SHORT REPORT

Association of *FGFR2* gene polymorphisms with the risk of breast cancer in population of West Siberia

Uljana A Boyarskikh¹, Natalja A Zarubina², Julia A Biltueva¹, Tatjana V Sinkina², Elena N Voronina¹, Aleksander F Lazarev², Valentina D Petrova², Yurii S Aulchenko³ and Maxim L Filipenko^{*,1}

¹Pharmacogenomics Group, Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia;

²Department of Epidemiology, Altai Branch of Russian N.N. Blokhin Cancer Research Centre, Barnaul, Russia;

³Department of Epidemiology and Biostatistics and Clinical Genetics, Erasmus MC Rotterdam, CA Rotterdam, The Netherlands

Polymorphisms within intron 2 of the *FGFR2* gene have been associated with increased risk of breast cancer (BC) in European and Asian populations. The study by Easton *et al* reported two *FGFR2* SNPs, rs2981582 and rs7895676, to be among those most strongly associated with BC risk. Statistical modeling suggested that rs7895676 was the variant responsible for the association observed in the region. In this work, we studied the association between seven *FGFR2* SNPs, including rs2981582 and rs7895676, and BC risk in the Russian population of 766 case and 665 control women from Siberia, Russian Federation. In our population, allelic frequencies and the magnitude of linkage disequilibrium (LD) were different from those observed in European and Asian populations. The following three SNPs were significantly associated with BC in our study: rs7895676[C] (odds ratio (OR) = 1.28 (1.12–1.43), $P = 1.7 \times 10^{-3}$), rs2981582[T] (OR = 1.46 (1.30–1.62), $P = 2 \times 10^{-6}$) and rs3135718[G] (OR = 1.43 (1.27–1.58), $P = 6 \times 10^{-6}$). The latter two SNPs were in strong ($r^2 = 0.95$) LD in our sample. Maximum likelihood analysis showed that the model, including rs7895676, only explains that the association is significantly ($P < 0.001$) worse than any of the models, including either rs2981582 or rs3135718. Thus, in addition to the confirmation of association of *FGFR2* with the BC risk in this new population, our study has suggested that rs7895676 is not likely to represent the causative variant.

European Journal of Human Genetics (2009) 17, 1688–1691; doi:10.1038/ejhg.2009.98; published online 17 June 2009

Keywords: breast cancer; *FGFR2*; molecular epidemiology; West Siberia

Introduction

Breast cancer (BC) is the most prevalent cancer in women. Several polymorphic loci were recently identified by two independent genome-wide association (GWA) studies as being associated with BC risk.^{1,2} One of the strongest associations was found in the *FGFR2* gene. Both GWA and

replication^{3,4} efforts suggest that the associated locus is limited to LD extending over ~20 kb in intron 2. However, the question of refining the boundaries of association and finding the causative variant within intron 2 of *FGFR2* remains open. In the study by Easton *et al.*,¹ the SNP rs2981582 located at the upstream boundary of intron 2 was the most significantly associated. In an attempt to identify the functional variant, additional 21 SNPs in LD with rs2981582 were genotyped in European (4426 cases, 4444 controls) and Asian (3918 cases, 3100 controls) datasets. The most significant association was observed for rs7895676 located at the downstream boundary of the

*Correspondence: Dr ML Filipenko, Pharmacogenomics Group, Institute of Chemical Biology and Fundamental Medicine, Lavrentieva 8, Novosibirsk, 630090 Russia.

Tel: +7 383 3333571; Fax: +7 383 3333677; E-mail: max@niboch.nsc.ru
Received 23 January 2009; revised 14 April 2009; accepted 7 May 2009; published online 17 June 2009

LD block; this SNP provided a much better (odds > 100) fit to the data than the original upstream SNP rs2981582.

The aim of this work was to study the association between *FGFR2* polymorphisms and BC in Russian women from West Siberia, Russian Federation. Specifically, we aimed at replicating previously described associations and refining the boundaries of association using an ethnic group, which has not been investigated for such an association before.

Materials and methods

The case group included 766 and the control group included 665 women. All the women signed an informed consent to participate in the study. The study protocol was approved by Medical Ethics Committee of Altai Oncological Centre. The case group included patients with sporadic BC. The control group included women randomly selected from blood bank donors. All women were ethnic Russian residents of the Altai Krai, Russian Federation.

For this study, we chose seven SNPs. Three of these (rs3135718, rs2981582 and rs7895676) were in strong LD and showed an association with BC in the study by Easton *et al.*¹ To study whether in the Russian population, the LD block is the same as in the other European populations, the other four SNPs were chosen from these flanking the European LD block at the 5'- (rs4647913 and rs3135715) and 3'-ends (rs2981428 and rs10510097).

Two SNPs, rs7895676 and rs2981428, were genotyped in part by allele-specific real-time PCR and in part by direct sequencing. Other SNPs were genotyped by real-time PCR using TaqMan probes (ICBFM, Novosibirsk, Russia).

All statistical analyses were carried out using the free R-2.6.0 software (<http://www.r-project.org>) and associated libraries GenABEL,⁵ haplo.stats,⁶ genetics and rmeta.

For further details, see Supplementary Materials and methods.

Results

We identified genotypes of rs4647913, rs3135715, rs3135718, rs2981582, rs7895676, rs2981428 and rs10510097 in Siberian cases and controls (Table 1). Of the seven typed SNPs, four are present in the HapMap database (<http://www.hapmap.org/>). For the SNP rs2981428, allelic frequencies observed in the West Siberian population were close to those observed in the HapMap Asian samples (Table 1); rs2981582 has shown a frequency that was intermediate between European and Asian populations; the frequency of rs10510097[T] was significantly higher in the Siberian compared with that in the Asian ($P=0.02$ with HCB, $P=10^{-5}$ with JPT) and the European ($P=10^{-8}$) populations and the frequency of rs3135715[C] was significantly lower in the Siberian

Table 1 Association between breast cancer and SNPs from the *FGFR2* region tested in the population of Western Siberia

dbSNP_rs	Number of		OR (95% CI)	P	P _{HWE}	CA	OR (95% CI)	P	West Siberia controls	Coded allele frequency	
	Controls	Cases								HapMap HCB/JPT	HapMap CEU
rs10510097	C/C 247 223 75	T/T 274 295 93	Ref. 1.19 [0.95–1.44] 1.12 [0.77–1.47]	— 0.16 0.53	0.03	T	1.09 [0.93–1.2]	0.29	0.34	0.17 (0.22/0.11)	0.09
rs2981428	C/C 249 309	C/A 253 295	Ref. 0.81 [0.49–1.12] 0.86 [0.54–1.18]	— 0.18 0.35	0.18	A	1.05 [0.90–1.20]	0.54	0.38	0.41 (0.51/0.31)	0.48
rs7895676	A/A 188 325	T/T 179 379	Ref. 1.22 [0.97–1.48] 1.64 [1.33–1.94]	— 0.11 1.6×10^{-3}	0.48	C	1.28 [1.12–1.43]	1.7×10^{-3}	0.45	—	—
rs2981582	C/C 284 273	C/C 247 371	Ref. 1.56 [1.33–1.79] 2.04 [1.70–2.38]	— 1.6×10^{-4} 3.4×10^{-5}	0.95	T	1.46 [1.30–1.62]	2×10^{-6}	0.33	0.25 (0.30/0.20)	0.42
rs3135718	A/A 299 285	A/A 251 375	Ref. 1.57 [1.34–1.80] 1.91 [1.58–2.23]	— 10^{-4} 10^{-4}	0.65	G	1.43 [1.27–1.58]	6×10^{-6}	0.34	—	—
rs3135715	C/G 80 479	C/G 128 533	Ref. 1.06 [0.82–1.30] 1.18 [0.44–1.91]	— 0.61 0.67	0.68	C	1.07 [0.86–1.28]	0.53	0.15	0.24 (0.27/0.22)	0.23
rs4647913	C/C 600 41	G/G 525 36	Ref. 2.63 [0.33–4.94] 2.63 [0.36–4.89]	— 0.41 0.40	0.06	C	0.93 [0.51–1.35]	0.73	0.04	—	—

OR, odds ratio; CI, confidence interval; CA: coded allele; HWE: Hardy–Weinberg equilibrium.

compared with that in the Chinese ($P=0.003$) and the European ($P=0.01$) populations, and marginally lower ($P=0.06$) compared with the Japanese population.

We estimated pairwise LD for the seven SNPs using our data (Figure 1, Supplementary Table 1a and b). Similar to the HapMap results, in the Russian population, SNPs rs4647913, rs2981428 and rs10510097 were not contained in the intron 2 LD block that was previously reported to be associated with BC risk. According to HapMap, rs3135715 and rs2981582 were in LD ($r^2=0.44$ for Caucasians and

$r^2=0.77$ for Asians) and, therefore, rs3135715 might be located inside the same LD block. In our data the LD between SNPs rs3135715 and rs2981582 was lower ($r^2=0.24$).

Table 1 and Figure 1 show that a significant association was observed between the risk of BC and three SNPs: rs3135718 (G allele odds ratio (OR)=1.43, $P=6 \times 10^{-6}$), rs2981582 (T allele OR=1.46, $P=2 \times 10^{-6}$) and rs7895676 (C allele OR=1.28, $P=1.7 \times 10^{-3}$); this association was still highly significant when multiple testing corrections for the number of SNPs studied were made. On the basis of Akaike's information criterion, the additive model was the best for all three SNPs (Supplementary Table 2). The other four SNPs (rs4647913, rs3135715, rs2981428, rs10510097) did not show a significant association with BC (Table 1, Figure 1).

We have estimated the frequencies of haplotypes, including rs3135718, rs2981582 and rs7895676, in the case and control groups (Table 2). A statistically significant association was shown only for the GTT haplotype consisting of the risk alleles of each locus identified in a single-locus analysis (Table 2). Thereafter, we tested whether all three SNPs are necessary to explain the observed association. For this, we have contrasted models, including each one of the significant SNPs and all pairs of SNP to the general model, including all three SNPs (rs7895676, rs2981582 and rs3135718). Supplementary Table 3 shows that the model, including rs7895676, only was significantly ($P=0.003$) worse than the model, including all three SNPs, whereas models, including rs2981582 and rs3135718 (individually or together), did not show a significant deviation from the general model.

Discussion

In this work, we studied the association between seven *FGFR2* SNPs, including rs2981582 and rs7895676, and BC risk in the Russian population of cases and controls from the Altay region of Siberia, Russian Federation. The study was sufficiently powered to carry out a replication study; for example, the power to detect the association with rs2981582 at a P -value of 0.05 was 93%. Of the seven SNPs typed, four were present in HapMap. In our population,

Figure 1 Association analysis of SNPs across intron 2 of *FGFR2* locus. (a) Association between breast cancer and SNPs in the *FGFR2* region in the West Siberian population. (b) Linkage disequilibrium in the *FGFR2* region in the HapMap CEU population and (c) the West Siberian control population.

Table 2 Frequency of *FGFR2* intron 1 and intron 2 haplotypes in cases and controls and association with breast cancer risk

Haplotype			Frequency		OR (95% CI)	P
rs3135718	rs2981582	rs7895676	Controls	Cases		
A	C	T	0.53	0.47	1.00	
A	C	C	0.13	0.10	0.85 [0.66–1.08]	0.19
G	T	T	0.02	0.01	0.64 [0.31–1.29]	0.20
G	T	C	0.31	0.40	1.42 [1.21–1.68]	2.1×10^{-6}

OR, odds ratio; CI, confidence interval.

allelic frequencies were different from those observed in the European HapMap population, although the minor allele was always the same (Table 1). These differences may reflect the history of the Russian population, which geographically lies between Western Europe and Asia and is known to form the “eastern end” of the European geographic and genetic distributions.⁷ Although the LD pattern observed in our sample is similar to that observed in HapMap and other studies, the magnitude of LD in *FGFR2* intron 2 generally seems to be lower in our population compared with that in the HapMap European population.

Three SNPs—rs7895676[C] (OR = 1.28 (1.12–1.43), $P = 1.7 \times 10^{-3}$), rs2981582[T] (OR = 1.46 (1.30–1.62), $P = 2 \times 10^{-6}$) and rs3135718[G] (OR = 1.43 (1.27–1.58), $P = 6 \times 10^{-6}$)—were significantly associated with BC in our study. In the study by Easton *et al*,¹ the SNP rs7895676 located at the downstream boundary of intron 2 had a stronger association with BC risk than rs2981582 (odds > 100), which is located at the upstream boundary. On the contrary, we found that in our sample, rs2981582 explained the association much better than rs7895676 (odds > 1000). If we hypothesize that the actual causative variant lies in the LD block, including rs7895676 and rs2981582/rs3135718, the apparent contradiction may be explained by different LD structures in different populations. Furthermore, for rs2981582, the OR observed in the Russian population was significantly higher than that observed for Asians ($P = 0.02$), but did not differ significantly ($P = 0.1$) from the estimates obtained from European samples (see Supplementary Materials and methods). This finding may be explained by assuming that the LD structure in the Russian population resembles that of Europeans more closely than that of Asians, which is also consistent with recent studies of gene-geography.⁷

Meyer *et al*⁸ have recently shown that the haplotype marked by the minor allele of rs2981582 is associated with a higher level of *FGFR2* transcription both in cell lines and in tumors. It was concluded that two SNPs, rs7895676 and rs2981578, located in the transcription factor binding sites (C/EBP β and Oct-1/Runx2) could affect the binding efficiency of these factors. The authors hypothesized that this could lead to an alteration of the transcription level of the corresponding allele and proposed that rs7895676 and rs2981578 are the functionally significant SNPs. Our study suggested that rs7895676 could not explain all the observed associations. However, our results, although strong, are based on statistical evidence, and the functionality of rs7895676 cannot be excluded. Further functional

studies are required to elucidate the actual causative variant located at intron 2 of the *FGFR2* gene.

In summary, we have shown that SNP polymorphisms within the LD block in intron 2 of the *FGFR2* gene are associated with BC risk in the Russian female residents of West Siberia, Russia. SNPs rs2981582 and rs3135718, located at the upstream boundary of intron 2 had a stronger (odds > 1000) association with BC risk than rs7895676, which is located at the upstream boundary of the block. We conclude that rs7895676 is not likely to be the causative variant responsible for the association in the region. Our study underlines the utility of different populations—including Russian—to fine-map causative variants underlying the variation of complex traits.

Acknowledgements

We thank Ms M Udler, Professor Douglas F Easton and Professor Cornelia M van Duijn for useful discussion. We thank Julia N Dimitriadis, Inna A Selezneva, Svetlana A Terekhova and Elena L Sekerzhinskaya who referred patients to this study. This study was supported by the Siberian Branch of Russian Academy of Sciences and the joint grant from the Netherlands Organization for Scientific Research (NWO) and the Russian Foundation for Basic Research (NWO-RFBR, 047.017.043).

Conflict of interest

The authors declare no conflict of interest.

References

- Easton DF, Pooley KA, Dunning AM *et al*: Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007; **447**: 1087–1093.
- Hunter DJ, Kraft P, Jacobs KB *et al*: A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007; **39**: 870–874.
- Raskin L, Pinchev M, Arad C *et al*: *FGFR2* is a breast cancer susceptibility gene in Jewish and Arab Israeli populations. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 1060–1065.
- Liang J, Chen P, Hu Z *et al*: Genetic variants in fibroblast growth factor receptor 2 (*FGFR2*) contribute to susceptibility of breast cancer in Chinese women. *Carcinogenesis* 2008; **20**: 2341–2346.
- Aulchenko YS, Ripke S, Isaacs A, van Duijn CM: GenABEL, an R library for genome-wide association analysis. *Bioinformatics* 2007; **23**: 1294–1296.
- Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet* 2002; **70**: 425–434.
- Heath SC, Gut IG, Brennan P *et al*: Investigation of the fine structure of European populations with applications to disease association studies. *Eur J Hum Genet* 2008; **16**: 1413–1429.
- Meyer KB, Maia A, O'Reilly M *et al*: Allele-specific up-regulation of *FGFR2* increases susceptibility to breast cancer. *PLOS Biol* 2008; **6**: 1098–1103.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)