Haplotype analysis of chemokine CXCL12 polymorphisms and susceptibility to premature ovarian failure in Chinese women

Binbin Wang¹,²,†, Peisu Suo¹,²,†, Beili Chen³, Zhaolian Wei², Lu Yang³, Sirui Zhou¹,², Jing Wang¹,², Yunxia Cao³, and Xu Ma¹,²,4,*

¹Center for Genetics, National Research Institute for Family Planning, 12, Dahuisi Road, Haidian, Beijing 100081, China ²Peking Union Medical College, Beijing, China ³Reproductive Medicine Center, First Affiliated Hospital, Anhui Medical University, Hefei, China ⁴World Health Organization Collaborating Centre for Research in Human Reproduction, Beijing, China

*Correspondence address. Tel. +86-10-62176870; Fax: +86-10-62179151; E-mail: genetic@263.net.cn

Submitted on June 5, 2010; resubmitted on December 19, 2010; accepted on January 5, 2011

BACKGROUND: Chemokine (C-X-C motif) ligand 12 (CXCL12/stromal cell-derived factor 1) has been suggested to play an essential role in primordial germ cell migration, colonization and survival, and in the primordial to primary follicle transition. This study was performed to investigate an association of polymorphisms in CXCL12 with the risk of premature ovarian failure (POF) in Chinese patients.

METHODS: Tagging single nucleotide polymorphisms (SNPs) were selected using the Chinese HapMap database. Five SNPs (rs4948878, rs1801157, rs266087, rs266093 and rs1029153) were genotyped by direct sequencing in 111 patients with POF and 183 healthy controls recruited from the First Affiliated Hospital, Anhui Medical University, China.

RESULTS: Compared with controls, there were significantly higher frequencies of the rs1801157 A allele and haplotype C-T-A-T-T in cases with POF [P = 6.38E−07, odds ratio (OR) = 3.10, 95% confidence interval (CI) 1.955–4.890 by allele; P = 7.0E−04, OR = 2.39, 95% CI 1.43–3.97 by haplotype]. No differences were observed for the other four SNPs between POF cases and controls.

CONCLUSIONS: A strong association between a CXCL12 polymorphism and POF was established in Chinese patients, suggesting that CXCL12 might be a new candidate gene involved in POF. The A allele of CXCL12 polymorphism rs1801157 is a possible risk factor for developing POF. However, further independent studies are necessary to confirm our findings.

Key words: premature ovarian failure / CXCL12 / haplotype / rs1801157 / rs4948878

Introduction

Premature ovarian failure (POF) is a common disorder that leads to infertility, and is characterized by primary or secondary amenorrhea under the age of 40 years together with elevated plasma FSH (>40 IU/l) and decreased levels of estrogens. POF affects ~1% of women under the age of 40 years, 0.1% of women under the age of 30 years and 0.01% of women under the age of 20 years (Panay and Kalu, 2009).

The pathogenesis of POF is highly heterogeneous and a wide spectrum of pathogenic mechanisms potentially leads to POF. The incidence of familial cases among affected women with POF varies from 4 to 30% (Cramer et al., 1995; Torgerson et al., 1997), suggesting that POF has a strong genetic component. Linkage searches in two POF families identified regions on chromosome Xq21.1–q21.33 and 5q14.1–q15 (Lacombe et al., 2006; Oldenburg et al., 2008). Deletion analysis and molecular studies revealed that the POF candidate genes might be clustered on Xq26–q28 and Xq13.3–q22 (Fitch et al., 1982; Krauss et al., 1987; Tharapel et al., 1993; Powell et al., 1994). On the basis of candidate gene research approaches, an increasing number of autosomal candidate genes involved in POF have been reported, such as NOBAX9 (Qin et al., 2007), FIGLA (Zhao et al., 2008) and NR5A1 (Lourenco et al., 2009). However, mutation screening of the reported candidate genes only explains a small proportion of POF cases, indicating that other, as yet undiscovered, genes are linked to the development of POF (Laiissue et al., 2008). The
chemokine (C-X-C motif) ligand 12 gene (CXCL12), located on chromosome 10q11.1, encodes a member of the large family of chemokines that act through interactions with a subset of 7-transmembrane protein-coupled receptors (Bagni et al., 1997). Interactions between CXCL12 (also known as stromal cell-derived factor) and its receptor, CXCR4, have been suggested to play an essential role in primordial germ cell (PGC) migration, colonization and survival (Doitsidou et al., 2002; Knaut et al., 2003; Molyneaux et al., 2003; Stebler et al., 2004; Herpin et al., 2008). A subsequent study suggested that CXCL12 inhibits the primordial to primary follicle transition in the neonatal mouse ovary (Holt et al., 2006).

Recently, the first reported genome-wide association study (GWAS) performed in 99 unrelated Caucasian patients with idiopathic POF and 235 unrelated Caucasian female controls showed that polymorphism rs4948878 in CXCL12 has a suggestive association with POF (Knauff et al., 2009), indicating that CXCL12 might be a new candidate gene involved in POF.

The CXCL12 3′-untranslated region (UTR) polymorphism rs1801157 (G>A) has been reported to be associated with the risk of human immunodeficiency virus (HIV) infection, myocardial infarction and cancer progression in breast cancer (Modi et al., 2005; Lin et al., 2009; Luan et al., 2010). The A allele is hypothesized to be a possible cis-acting factor and might up-regulate the expression of CXCL12 (Winkler et al., 1998; Watanabe et al., 2003).

To investigate a possible association between CXCL12 polymorphisms and POF in the Chinese Han population, four tagging single nucleotide polymorphisms (SNPs) (rs1801157, rs266978, rs266093 and rs1029153), which capture common patterns of genetic variation in the CXCL12 gene, were selected in the present study. Moreover, SNP rs4948878 which has been reported to associate with POF in Caucasian women was chosen in our study. A total of five SNPs were genotyped in 111 patients with POF and 183 controls.

### Materials and Methods

#### Subjects

A total of 111 unrelated and well-characterized cases with POF and 183 normal controls were recruited from the First Affiliated Hospital, Anhui Medical University, China. All patients and controls were Chinese. The 111 patients (mean age at screening $31.1 \pm 6.4$ years, range 16–39) included 13 (11.7%) cases affected with primary amenorrhea and 98 (88.3%, mean age at menopause $26.1 \pm 6.9$ years) with secondary amenorrhea. The following criteria for diagnosing POF were used: secondary amenorrhea for >6 months together with repeated high levels of FSH (>40 mIU/ml) and low plasma estradiol levels. All patients with POF underwent karyotype analysis and those with chromosomal abnormalities were excluded. Women with associated endocrinopathies and autoimmune disorders were also excluded. All controls (mean age at screening $31.2 \pm 4.5 $ years, range 21–52) had proven fertility, with normal menstrual cycles and ovary morphology, and no history of subfertility treatment. Phenotype characteristics of cases and controls are listed in Table I. Informed consent was obtained from all participants, and the study was approved by the Ethics Committee of The National Research for Family Planning.

#### Methods and genetic analysis

Genomic DNA was extracted from peripheral blood using a TIANamp Blood DNA Kit (Tiangen, Beijing, China). A total of five SNPs were

---

### Table I

**Phenotype characteristics of cases with POF and healthy controls.**

<table>
<thead>
<tr>
<th>Group</th>
<th>Age range (years)</th>
<th>Primary amenorrhoea (n)</th>
<th>Secondary amenorrhoea (n)</th>
<th>Total (n)</th>
<th>Age at screening (years)</th>
<th>Age at menarche (years)</th>
<th>Age at amenorrhoea (years)</th>
<th>FSH (IU/l)</th>
<th>46 XX karyotype (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>16–39</td>
<td>13</td>
<td>98</td>
<td>111</td>
<td>31.1 ± 6.4</td>
<td>14.6 ± 2.0</td>
<td>26.1 ± 6.9</td>
<td>73.2 ± 31.3</td>
<td>100</td>
</tr>
<tr>
<td>Controls</td>
<td>21–52</td>
<td></td>
<td></td>
<td>183</td>
<td>31.2 ± 4.5</td>
<td>14.4 ± 1.3</td>
<td>26.1 ± 6.9</td>
<td>6.5 ± 2.1</td>
<td></td>
</tr>
</tbody>
</table>

Data are mean ± SD.

### Table II

**Results of association of single CXCL12 SNPs with POF under the additive model.**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Alleles</th>
<th>Risk allele</th>
<th>MAF cases (n = 111)</th>
<th>MAF controls (n = 183)</th>
<th>HWE cases</th>
<th>HWE controls</th>
<th>OR</th>
<th>95% CI</th>
<th>P-value (^b)</th>
<th>P-value from permutation (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs266093</td>
<td>C/G</td>
<td>C</td>
<td>0.18</td>
<td>0.216</td>
<td>0.30</td>
<td>0.27</td>
<td>1.25</td>
<td>0.82–1.91</td>
<td>0.30</td>
<td>0.96</td>
</tr>
<tr>
<td>rs1029153</td>
<td>C/T</td>
<td>T</td>
<td>0.167</td>
<td>0.227</td>
<td>0.95</td>
<td>0.86</td>
<td>0.68</td>
<td>0.44–1.05</td>
<td>0.08</td>
<td>0.44</td>
</tr>
<tr>
<td>rs1801157</td>
<td>A/G</td>
<td>A</td>
<td>0.252</td>
<td>0.098</td>
<td>0.12</td>
<td>0.14</td>
<td>3.09</td>
<td>1.96–4.90</td>
<td>6.37E–07</td>
<td>&lt;1.00E–5</td>
</tr>
<tr>
<td>rs266087</td>
<td>C/T</td>
<td>T</td>
<td>0.491</td>
<td>0.423</td>
<td>0.50</td>
<td>0.06</td>
<td>0.76</td>
<td>0.54–1.07</td>
<td>0.12</td>
<td>0.66</td>
</tr>
<tr>
<td>rs4948878</td>
<td>C/T</td>
<td>C</td>
<td>0.068</td>
<td>0.049</td>
<td>0.46</td>
<td>0.48</td>
<td>1.4</td>
<td>0.69–2.84</td>
<td>0.35</td>
<td>0.98</td>
</tr>
</tbody>
</table>

OR, odds ratio; CI, confidence interval; MAF, minor allele frequency; HWE, Hardy–Weinberg equilibrium.

\(^b\)-values were calculated using exact test method.

\(^c\)-values were calculated using \(^b\)-test.

\(^c\)-values were adjusted after 10⁴ permutation test.
selected and genotyped in 111 patients with POF and 183 controls. Four tagging SNPs (rs1801157, rs266087, rs266093 and rs1029153) which capture common patterns of genetic variation of CXCL12 (spanning 8 kb) were selected using the CHB HapMap database (http://www.hapmap.org). We used polymorphisms with a minor allele frequency >0.2 to assign unphased genotypic data into haplotypes. The tag algorithm, with a Pearson correlation $r^2$ threshold of 0.8, was employed to choose a set of tagging SNPs across the gene. Also, SNP rs4948878, which has been reported to associate with POF in Caucasian women, was selected for our study. DNA regions containing five selected SNPs (rs4948878, rs1801157, rs266087, rs266093 and rs1029153) were amplified from all patients and controls individually by PCR, using the primer set listed in Supplementary data, Table S1. The PCR products were directly sequenced on an ABI 3730XL automated sequencer (Applied Biosystems, Foster City, CA, USA).

The allelic and genotypic distributions of polymorphisms rs4948878, rs1801157, rs266087, rs266093 and rs1029153 were estimated by allele counting and compared in the POF and control groups by $\chi^2$ tests using the Statistical Package for Social Science Version 10.0 (SPSS 10.0).

Haplotype association tests were performed using HaploView4.2 software. Haplotypes with a frequency <0.05 in our study population were excluded from analysis owing to small sample size. Linkage disequilibrium (LD) was calculated using the absolute association ($r^2$). The frequencies closely agreed with results from a maximum-likelihood method implemented via an expectation-maximization algorithm. Permutation tests were performed to exclude false positives related to multiple tests using HaploView4.2 software. An exact test was applied to estimate the Hardy–Weinberg equilibrium (HWE) of all polymorphisms. Power calculations were performed using the online resource (http://pngu.mgh.harvard.edu/~purcell/gpc/#cc_ins). Statistical significance was set at $P < 0.05$.

## Results

The distribution of alleles of the five SNPs in the study groups is shown in Table II. All SNPs are in HWE in the case and control groups with a significance threshold of 0.05. All five SNPs were estimated under an additive model (Table II). A strong association was established between polymorphism rs1801157 and POF in our Chinese population.

Compared with the controls, the occurrence of the A allele was higher in cases ($P = 6.38E−07$ odds ratio (OR) = 3.10, 95% confidence interval (CI) 1.955–4.890). The association still reached significance after $10^7$ permutation tests were performed ($P < 10E−5$). There were no differences for the other four SNPs (rs266087, rs4948878, rs1029153 and rs266093) between cases and controls under an additive model (Table II). Association was estimated under genotypic, recessive and dominant models (Supplementary data, Table S2). A significant difference for SNP rs266087 was observed between cases and controls under the dominant model ($P = 0.03$, OR = 1.80, 95% CI 1.06–3.05).

LD was present, but no strong LD was observed between any two SNPs (Fig. 1). The distributions of five common (frequency $\geq 0.05$) haplotypes of the five SNPs are listed in Table III. Significant differences in the counts of haplotypes C-T-A-T-T and C-C-G-C-T between cases and the control group were observed; however, an association was only established for C-T-A-T-T after the permutation test ($P = 0.0029$).

We compared the LD patterns between rs1801157 and rs4948878 in different populations, based on information in the HapMap database. The LD pattern between the two SNPs was weak in both the HCB (Han Chinese in Beijing, China) sample ($r^2 = 0.051$) and the

### Table III Haplotype distribution of the CXCL12 polymorphisms in cases with POF and controls.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Case (%)</th>
<th>Control (%)</th>
<th>OR*</th>
<th>95% CI</th>
<th>P-valueb</th>
<th>P-value from permutationc</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-T-G-T-T</td>
<td>0.265</td>
<td>0.276</td>
<td>0.95</td>
<td>0.65–1.38</td>
<td>0.7659</td>
<td>I</td>
</tr>
<tr>
<td>C-T-C-G-T</td>
<td>0.207</td>
<td>0.234</td>
<td>0.85</td>
<td>0.57–1.28</td>
<td>0.4476</td>
<td>0.999</td>
</tr>
<tr>
<td>C-T-A-T-T</td>
<td>0.174</td>
<td>0.081</td>
<td>2.39</td>
<td>1.43–3.97</td>
<td>7.00E−04</td>
<td>0.0029</td>
</tr>
<tr>
<td>G-T-G-C-T</td>
<td>0.108</td>
<td>0.115</td>
<td>0.94</td>
<td>0.55–1.59</td>
<td>0.8021</td>
<td>I</td>
</tr>
<tr>
<td>C-C-G-C-C</td>
<td>0.068</td>
<td>0.133</td>
<td>0.47</td>
<td>0.26–0.86</td>
<td>0.0134</td>
<td>0.0745</td>
</tr>
</tbody>
</table>

Other haplotypes with frequency <0.05 are: C-T-A-C-T, G-C-C-G-C, C-C-G-C-T, G-C-G-C-T, G-T-G-T-T and C-C-G-C-C.

*Compared with all other haplotypes.

b$P$-values were calculated using $\chi^2$ test.

c$P$-values were adjusted after $10^7$ permutation test.
Chemokine (CXCL12) polymorphisms and ovarian failure

Discussion

CXCL12 has been suggested to play an essential role in PGC migration, colonization and survival, and in the primordial to primary follicle transition. Five SNPs (rs1801157, rs266087, rs266093, rs1029153 and rs4948878) were selected to investigate the possible association between CXCL12 polymorphisms and POF in the Chinese Han population.

Interestingly, we identified a strong association between the 3'UTR functional polymorphism rs1801157 (G>A) and POF (P = 6.38E−07). The association still reached significance after 10^4 permutation tests (P < 10E−5). Accordingly, haplotype C-T-A-T-T carrying the rs1801157 A allele was associated with POF (P = 7.0E−04). These results suggest that the 3'UTR A polymorphism might be a risk factor contributing to POF in the Chinese population. This is consistent with a previous GWAS study performed in Caucasians which suggested an association for CXCL12 polymorphism rs4948878 (Knauth et al., 2009), and therefore both studies suggest that CXCL12 might be a candidate gene involved in POF.

However, we failed to establish an association for rs4948878 in our Chinese population despite having >80% power to detect the previously reported association. Polymorphisms rs1801157 and rs4948878 are in different LD, as the LD between rs4948878 and rs1801157 was weak in both the HapMap HCB sample (r^2 = 0.051) and the HapMap CEU sample (r^2 = 0.013). It is possible that CXCL12 is associated with POF via different SNPs in different ethnic groups, and the susceptible site contributing to POF could show a different LD in the two populations. Further studies are warranted to investigate the discrepancy in different populations.

The 3'UTR polymorphism A allele has been hypothesized to serve as a target for cis-acting factors influencing transcript abundance, synthesis, transport, stability or splice product abundance, and was reported to induce increased CXCL12 protein levels as a result of enhanced mRNA stability (Winkler et al., 1998; Watanabe et al., 2003; Garcia-Moruja et al., 2009). Importantly, polymorphism rs1801157 has been reported to be associated with certain diseases. The 3'UTR A allele was identified as a protective factor against HIV infection and was recently reported to decrease risk for myocardial infarction in a Chinese population (Modi et al., 2005; Luan et al., 2010). In contrast, the A allele was reported to be associated with an increased risk of breast and lung cancer (Razmkhah et al., 2005a,b).

CXCL12 and its receptor CXCR4 are essential in the development of PGCs. PGCs were randomly scattered in zebrafish embryos in which expression of CXCL12 and CXCR4 were blocked, suggesting this signal provides a key directional cue for PGC migration (Doitsidou et al., 2002; Knaut et al., 2003). In mouse embryo, the interaction of CXCL12 and CXCR4 plays a role in homing PGCs to genital ridges. In CXCL12/−/− mutant mice, PGCs were decreased and displayed a delayed migration (Molyneaux et al., 2003; Stebler et al., 2004). A similar signaling axis was established in chicks and Medaka (Japanese killifish) (Stebler et al., 2004; Herpin et al., 2008). Importantly, the addition of exogenous CXCL12 causes aberrant germ cell migration in mice. A subsequent study suggested that the high level of CXCL12 in the culture resulted in higher primordial follicle densities and follicles of smaller size, indicating that the chemokine was acting as an inhibitor of follicle activation (Holt et al., 2006).

Defects in migration or proliferation of primary germ cells, the founders of the sperm or oocytes, have been previously proposed as a cause of idiopathic POF (Rebar, 1982; Razmkhah et al., 2005a,b). Primary germ cell deficiency has been documented as a model of POF (Duncan et al., 1993).

Therefore, a plausible supposition might be that patients with POF carrying the rs1801157 A allele have an increased expression of CXCL12, which results in aberrant PGC migration and an inhibition of transition from primordial to primary follicle, both of which were proposed to lead to POF. Further investigations are needed to elucidate the role of CXCL12 in the pathogenesis of POF.

In conclusion, our study established a strong association between CXCL12 polymorphisms and POF in a Chinese population. The CXCL12 polymorphism rs1801157 A allele might increase the risk of developing POF. However, rs4948878, an SNP implicated in POF in Caucasian populations, might not be a potential marker in the Chinese population. Of course, in our study, the sample size is small and further studies are necessary to confirm the association of polymorphism rs1801157 with a risk of POF, as well as to elucidate the discrepancy observed in results from different populations.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgements

The authors thank the family and all subjects for taking part in this study. Professors Xu Ma (genetics@263.net.cn) and Yunxia Cao (caoyunxia6@126.com) contributed equally to the research project and can be considered co-corresponding authors.

Funding

This work was supported by the National Basic Research Program of China (2010CB945100), the National Natural Science Foundation of China (30973197), the National Science & Technology Pillar Program of China (No.2008BAH24B05), National Basic Research Program of China (2007CB947400), the National Infrastructure Program of Chinese Genetic Resources (2006DKA21300).

References

Fitch N, de Saint Victor J, Richer CL, Pinsky L, Sitahal S. Premature menopause due to a small deletion in the long arm of the


