, Al-Attas OS, Mohammed AK, et al. Assessing the contribution of 38 genetic loci to the risk of type 2 diabetes in the Saudi Arabian Population. *Clin Endocrinol (Oxf).* 2014; 80: 532–537
32. Cauchi S, Ezzidi I, El Achhab Y, Mtraoui N, Chaieb L, Salah D, et al. European genetic variants associated with type 2 diabetes in North African Arabs. *Diabetes Metab.* 2012; 38: 316–323 doi: [10.1016/j.diabet.2012.02.003](https://doi.org/10.1016/j.diabet.2012.02.003) PMID: [22463974](https://pubmed.ncbi.nlm.nih.gov/22463974/)
33. Alsmadi O, Al-Rubeaan K, Mohamed G, Alkayal F, Al-Saud H, Al-Saud NA, et al. Weak or no association of TCF7L2 variants with Type 2 diabetes risk in an Arab population. *BMC medical genetics.* 2008; 9: 72 doi: [10.1186/1471-2350-9-72](https://doi.org/10.1186/1471-2350-9-72) PMID: [18655717](https://pubmed.ncbi.nlm.nih.gov/18655717/)

34. Ezzidi I, Mtraoui N, Cauchi S, Vaillant E, Dechaume A, Chaieb M, et al. Contribution of type 2 diabetes associated loci in the Arabic population from Tunisia: a case-control study. *BMC medical genetics*. 2009; 10: 33 doi: [10.1186/1471-2350-10-33](https://doi.org/10.1186/1471-2350-10-33) PMID: [19368707](https://pubmed.ncbi.nlm.nih.gov/19368707/)
35. Tomei S, Mamtani R, Al Ali R, Elkum N, Abdulmalik M, Ismail A, et al. Obesity susceptibility loci in Qataris, a highly consanguineous Arabian population. *J Transl Med*. 2015; 13: 119 doi: [10.1186/s12967-015-0459-3](https://doi.org/10.1186/s12967-015-0459-3) PMID: [25890290](https://pubmed.ncbi.nlm.nih.gov/25890290/)
36. Goddard KA, Hopkins PJ, Hall JM, Witte JS. Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *American journal of human genetics*. 2000; 66: 216–234 PMID: [10631153](https://pubmed.ncbi.nlm.nih.gov/10631153/)
37. Gu S, Pakstis AJ, Li H, Speed WC, Kidd JR, Kidd KK. Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations. *European journal of human genetics: EJHG*. 2007; 15: 302–312 PMID: [17202997](https://pubmed.ncbi.nlm.nih.gov/17202997/)
38. Bener A, Zirie M, Al-Rikabi A. Genetics, obesity, and environmental risk factors associated with type 2 diabetes. *Croat Med J*. 2005; 46: 302–307 PMID: [15849854](https://pubmed.ncbi.nlm.nih.gov/15849854/)
39. Popkin BM. Nutrition Transition and the Global Diabetes Epidemic. *Current diabetes reports*. 2015; 15: 64 doi: [10.1007/s11892-015-0631-4](https://doi.org/10.1007/s11892-015-0631-4) PMID: [26209940](https://pubmed.ncbi.nlm.nih.gov/26209940/)
40. Ma RC, Chan JC. Type 2 diabetes in East Asians: similarities and differences with populations in Europe and the United States. *Ann N Y Acad Sci*. 2013; 1281: 64–91 doi: [10.1111/nyas.12098](https://doi.org/10.1111/nyas.12098) PMID: [23551121](https://pubmed.ncbi.nlm.nih.gov/23551121/)
41. Haiman CA, Fesinmeyer MD, Spencer KL, Buzkova P, Voruganti VS, Wan P, et al. Consistent directions of effect for established type 2 diabetes risk variants across populations: the population architecture using Genomics and Epidemiology (PAGE) Consortium. *Diabetes*. 2012; 61: 1642–1647 doi: [10.2337/db11-1296](https://doi.org/10.2337/db11-1296) PMID: [22474029](https://pubmed.ncbi.nlm.nih.gov/22474029/)
42. Waters KM, Stram DO, Hassanein MT, Le Marchand L, Wilkens LR, Maskarinec G, et al. Consistent association of type 2 diabetes risk variants found in europeans in diverse racial and ethnic groups. *PLoS genetics*. 2010; 6:
43. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nature genetics*. 2008; 40: 695–701 doi: [10.1038/ng.f.136](https://doi.org/10.1038/ng.f.136) PMID: [18509313](https://pubmed.ncbi.nlm.nih.gov/18509313/)
44. Doria A, Patti ME, Kahn CR. The emerging genetic architecture of type 2 diabetes. *Cell metabolism*. 2008; 8: 186–200 doi: [10.1016/j.cmet.2008.08.006](https://doi.org/10.1016/j.cmet.2008.08.006) PMID: [18762020](https://pubmed.ncbi.nlm.nih.gov/18762020/)