Gene 548 (2014) 22–28

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Utility of large consanguineous family-based model for investigating the genetics of type 2 diabetes mellitus☆

Sawsan Al-Sinani [a,*], Mohammed Othman Hassan [a], Fahad Zadjali [a], Said Al-Yahyaee [a], Sulayma Albarwani [a], Syed Rizvi [a], Deepali Jaju [a], Anthony Comuzzie [b], Venkata Saroja Voruganti [c], Riad Bayoumi [a]

[a] College of Medicine & Health Sciences, Sultan Qaboos University, Muscat, Oman
[b] Department of Genetics, Texas Biomedical Research Institute, TX, USA
[c] Department of Nutrition and UNC Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC, USA

ARTICLE INFO

Article history:
Received 8 September 2013
Received in revised form 16 June 2014
Accepted 23 June 2014
Available online 30 June 2014

Keywords:
Genetics
Type 2 diabetes
Family
Pedigree
Arab
Oman

ABSTRACT

Objectives: This study examined the utility of a family-based model for replicating the results of genome-wide association studies (GWAS) of type 2 diabetes (T2D).
Methods and results: In a total of 232 members of a large consanguineous Omani Arab pedigree (age: 16–80 years), there were 27 diabetics and 50 prediabetics (17 with impaired fasting glucose and 33 with impaired glucose tolerance). All 232 individuals underwent anthropometric and biochemical investigations and genotyped for 14 known common gene variants of modest effect on T2D risk. Power analysis at a LOD score of 3, gave 80% power to locate a single specific locus that accounts for 52% of the total phenotypic variation. Measured genotype analysis (MGA) was used to determine heritability of various quantitative traits (QTs) which ranged 25–56%. Using MGA, some common gene variants were found to have little (<5%) but significant impact on the heritability of T2D related QTs [*KCNJ11* (rs5219), p = 0.004]; [*IGF2BP2* (rs4402960), p = 0.02]; [*SLC30A8* (rs13266634), p = 0.05]; [*CAPN10* (rs2975760), p = 0.031]; [*FTO* (rs8050136), p = 0.023]; [*FTO* (rs9939609), p = 0.018] and [*SLC30A8* (rs13266634), p = 0.05]. Sib-TDT analysis showed that some gene variants were significantly associated with T2D risk but didn't reach the level of significance after Bonferroni correction [*KCNJ11* (rs5219), p = 0.047] and [*CAPN10* (rs41266971), p = 0.035].
Conclusion: We have demonstrated that, in principle, a family-based model with minor limitations could be used to replicate some of the results of large GWAS case–control studies. This model could successfully be applied for the future discovery, by deep sequencing, of rare gene variants.

1. Introduction

Most common diseases are complex or multifactorial with both environmental and genetic contributions along with their interaction effects. A "common disease/common variant" (CDCV) hypothesis, which presupposes that different combinations of common allele aggregate in specific individuals to increase disease risk, has been popularized and was supported as an explanation for common disorders (Reich and Lander, 2001; Risch and Merikangas, 1996). Common ancient polymorphisms (frequency > 5%) are shared by all human populations and account for approximately 90 to 95% of human variation (McClellan and King, 2010). The CDCV hypothesis led to a generation of genome-wide association studies (GWAS) identifying regions influencing disease status or risk factor levels (Lupski et al., 2011). However, only <15% of heritability could be accounted for by these variants, suggesting that much remains to be discovered (Wheeler and Barroso, 2011).

Type 2 diabetes (T2D) is a complex common disease. Genetic, environmental and demographic factors, and their interaction, determine an individual's risk for T2D; its heritability has been estimated at approximately 25% (Poulsen et al., 1999). In 2006, common intronic variants within *TCF7L2* gene were found to hold a strong genetic risk for T2D, which was subsequently confirmed in GWAS and replicated in different ethnicities (Cauchi et al., 2007; Grant et al., 2006; Sladek et al., 2007). Since 2007, further GWAS identified and confirmed the following as T2D risk genes or loci: *SLC30A8*, *HHEX*, *CDKAL1*, *IGF2BP2*, *CDKN2A/B* and *FTO* (Frayling et al., 2007; Saxena et al., 2007; Scott et al., 2007; Zeggini et al., 2007). Most of the above identified and replicated gene variants have odds ratios of 1.14 to 1.4 on T2D risk, which is considered to be a fairly modest risk (Morris et al., 2012).

Abbreviations: GWAS, genome-wide association studies; T2D, type 2 diabetes; MGA, measured genotype analysis; QTs, quantitative traits; CDCV, common disease/common variant; CDRV, common disease/rare variant; OFS, Oman Family Study; IGT, impaired glucose tolerance; IFG, impaired fasting glucose; PCR, polymerase chain reaction; SOLAR, Sequential Oligogenic Linkage Analysis Routines; PEDSYS, Pedigree Data Management System; S-TDT, sibship transmission/disequilibrium test.

☆ Ethics guidelines: The study was approved by the Ethics and Research Committee of the College of Medicine, Sultan Qaboos University, Muscat, Oman; in full compliance with the guidelines for human experimentation set by the Ministry of Health in Oman (www.moh.gov.om).

* Corresponding author.
E-mail address: Sawsan.alsinani@gmail.com (S. Al-Sinani).

http://dx.doi.org/10.1016/j.gene.2014.06.053
0378-1119/

These recent GWAS have thus demonstrated that the population from which one comes and its collection of older common variants may have less influence on an individual's disease susceptibility than the collection of recently arisen rare variants (Lupski et al., 2011). Rare large-effect mutations have been suggested as causes of many complex diseases and is described under the alternative common disease/rare variant (CDRV) hypothesis (Hall, 2010; Lyssenko and Groop, 2009; McClellan and King, 2010). Therefore, study models that can be developed to investigate rare variants are needed.

The main difficulties associated with studies of complex traits are to achieve the necessary statistical power and reproducibility. In heterogeneous populations, very large sample sizes are required for successful detection of the effects of multiple genes, both for association and linkage studies (Frayling, 2007; Groop and Lyssenko, 2008; Zeggini, 2007). However, an ideal population to provide statistical power required to study complex diseases with confidence preferably consists of multigenerational pedigrees, descended from a small number of founders just a few generations ago, and with environmental homogeneity, restricted geographical distribution, detailed records, well-ascertained and validated pedigrees, and inbreeding as a norm (Arcos-Burgos and Muenke, 2002). These large pedigrees have a greater chance of observing multiple copies of disease associated gene variants among related individuals.

Hassan et al. (2005) introduced a unique model, Oman Family Study (OFS), for examining the genetics of hypertension, the dyslipidemias, obesity, diabetes and the metabolic syndrome. In this model they chose a homogeneous, genetically and geographically, isolated Arab population consisting of 5 pedigrees of very large size (160–325 individuals each), with few recent founders (5th–10th generations), a very high degree of inbreeding, polygamy and excellent genealogy records. The statistical power of this model was evaluated and indicated that it has an outstanding power to detect susceptibility loci for common diseases.

This study is a continuation of OFS. It examined the utility of one pedigree of 232 individuals in replicating the association of 14 known common gene variants, of moderate effect sizes, with susceptibility to T2D risk. We believe that this model is at the limit of the power required to examine known common gene variants. But we anticipate that this model could be successfully used to uncover rare gene variants of large-effect with high mean effects on complex disease risk.

2. Subjects and methods

2.1. Study population (pedigree)

The OFS is a homogeneous Arab population made of five large, extended and highly consanguineous pedigrees (Bayoumi et al., 2007; Hassan et al., 2005; Lopez-Alvarenga et al., 2008). The numbers of subjects in these five pedigrees were 325, 160, 230, 281 and 279, totaling 1275 subjects (Hassan et al., 2005). Although the total number of founders ranged between 70 and 100 in each pedigree, most of these were due to marriages outside the pedigrees. The 5th to 10th generation founders were usually very few and ranged between 3 and 17 individuals. The rapid population growth made these pedigrees fairly young isolates of 7 to 12 generations each (Bayoumi et al., 2007).

Detailed inquiry revealed that all individuals in OFS seemed to be related to common ancestry, and intermarriage is common among them. Cousin marriages represent >35% of all marriages and a further 20% between tribal groupings (Hassan et al., 2005). Polygamy is widely practiced, with some men marrying up to four wives. Family relationships were ascertained initially by local staff, volunteers, and elders from each village. A questionnaire was filled in for each subject, each of whom was given a unique identification number that was used for creating a master file for all tests and analyses (Bayoumi et al., 2007).

The present study was carried out on only one of the 5 pedigrees of OFS (pedigree 3). It consisted of nine generations with a total of 511 closely related individuals. It was selected because of a high number of individuals within the pedigree with T2D and prediabetes. Phenotypic data were collected from 232 living subjects (above age of 16 years). Pedigree data are summarized in Table 1.

Participants were informed about the project and written or thumb stamped consents were obtained. The study was approved by the Ethics and Research Committee of the College of Medicine, Sultan Qaboos University, Muscat, Oman.

2.2. Demographic and biochemical data collection

A satellite Research Centre was established in Nizwa Polyclinic, 150 km from the capital Muscat. This center was used for anthropometric measurements, data and sample collection (Bayoumi et al., 2007). Height and weight were measured using standard methods. Waist circumference was measured by a soft tape at the largest circumference between the lowest rib and iliac crest. Body fat percentage was assessed using electrical impedance (Tanita, Tokyo, Japan).

All participants went through a standard 75 g oral glucose tolerance test. Biochemical tests included fasting blood glucose and 2 h post-glucose load, fasting and 2-hour serum insulin levels. All biochemical tests were done the same day of sample collection or frozen at −80 °C until done at Sultan Qaboos University Hospital laboratories using automated equipment (Synchron 7, Access II and Image; Beckman Coulter, Fullerton, CA). Quality is assured by participation in international and local quality control programs.

Diabetes and prediabetes (IGT and IFG) were defined according to 2006 World Health Organization (WHO) criteria (diabetes: fasting plasma glucose ≥7.0 mmol/l or 2-h postprandial plasma glucose ≥11.1 mmol/l; IGT: fasting plasma glucose <7.0 mmol/l and 2-h postprandial plasma glucose ≥7.8 and <11.1 mmol/l; IFG: fasting plasma glucose 6.1 to 6.9 mmol/l and 2-h postprandial plasma glucose <7.8 mmol/l). Obesity status was defined among participants using the international classification of adult normal weight, overweight and obesity according to BMI (global database on body mass index), using criteria of the WHO-1995, WHO-2000 and WHO-2004 [Normal range: 18.5–24.99, overweight: 25.00–29.99 and obese ≥30.00 kg/m^2] (WHO, 1996). Anthropometric and clinical characteristics of all participants are summarized in Table 2.

2.3. Genotyping

Selection of T2D common gene variants was based on GWAS studies that investigated T2D extensively (Altshuler et al., 2000; Cauchi et al., 2007; Frayling et al., 2007; Gloyn et al., 2003; Grant et al., 2006; Hanis et al., 1996; Horikawa et al., 2000; Nielsen et al., 2003; Saxena et al., 2007; Scott et al., 2007; Sladek et al., 2007; Weedon et al., 2003; Zeggini et al., 2007). The variants which were selected were those of modest effect sizes on T2D risk; investigated repeatedly in different populations and their association with susceptibility to T2D was significantly confirmed.

The 232 participants were genotyped for the following gene variants: *TCF7L2* (rs7903146 & rs7901695), *CDKN2A/B* (rs10811661), *PPARG* (rs1801282), *FTO* (rs8050136 & rs9939609), *IGF2BP2* (rs4402960), *KCNJ11* (rs5219), *HHEX* (rs1111875), *SLC30A8* (rs13266634), *CDKAL1* (rs10946398) and *CAPN10* (rs41266971, rs2975760 and rs3792267). DNA fragments containing polymorphic sites were amplified and

Table 1
Pedigree data.

Generations	9
Total closely related individuals	511
Individuals tested	230
Founders	3
Nuclear families	98
Mean inbreeding coefficient	0.0206