Tumor Biology 45 (2023) 15–21
DOI:10.3233/TUB-220025
IOS Press

Medium and large alleles of the *PGC* gene are risk factors for gastric cancer

Josefina Yoaly Sánchez-López[a,*], Katia Carolina Vázquez-Ibarra[a,b], Andrea Marlene García-Muro[a,b], Azaria García-Ruvalcaba[a,b], Sergio Pacheco-Sotelo[c], Luis Carlos Díaz-Herrera[d] and María Eugenia Marin-Contreras[d]

[a]*Division of Genetics, Western Biomedical Research Center, Mexican Institute of Social Security, Guadalajara, Jalisco, Mexico*

[b]*Doctoral Program in Human Genetics, Department of Molecular Biology and Genomics, University Center of Health Sciences, University of Guadalajara, Guadalajara, Jalisco, Mexico*

[c]*Pediatric Hospital, National Medical Center of the West, Mexican Institute of Social Security, Guadalajara, Jalisco, Mexico*

[d]*Department of Gastroenterology, National Medical Center of the West, Mexican Institute of Social Security, Guadalajara, Jalisco, Mexico*

Received 5 August 2022
Accepted 9 January 2023

Abstract.
BACKGROUND: A 100-bp insertion/deletion polymorphism in the pepsinogen C gene has been associated with the risk of gastric cancer (GC).
OBJECTIVE: We analyzed the relationships of the 100-bp insertion/deletion polymorphism with GC, atrophic gastritis (AG), and intestinal metaplasia (IM) in the Mexican general population (MGP).
METHODS: We studied the genomic DNA of subjects with GC $n = 80$, AG and IM $n = 60$, controls $n = 110$, and the MGP $n = 97$. PGC gene insertion/deletion polymorphism was identified by means of PCR, capillary electrophoresis and GeneScan software.
RESULTS: Different allele sizes of PGC polymorphism were observed in the studied groups, from 266 bp to 499 bp, which were grouped for the analysis as short alleles of 266–399 bp, medium alleles of 400–433 bp and large alleles of 434–499 bp. Carriers of one or two medium alleles, had an increased risk of GC, with OR of 1.99 (CI95%1.08–3.67 $p = 0.026$) compared to homozygotes (no medium/no medium).
CONCLUSIONS: Previous studies have related PGC short alleles to risk for or protection against GC depending on the ethnic origin of the population. In our study, medium alleles were related to risk for GC. Further studies are required to establish the importance of this polymorphism in the origin of gastric neoplasia.

Keywords: Gastric cancer, polymorphisms, neoplasia

*Corresponding author: Josefina Yoaly Sánchez-López, PhD, Division of Genetics, Western Biomedical Research Center, Mexican Institute of Social Security, Sierra Mojada 800 Col. Independencia, Guadalajara, CP44340 Jalisco, México. Tel.: +52336389410; E-mails: yosalo1795@yahoo.com; josefina.sanchezl@imss.gob.mx. ORCID ID: 0000-0002-0876-6225.

1. Introduction

Pepsinogen C or Pepsinogen II (PGC or PG-II) is a 42-kDa protein that belongs to the aspartic protease family and is synthesized by chief cells of the gastric mucosa and secreted to the stomach under acidic conditions. PGC is activated to pepsin C to digest polypeptides and amino acids. This protein is encoded by the progastricsin (*PGC*) gene, which is located on chromosome 6p21.1 and consists of 9 exons [1].

In situ expression of PGC occurs mainly in the normal gastric mucosa [2]. PGC expression has been detected in different gastric diseases by immunohistochemistry; positive PGC expression (100%) is observed in normal gastric mucosa, whereas its positivity declines considerably in GC [3].

Several variants in the *PGC* gene can influence PGC expression and regulation, for instance, rs6912200, rs6941539, rs9471643, and rs6458238 [4–6]; however, the most studied polymorphism is the 100-bp insertion/deletion located between exons 7 and 8 of the *PGC* gene. Based on fragment lengths, following agarose gel electrophoresis of PCR products, four alleles with different molecular weights were initially obtained: 310 bp, 400 bp, 450 bp and 480 bp [7].

A 100-bp insertion/deletion polymorphism is associated with the risk of peptic ulcer and GC, and individuals homozygous for the 310-bp allele have downregulated PGC expression [8, 9]; more recently, Kumar used similar methods and further concluded that being homozygous for the 310-bp allele of PGC may lead to elevated serum PGC levels in cases of GC, especially in patients with *Helicobacter pylori* infection and intestinal metaplasia [10]. Moreover, Pinto-Correia reported that being homozygous for the 310-bp allele was related to the upregulation of PGC expression, which served as a protective factor in the development of gastric disease [11].

In this work, we analyzed the 100-bp insertion/deletion polymorphism in subjects with GC, atrophic gastritis and intestinal metaplasia, controls, and the Mexican general population by means of PCR-capillary electrophoresis and GeneScan, which represents an easy, quick and inexpensive method for more precisely identifying the size of alleles for this polymorphism.

2. Methods

We studied the genomic DNA of Mexican subjects with gastric cancer (GC group, $n = 80$), atrophic gastritis and intestinal metaplasia (AG/IM group, $n = 60$), and no atrophic gastritis (control group NAG, $n = 110$) and the Mexican general population (MGP group, $n = 97$).

Diagnoses of GC, AG, IM and NAG were made by a pathologist through analysis of gastric biopsies obtained by endoscopy. NAG (controls) were subjects with chronic gastritis without athrophy, none of them had healthy gastric mucosa. MGP individuals were healthy donors and they did not undergo upper endoscopy (this group represent the western region of Mexico). All subjects were recruited from the gastroenterology department from two hospitals, the Pediatric Hospital and Specialty Hospital, as well blood bank, from the Western National Medical Center, Mexican Institute of Social Security, in the city of Guadalajara, Jalisco, Mexico.

The study was approved by the National Committee of Health Research, IMSS (CNIC-2017-785-087). All subjects provided written informed consent. The study conforms with the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the British Medical Journal (18 July 1964).

DNA was extracted from leucocytes of peripheral blood by the Miller method [12]. The PGC gene insertion/deletion polymorphism was identified by means of PCR and capillary electrophoresis using the following primers: forward 5'-FAM/GGCCAGATCTGCGTGTTTTA-3' and reverse 5'-AGCCCTAAGCCTGTTTTTGG-3' [13]. PCR was performed in a 15-μL reaction mixture comprising

Table 1
PGC size alleles observed in the studied groups

	GC ($n=80$)	AG/IM ($n=60$)	Controls ($n=110$)	MGP ($n=97$)	p-value
Sex					**CG vs controls $p=$ 0.000**
Female	(28) 35.0%	(35) 58.3%	(69) 62.7%	(43) 44.3%	OR 0.32 (0.17–0.58)
Male	(52) 65.0%	(25) 41.7%	(41) 37.3%	(54) 55.7%	OR 3.125 (1.71–5.69)
Age, years median	59.0	58.5	49.0	32.0	GC vs controls **$p=$ 0.000** AG/IM vs controls **$p=$ 0.001**
Number total of different alleles	28	24	34	22	
Short alleles (266–399bp)	(51) 31.9%	(51) 42.5%	(101) 45.9%	(64) 33.0%	
Medium alleles (400–433 bp)	(85) 53.1%	(54) 45.0%	(91) 41.4%	(86) 44.3%	
Large alleles (434–499 bp)	(24) 15.0%	(15) 12.5%	(28) 12.7%	(44) 22.7%	

GC: Gastric cancer, AG: Atrophic gastritis, IM: Intestinal metaplasia, NAG: No atrophic gastritis, MGP: Mexican general population.

200 ng genomic DNA, each primer at 5 pM, 0.5 U Taq polymerase (Invitrogen, USA), 1X PCR buffer, 1.5 mM MgCl2, 2.0 mM deoxynucleotide triphosphate mix (dNTP Set; Vivantis, Malaysia), and 5% DMSO. PCR conditions were 94°C for 5 min, followed by 28 cycles of 94°C for 30 s, 57°C for 30 s, and 72°C for 30 s and a final step at 72°C for 5 min; PCR was carried out in a 2720 Thermal Cycler (Applied Biosystems, USA). PCR fluorescent products were resolved by capillary electrophoresis in a 310 Genetic Analyzer (Applied Biosystems, USA) with Hi-Di Formamide (Applied Biosystems, USA) and GeneScan 600 LIZ Size Standard v2.0 (Applied Biosystems, USA); the fragments were analyzed by GeneScan software (ThermoFisher Scientific).

3. Results

The median age and sex of the Mexican subjects here studied are shown in Table 1. We found differences in terms of sex between GC group and controls ($p=0.000$), Odds Ratio analysis showed protection against GC for females with OR 0.320 (0.17–0.58), and an increased risk for GC for males with OR 3.125 (1.71–5.69). For AG/IM vs controls was not significant ($p=$ NS). Also, we observed differences regarding age, between GC and controls ($p=0.000$), as well AG/IM and controls ($p=0.001$), adults > 60 years had an increased risk for GC with OR of 1.98 (1.06–3.68) $p=0.030$ as well for AG/IM with OR of 2.03 (1.3–3.99) $p=0.038$ (Table 1).

Different sizes of alleles of PGC polymorphism were observed in the studied groups, from 266 bp to 499 bp, which were grouped for the analysis as follows: short alleles of 266–399 bp, medium alleles of 400–433 bp and large alleles of 434–499 bp (Fig. 1). The control group had the highest diversity of alleles, with 34 different alleles (in a range of sizes from 267 to 470 bp), followed by the GC group

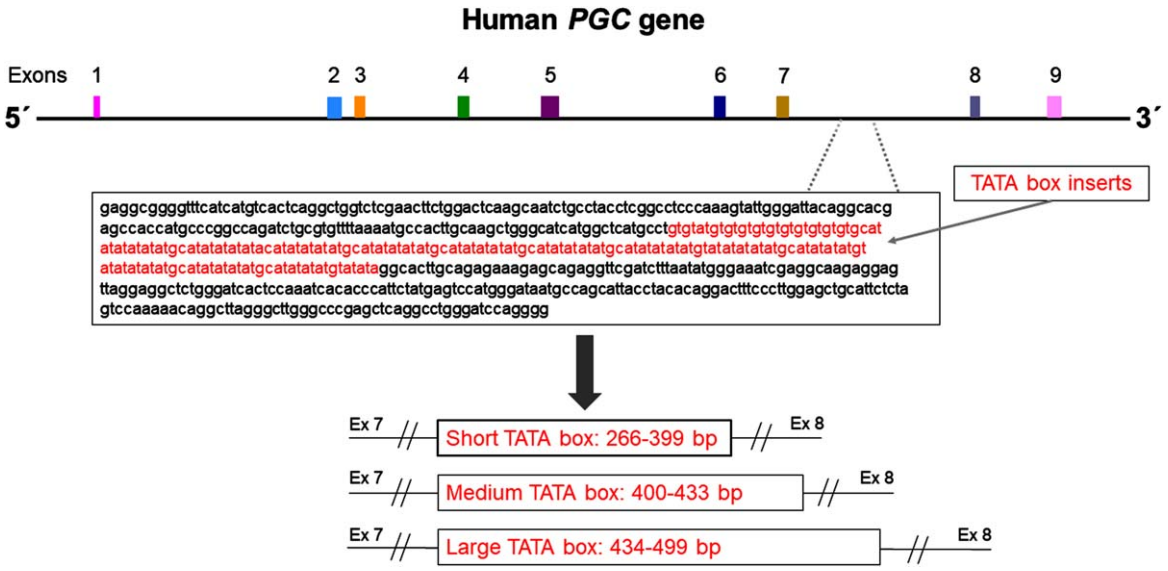

Fig. 1. Structure of human Progastricsin (*PGC*) gene. The 100-bp insertion/deletion polymorphism located between exons 7 and 8 is shown, different size of alleles/repeats in TATA box are grouped as short, medium and large. [Modified from 13].

with 28 alleles (305–499 bp), the AG/IM group with 24 alleles (305–471 bp) and the MGP group with 22 alleles (266–496 bp).

The three most frequent alleles observed in each group were as follows: in the GC group, 305 bp 29.4%, 431 bp 28.1%, and 406 bp 5.0%; in the AG/IM group, 305 bp 40.0%, 432 bp 15.0%, and 431 bp 11.7%; in controls, 305 bp 42.3%, 431 bp 12.3%, and 430 bp 5.5%; and in the MGP, 433 bp 32.0%, 305 bp 30.9%, and 411 bp 4.6%.

Short alleles were more frequently observed in the control group, medium alleles were more frequent in the GC group, and large alleles were more frequent in the MGP group. Significant differences were observed between the GC and control groups ($p = 0.021$) as well as the MGP and control groups ($p = 0.0055$) (Table 1).

The odds ratio analysis performed between the GC group and the control group showed that short alleles had an OR of 0.55 (CI95% 0.36–0.84, $p = 0.006$), meaning that short alleles (266–399 bp) confer protection against GC; conversely, the medium and large alleles showed an OR of 1.81 (CI95% 1.18–2.77, $p = 0.006$), thus conferring an increased risk for GC. With Cochrane-Armitage trend test [14], we found that under the recessive inheritance model, carriers of two short alleles (homozygotes) had protection against GC, with an OR of 0.42 (CI95% 0.22–0.812, $p = 0.009$), compared to those with only one or no short alleles. Under the dominant model, carriers of one or two medium alleles, had an increased risk of GC, with OR of 1.99 (CI95% 1.08–3.67) $p = 0.026$ compared to homozygotes (no medium/no medium) (Table 2).

4. Discussion

The *PGC* gene coding for the pepsinogen II protein consists of nine exons, and different lengths of the TATA box are represented by short (308–378 bp), medium (400–413 bp) and large (434–479 bp) repeats, which are located between exons 7 and 8 [13]. This polymorphic length of the TATA box is better known as the 100-bp insertion/deletion polymorphism and has been related to peptic ulcers and gastric cancer. In this work, we analyzed the relationships of the 100-bp insertion/deletion

Table 2
Cochrane-Armitage trend test analysis comparing short, medium and large PGC genotypes in Gastric Cancer cases versus controls and Atrophic Gastritis and Intestinal Metaplasia versus controls

	Genotypes of PGC			GC vs controls Odds Ratio, (CI95%), p	AG/IM vs controls Odds Ratio, (CI95%), p
	S/S	S/ML	ML/ML	Short alleles vs Medium+Large alleles	
GC ($n=80$)	17	19	44	Dominant OR 0.73 (0.41–1.30), $p=0.292$	Dominant OR 0.89 (0.48–1.68), $p=0.733$
AG/IM ($n=60$)	21	9	30	**Recessive OR 0.42 (0.22–0.81), $p=0.009$**	Recessive OR 0.83 (0.44–1.61), $p=0.598$
Controls ($n=110$)	43	15	52	Codominant OR 1.97 (0.93–4.17), $p=0.072$	Codominant OR 1.11 (0.46–2.73), $p=0.807$
	M/M	M/SL	SL/SL	Medium alleles vs Short+Large alleles	
GC ($n=80$)	28	29	23	**Dominant OR 1.99 (1.08–3.67), $p=0.026$**	Dominant OR 1.20 (0.64–2.28), $p=0.567$
AG/IM ($n=60$)	18	18	24	Recessive OR 1.37 (0.74–2.54), $p=0.315$	Recessive OR 1.09 (0.55–2.17), $p=0.802$
Controls ($n=110$)	31	30	49	Codominant OR 1.51 (0.82–2.81), $p=0.186$	Codominant OR 1.14 (0.57–2.28), $p=0.705$
	L/L	L/SM	SM/SM	Large alleles vs Short+Medium alleles	
GC ($n=80$)	4	16	60	Dominant OR 1.26 (0.64–2.49), $p=0.505$	Dominant OR 0.49 (0.20–1.24), $p=0.130$
AG/IM ($n=60$)	2	5	53	Recessive OR 1.39 (0.34–5.75), $p=0.644$	Recessive OR 0.91 (0.16–5.14), $p=0.918$
Controls ($n=110$)	4	19	87	Codominant OR 1.19 (0.57–2.50), $p=0.632$	Codominant OR 0.43 (0.15–1.23), $p=0.109$

S: short alleles 266–399 bp; M: medium alleles 400–433 bp; L: large alleles 434–499 bp.

polymorphism with GC, AG, and IM in the Mexican population. We found different allele sizes of PGC polymorphism which were grouped as short alleles of 266–399 bp, medium alleles of 400–433 bp and large alleles of 434-499 bp. Carriers of one or two medium alleles, had an increased risk of GC, with OR of 1.99 (CI95%1.08–3.67 $p=0.026$). To our knowledge, up to now there is no consensus on the cut-off point for the size of the alleles of the *PGC* gene, only one study has reported the cut off for size of alleles in *PGC* gene, group 1 or short 308–378 bp, group 2 or medium 400–413 bp, and group 3 or large 434–479 bp [13], therefore we adapted a classification system for PGC alleles very similar to that of those authors, though we found more diversity in the size of PGC alleles than these authors.

Some authors have reported that individuals homozygous for the 310-bp allele have downregulated PGC expression, and this status has been related to gastric cancer in China [8]. More recently, other authors concluded that being homozygous for the 310-bp allele led to elevated serum PGC levels and was associated with gastric cancer in Indian patients [10]. Pinto-Correia et al. 2006 hypothesized that the 100-bp insertion/deletion polymorphism could interfere directly with the number of TATA-box repeats accessible for the activation of *PGC* gene expression. A large number of TATA-boxes together in one sequence might function as confounders for transcriptional activation factors. Thus, short alleles would stabilize the activation of gene expression, increasing its levels [11].

Discordance is observed among results for the frequency 310-bp alleles between Caucasian and Asian populations, conferring protection or risk, respectively, which can be explained by the distinct distributions of frequencies of these alleles between these populations. Specifically, in Caucasian populations, the frequency of the 310-bp allele is lower in gastric cancer patients (0.509 compared

to 0.724 in controls), and this allele can even serve as a protective factor against the development of gastric cancer in the Portuguese population (OR 0.39, CI95% 0.21–0.75, $p = 0.004$) [11]. In contrast, in Asian populations, the frequency of the 310-bp allele is higher in gastric cancer cases than in controls (0.516 vs. 0.357), while homozygotes for the 310-bp allele have an increased risk for gastric cancer with an OR of 2.93 (CI95% 1.08–7.93, $p = 0.033$) [8].

In our population, short alleles (266–399 bp) conferred protection against gastric cancer (OR 0.55, CI95% 0.36–0.84, $p = 0.006$), which is in concordance with the results of Pinto-Correia et al. 2006 [11], and also we observed in our study that carriers of one or two medium alleles had an increased risk of gastric cancer with an OR of 1.99 (CI95% 1.08–3.67, $p = 0.026$). As the methodology used here (capillary electrophoresis) is more precise than agarose or polyacrylamide electrophoresis, we think that the 310-bp allele is represented in our study by the 305-bp allele, which had a lower frequency in the gastric cancer group than in the controls (29.4% vs. 42.3%), as was observed by Pinto-Correia et al. in the Caucasian population [11]. It should be noted that the genetic ancestry of the Mexican Mestizo population is 56% Native American, 37% Caucasian and 5% African [15], which may explain the similarity between our results and those observed for the Caucasian population [11].

Although our results are interesting, the study has some limitations. It is important to note that the observed differences in sex and age between GC and controls, could be attributed to the fact that most of the cases of GC are men (65%) older than 50 years of age (74.7%), while the control group included mainly women (62.7%) under 60 years (72.7%). These differences can be attributed to some hormonal issue in which women have protection against gastric cancer while men do not, or to the fact that more women seek medical attention than men when they present gastric symptoms. The fact that there are differences in terms of age and sex between the GC and control groups suggests that the results could have a bias, which could be corrected by conducting the study with samples matched by age and sex; on the other hand, there is a lack of information about the relationships between PGC alleles and the expression or regulation of PGC levels, which would allow a better approach to the role of this variant in gastric cancer in our population.

5. Conclusions

Previous studies have related short alleles as a risk for or protection against gastric cancer. In our population, short alleles were related to protection against gastric cancer, but we also analyzed and emphasize the importance of medium alleles, which were related to risk for gastric cancer in our population. Further studies are required to establish the importance of this polymorphism in the origin of gastric neoplasia.

Acknowledgments

The authors thank Fondo de Investigación en Salud for grant FIS/IMSS/PROT/PRIO/18/074.

Author contributions

CONCEPTION: JYSL
ANALYSIS OF DATA: JYSL, KCVI, LCDH, AMGM and AGR
PREPARATION OF THE MANUSCRIPT: JYSL and KCVI
REVISION FOR IMPORTANT INTELLECTUAL CONTENT: JYSL, MEMC and SPS
SUPERVISION: JYSL, MEMC and SPS

Conflict of interest

The Authors declare that they have no conflict of interests.

References

[1] Gene [homepage on the Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [updated 2022 May 13; cited 2022 Aug 01]. Available from: https://www.ncbi.nlm.nih.gov/gene/5225

[2] Kageyama T. Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development. Cell Mol Life Sci. 2002;59(2):288–306. DOI: 10.1007/s00018-002-8423-9

[3] Ning PF, Liu HJ, Yuan Y. Dynamic expression of pepsinogen C in gastric cancer, precancerous lesions and Helicobacter pylori associated gastric diseases. World J Gastroenterol. 2005;11(17):2545–8. DOI: 10.3748/wjg.v11.i17.2545

[4] He C, Tu H, Sun L, Xu Q, Li P, Gong Y, et al. Helicobacter pylori related host gene polymorphisms associated with susceptibility of gastric carcinogenesis: a two-stage case–control study in Chinese. Carcinogenesis. 2013;34(7):1450–7. DOI: 10.1093/carcin/bgt079

[5] He CY, Sun LP, Xu Q, Liu JW, Jiang JY, Dong NN, et al. PGC TagSNP and its interaction with H. pylori and relation with gene expression in susceptibility to gastric carcinogenesis. PLoS ONE. 2014;9(12):e115955. DOI: 10.1371/journal.pone.0115955

[6] He C, Xu Q, Tu H, Sun L, Gong Y, Liu J, et al. Polymorphic rs9471643 and rs6458238 upregulate PGC transcription and protein expression in overdominant or dominant models. Mol Carcinog. 2016;55(5):586–99. DOI: 10.1002/mc.22305

[7] Azuma T, Pals G, Taggart RT. RFLP for the human pepsinogen C gene (PGC). Nucleic Acids Res. 1988;16(19):9372. DOI: 10.1093/nar/16.19.9372

[8] Liu HJ, Guo XL, Dong M, Wang L, Yuan Y. Association between pepsinogen C gene polymorphism and genetic predisposition to gastric cancer. World J Gastroenterol. 2003;9(1):50–3. DOI: 10.3748/wjg.v9.i1.50

[9] Sun LP, Gong YH, Dong NN, Wang L, Yuan Y. Correlation of pepsinogen C (PGC) gene insertion/deletion polymorphism to PGC protein expression in gastric mucosa and serum. Ai Zheng. 2009;28(5):487–92.

[10] Kumar S, Kumari N, Mittal RD, Ghoshal UC. Pepsinogen-II 100 bp ins/del gene polymorphism and its elevated circulating levels are associated with gastric cancer, particularly with Helicobacter pylori infection and intestinal metaplasia. Gastric Cancer. 2016;19(3):808–16. DOI: 10.1007/s10120-015-0550-8

[11] Pinto-Correia AL, Sousa H, Fragoso M, Moreira-Dias L, Lopes C, Medeiros R, et al. Gastric cancer in a Caucasian population: role of pepsinogen C genetic variants. World J Gastroenterol. 2006;12(31):5033-6. DOI: 10.3748/wjg.v12.i31.5033

[12] Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215. DOI: 10.1093/nar/16.3.1215

[13] Re V, Zorzi M, Caggiari L, Repetto O, Brisotto G, Magris R, et al. Polymorphisms in Pepsinogen C and miRNA Genes Associate with High Serum Pepsinogen II in Gastric Cancer Patients. Microorganisms. 2021;9(1):126. DOI: 10.3390/microorganisms9010126

[14] Ziegler A, König IR. A Statistical Approach to Genetic Epidemiology. Concepts and applications. 1st ed. Weinheim, Germany: Wiley-VCH 2006. p. 199–214.

[15] Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, et al. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet. 2014;10(9):e1004572. DOI: 10.1371/journal.pgen.1004572