The Impact of Heat Shock Protein 70 Gene Variations on Clinical Presentation and Outcome in Schizophrenic Inpatients

Chi-Un Pae\textsuperscript{a,c}  Antonio Drago\textsuperscript{d}  Jung-Jin Kim\textsuperscript{b}  Laura Mandelli\textsuperscript{d}

Diana De Ronchi\textsuperscript{d}  Alessandro Serretti\textsuperscript{d}

\textsuperscript{a}Department of Psychiatry, Holy Family Hospital, Catholic University of Korea College of Medicine, Bucheon, and \textsuperscript{b}Department of Psychiatry, Kangnam St. Mary’s Hospital, Catholic University of Korea College of Medicine, Seoul, Republic of Korea; \textsuperscript{c}Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, N.C., USA; \textsuperscript{d}Institute of Psychiatry, University of Bologna, Bologna, Italy

Key Words
Heat shock proteins · Schizophrenia · Gene association

Abstract
We previously investigated a group of single-nucleotide polymorphisms of a set of genes coding for heat shock proteins (HSPA1A, HSPA1B and HSPA1L) and found a significant association between one HSPA1B variation and schizophrenia (SZ). We now report an association between a set of variations (rs2227956, rs2075799, rs1043618, rs562047 and rs539689) within the same genes and a larger sample of schizophrenic inpatients. A single variation, rs539689 (HSPA1B), was found to be marginally associated with Positive and Negative Syndrome Scale (PANSS) positive scores at discharge, and haplotype analysis revealed a significant association between improvement in PANSS scores with both A-C-G-G and A-C-G-G haplotypes. These findings further support a role of heat shock proteins in the pathophysiology of SZ.

Introduction
We previously reported an association between the heat shock protein (Hsp) gene HSPA1B variation rs1061581 (p = 0.026) and schizophrenia (SZ) [1]. The aim of this paper is to investigate a larger set of mutations (rs2227956, rs2075799, rs1043618, rs562047 and rs539689 covering 3 Hsp genes: HSPA1A, HSPA1B and HSPA1L) in a sample of 240 schizophrenic Korean inpatients.

Hsps represent a protein family dedicated to housekeeping and stress-induced cellular activities: both actions are partially depending on their chaperone ability, as they can bind and stabilize otherwise unstable protein, facilitating its correct fate in vivo, be it folding, oligomeric assembly, transport to a particular subcellular compartment or disposal to degradation (reviewed in Hartl [2]). Moreover, Hsps are involved in the uncoating of clathrin-coated vesicles, in mitochondrial activity and in the control of regulatory proteins [2–6]. Indeed, they...
probably also regulate other than chaperone functions (reviewed in Daugaard et al. [7]), such as the stabilization of the lysosomal membrane [8, 9]. Accordingly, Hsps may be appropriate targets for modulating cell death pathways [6], and the chaperone function is thought to be so relevant that it is likely that chaperone diseases exist [10]; of note, SZ might be one of these.

Hsps are classified into families on the basis of sequence homology and typical molecular weight as Hsp 110, Hsp 100, Hsp 90, Hsp 70, Hsp 40, Hsp 10 and small Hsp families. Within this group, the 70-kDa Hsp is one of the most relevant, mainly involved in the inducible cellular response to stress, and associated with one of the critical chaperone functions: the recovery of ribosomal proteins [11]. There are at least 8 unique gene products belonging to the human Hsp 70 family (HSPA1A, HSPA1B, HSPA1L, HSPA2, Hsp70-5/Bip, HSPA6, HSPA8, HSPA9) [7, 12]: Hsp 70-5 and Hsp 70-9 are confined to the lumen of the endoplasmic reticulum and of the mitochondrial matrix, respectively, others reside mainly in the cytosol and in the nucleus. The most relevant homologues are Hsp 70-1a, -1b and -1t, respectively, coded by genes HSPA1A, HSPA1B and HSPA1L. HSPA1A and HSPA1B are closely linked, stress-inducible and intronless genes localized in the major histocompatibility class III cluster between the complement and tumor necrosis factor locus on the short arm of chromosome 6, which is also the localization of the intronless HSPA1L [13, 14]. Of note, association studies identified the chromosome 6p as one of the most replicated vulnerability areas for SZ [for reviews, see 15–18]. More in detail, HSPA1A is located in position 6p21.33 and it spans about 2.4 kb. No introns can be found in this gene as well as in the HSPA1B which is coded in the same genetic position and which also spans about 2.4 kb. Indeed, these two genes code for two 99% identical proteins, and the HSPA1L (position 6p21.33, it spans about 5.5 kb) codes for an Hsp 70 which is 91% identical to Hsp 70-1a [7]. More genetic information is provided by international databases (http://www.genecards.org/index.shtml). A genetic perspective will be used in this study to verify if these mutated genes are influential on clinical presentation and drug response of a group of schizophrenic inpatients.

Patients and Methods

Sample Description

The sample was composed by 240 schizophrenic inpatients with an almost equal gender distribution (females = 55%). At the moment of hospitalization, the mean age was 34.8 ± 12.2 years, while age at onset was 23.3 ± 6.5 years. Duration of illness varied widely in the sample, 20 and 3 being the superior and inferior quartiles, respectively, and the standard deviation was found to be as large as 13.2 years. The number of hospitalizations ranged from once to 25 times in a lifetime (average 3.0 ± 2.7), and 46 subjects (19%) attempted suicide. Different antipsychotic drugs were used as treatment: 109 (45%) patients were treated with haloperidol, 67 (28%) with risperidone, 49 (20%) with trifluoperazine, 14 (6%) with clozapine, 9 (3%) with chlorpromazine. One patient was treated with pimozide and another with perphenazine. Chlorpromazine equivalents were calculated to perform a comparative analysis [19, 20]: during treatment, the average dose was 429.3 ± 191.9 mg (chlorpromazine equivalents). At the moment of discharge, the average dose was 574 ± 268.9 mg (chlorpromazine equivalents). Inclusion criterion was a diagnosis of SZ assessed by a Structured Clinical Interview, DSM-IV axis I disorders-clinical version, and comorbid axis I disorder other than SZ together with neurological and current medicosurgical illness were used as exclusion criteria. To assess symptom severity, patients were administered the Clinical Global Impression (CGI) [21] scale and the Positive and Negative Syndrome Scale (PANSS) [22, 23]. Scales were administered at the moment of hospitalization and at the moment of discharge, with an average period of time of about 1 month between the two assessments. Expert psychiatrists who worked outside the hospital and were blind to the genotype of patients were used as raters. Interrater evaluation among raters gave reliable results (κ > 0.80). All subjects were biologically unrelated, native Korean descendants residing in Korea. Written informed consent was provided by the subjects after the purpose and method of the study had been explained. The institutional review board of Kangnam St. Mary’s Hospital approved the study that was conducted in accordance with the Declaration of Helsinki. Of note, the Korean population is ethnically homogeneous [24]; thus, the likelihood of hidden stratification is low.

Genetic Analyses

Genomic DNA was extracted from blood by standard methods and quantified. The high-throughput genotyping method using a pyrosequencer (Biotage AB, Uppsala, Sweden) was used for genotyping 5 single-nucleotide polymorphisms (SNPs, i.e., rs2075799, rs1043618, rs562047, rs539689, rs2227956) of HSP genes, which were selected based on a public database (National Center for Biotechnology Information, dbSNP, http://www.ncbi.nlm.nih.gov/SNP/). PCR primers (Bioneer, Daejeon, Korea) and sequencing primers (Bioneer) used for the pyrosequencing assay were designed by using the Pyrosequencing Assay Design Software version 1 (Biotage AB), and 1 primer of each primer set was biotinylated.

Statistical Analyses

Haploview 3.2 was used to generate a linkage disequilibrium (LD) map and to test for Hardy-Weinberg equilibrium. Single genotype associations with PANSS and CGI scores were analyzed by analysis of variance; when including covariates or other factors, analysis of covariance and multivariate analysis of variance/covariance were employed. Baseline scores were included as covariates, along with the clinical variables associated with genotypes. Associations with other clinical variables in the sample were performed by analysis of variance or the χ2 test: the former...
to test continuous and the latter nominal variables. To limit the rate of first-time error, Bonferroni corrections were used for every positive result. The R software (Programming Environment for Data Analysis and Graphics version 2.2.1) was employed to analyze haplotype with both discrete and continuous traits and to include covariates. Permutation (50,000 permutations) was used to estimate the global significance of the results for haplotype analyses to confirm the expectation-maximization values. Only haplotypes with >1% prevalence were included in the analysis. Results were considered significant with an alpha level lower than 0.05. With this level of significance we had a power of 0.80 to detect an effect size $d = 0.18$ which corresponded to a detectable difference of about 2% in the PANSS scores between two main genotype variants [25].

**SNP Choice**

SNPs were chosen on the basis of the previous experience and significant findings of our group, along with some evidence coming from the nonpsychiatric literature. Moreover, one additional variation not closely linked to previous findings was also studied. The choice of SNPs is a highly debated issue with supporters of inductive or deductive models and supporters of whole-gene coverage [26–31]. Accordingly, rs2227956 is coding nonsynonyms, and it was found to be predictive of poor cellular response to stress [32] together with rs1043618 which runs in the 5'-untranslated region functional zone. rs2075799 is coding synonyms, we previously found a positive significance in a sample of depressed patients [33], as well as for the rs562047 which is a missense mutation and was found to be associated with SZ [1]. Moreover, it is reported to be associated with a propeptide cleavage site (http://snpeffect.vib.be/index.php). Finally, rs539689 is coding a synonymous variation.

**Results**

**Clinical Findings**

CGI baseline scores were 5.50 ± 0.65, and this was reduced by 45% by the end of treatment (3.01 ± 0.55). The PANSS global score was 93.81 ± 14.12 at baseline and 76.70 ± 9.00 at discharge. Positive and negative subscales of the PANSS were 24.81 ± 4.86 and 19.73 ± 4.06, respectively, at the beginning of the treatment, and 21.47 ± 5.16 and 20.22 ± 4.22 by the end of the treatment. Patients treated with typical antipsychotics had a better outcome when compared with patients treated with atypical antipsychotics (difference in PANSS positive score at discharge $p = 0.002$), probably because atypical antipsychotics were given to patients with a history of antipsychotic treatment resistance. Differences in time of illness are consistent with this. A reduction of CGI scores was found to be nonsignificant at the end of treatment ($p = 0.32$), whereas PANSS scores were found to be significantly decreased from the beginning to the end of treatment ($p < 0.0001$). Age at onset was found to be significantly and directly associated with CGI baseline scores ($p = 0.04$; $\rho = 0.13$), whereas it was significantly and inversely associated with CGI discharge scores ($p = 0.013$; $\rho = -0.15$).

**SNP Findings**

All genetic markers were in Hardy-Weinberg equilibrium in the sample, and rs2227956 and rs2075799 were found to be in complete LD; rs2227956 and rs1043618 are in complete LD; rs1043618 is in strong LD with rs562047 and with rs2075799. D’ values are shown.

![Fig. 1. Haplotype analysis. rs2227956 and rs2075799 are in complete LD; rs2227956 and rs1043618 are in complete LD; rs1043618 is in strong LD with rs562047 and with rs2075799. D’ values are shown.](image-url)
the treatment was entered in the analysis as a covariate, this association dropped to nonsignificance. When other variables were analyzed as covariates (age, age at onset, sex, chlorpromazine equivalents at the moment of discharge), the relationship between rs539689 and PANSS positive scores at discharge remained significant. On the other side, when clinical assessments at baseline (CGI at baseline and PANSS scores at baseline) were put in the statistical model as covariates, the significance of the above-mentioned relationship was lost again. The test revealed no significant association with each SNP and gender or suicide variables. No other significant associations were found for the other SNPs.

**Haplotype Findings**

As far as haplotypes are concerned, the main result is an association between the A-C-G-G haplotype and the difference of the scores of PANSS total and negative subscales from baseline to discharge (the haplotype was associated with the greatest difference), and an association between haplotype G-C-C-C and the difference of the PANSS total score from baseline to discharge (haplotype associated with the smallest difference). Covariate analysis showed that age, sex, duration of illness and drug class did not influence the significance of the above-mentioned associations. On the other hand, when age at onset was added to the model as a covariate, the association between haplotype A-C-G-G and the difference of the PANSS total score from baseline to discharge became nonsignificant. However, other results emerged: the A-C-G-G haplotype and PANSS negative scores at baseline (p = 0.0001, haplotype associated with highest scores), and G-G-G-G haplotype and CGI scores at discharge (p = 0.01, haplotype associated with lowest scores) as shown in tables 1 and 2. Finally, covariation with chlorpromazine equivalents during treatment made the association between PANSS total score difference from baseline to discharge and haplotype A-C-G-G nonsignificant, and covariation with PANSS total score at baseline made all the mentioned associations nonsignificant.

**Table 1. Haplotype analysis with no covariate**

<table>
<thead>
<tr>
<th>Test</th>
<th>Overall model significance p</th>
<th>Haplotype frequency</th>
<th>Haplotype score p</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-PN</td>
<td>0.005</td>
<td>A-C-G-G 0.01</td>
<td>2.42</td>
</tr>
<tr>
<td>D-PT</td>
<td>0.004</td>
<td>A-C-G-G 0.01</td>
<td>2.20</td>
</tr>
</tbody>
</table>

D-PN = Difference between scores of negative PANSS at baseline and at discharge; D-PT = difference between scores of PANSS at baseline and at discharge.

**Table 2. Haplotype analysis with age at onset as covariate**

<table>
<thead>
<tr>
<th>Test</th>
<th>Overall model significance p</th>
<th>Haplotype frequency</th>
<th>Haplotype score p</th>
</tr>
</thead>
<tbody>
<tr>
<td>PN-B</td>
<td>&lt;0.001</td>
<td>A-C-G-G 0.01</td>
<td>3.81</td>
</tr>
<tr>
<td>CGI-D</td>
<td>0.001</td>
<td>G-G-G-G 0.01</td>
<td>2.54</td>
</tr>
<tr>
<td>D-PN</td>
<td>0.02</td>
<td>A-C-G-G 0.01</td>
<td>2.00</td>
</tr>
<tr>
<td>D-PT</td>
<td>n.s.</td>
<td>A-C-G-G 0.01</td>
<td>2.13</td>
</tr>
</tbody>
</table>

D-PN = Difference between scores of negative PANSS at baseline and at discharge; D-PT = difference between scores of PANSS at baseline and at discharge; PN-B = negative PANSS score at baseline; CGI-D = CGI score at discharge; n.s. = not significant.

**Discussion**

The pathophysiology of SZ is not known so far. A neurodegenerative or neurodevelopmental path, a disruption in the autoimmune system as well as gene-environment interactions involving putative infective agents are current shared theories [34–46]. Thanks to their role in the viability of cells as a constitutive and induced mechanism facing adverse environmental conditions, and because of playing as cofactors in the antigen presentation mechanism (reviewed in Javid et al. [47]), Hsps represent good candidates in the study of the molecular mechanisms underlying SZ. Consistently, it has recently been reported that Hsp 72 (alias Hsp 70; http://www.genecards.org/index.shtml) protects against the neuronal hippocampal apoptosis rate in mice exposed to a model of chronic psychological stress [48], probably also via activation of the astrocytes, which have been demonstrated to enhance the synthesis of Hsp 70 under stress conditions and to release increasing amounts of Hsp 70 into the extracellular environment following stress [49, 50]. Furthermore, it has been demonstrated that induced Hsp activity after administration of arimoclomol was associated with a reduced neurodegenerative progression in animal models for amyotrophic lateral sclerosis [51]. Finally, Hsps have been demonstrated to be defensive against neurotrophic viruses [52]. Thus, we hypothesized that mutations running through the Hsp coding sequence might have an
influence on the clinical presentation and response to treatment of schizophrenic patients. Here we report that the polymorphism rs539689 at HSPA1B was found to be associated with PANSS positive scores at discharge. HSPA1B is one of the intronless genes that code for Hsp 70, and its product mediates the folding of newly translated proteins in the cytosol and in organelles. It is also involved in the ubiquitin-proteasome pathway. This finding is consistent with our previous association study about associations between Hsp and SZ [1], and with other findings in the literature, both focused on SZ [53, 54] and on other neurological disorders, such as the positive correlation between HSPA1B variations and noncognitive symptoms in patients suffering from Alzheimer’s disease [55]. This further evidence of association between SZ and HSPA1B strengthens the hypothesis that this Hsp might be involved in the pathophysiology of SZ. Nonetheless, it must be considered that this finding might be due to the LD reported between tumor necrosis factor (which refers to a group of cytokines that are associated with cellular apoptosis, their involvement in SZ has been repetitively reported [56–58]) and HSPA1B [59], or between HSPA1B and other relevant genes located in chromosome 6p. On the other hand, the genetic variation, albeit synonymous, might lead to secondary mRNA structure disruption, as well as genomic and posttranslational modifications possibly affecting the schizophrenic phenotypes.

As a second interesting result, haplotypes A-C-G-G and G-C-G-G were shown to be associated with a better clinical outcome as measured by PANSS total and negative scores, the A-C-G-G haplotype being associated with a greater score difference from baseline to discharge and the G-C-G-G haplotype with smaller ones. Interestingly, the association between A-C-G-G and negative PANSS subscale improvement remained significant after covariation with age at onset, while the association with PANSS negative scores at baseline became significant. This might underline that this haplotype in our sample is particularly sensitive to ‘age at onset’ as a variable, especially as far as the positive symptoms are concerned. Moreover, positive and negative symptoms seem to have an opposite behavior toward this genetic makeup, the former relying on the ‘age at onset’, and the latter being shadowed by that variant at baseline assessments. It is speculative whether such an opposite manner is suggestive of an Hsp mutation matching role in negative and positive phenotypes which might be consistent with the molecular role of these proteins: mutations leading to defective orchestration of Hsp activities might lead to a reduced viability and synaptic functions (reviewed in Buccellato et al. [52]) associated with a negative phenotype, whereas an imbalanced Hsp orchestration echoes the theory of an imbalanced aminergic tone in SZ brains (glutamatergic being down- and dopaminergic upregulated). When the effect of baseline PANSS scores is cleared from the statistical model using it as a covariate, all the associations become nonsignificant. This is suggestive of a small impact of the drug treatment alone in the improvement of these patients, and it may highlight that Hsp deregulation is closer to the pathophysiology of SZ than to its pharmacogenetics. Again, these considerations remain speculative, and more studies on this topic are needed. Covariate analysis also highlights the association between the G-G-G-G haplotype and lower CGI scores at discharge. This significant association was revealed only after the variable ‘age at onset’ was used as a covariate. It may be hypothesized that the CGI scores are sensitive to the negative symptom profile (the higher the negative symptoms, the lower the CGI scores): in this way, the result of the covariate analysis with the CGI scores at discharge would be consistent with the above-mentioned hypothesis about the influence of Hsp on negative and positive symptoms. Finally, chlorpromazine equivalents during treatment also proved to be associated with the significance of the statistical model, when considered as covariates; this event appears to be reasonably consistent with the findings concerning the PANSS baseline scores.

The main limitations of the present study are the low significance levels, the lack of a standardized treatment design, the incomplete coverage of all the Tag SNPs within the investigated sequences, the presence of possible stratification factors in the sample and the absence of a control group. In particular, a retrospective design and the presence of multiple antidepressant treatments may represent relevant limitations to the study, even though their impact has been limited by the use of the statistical analysis. Further, the main findings of our paper rely on rare haplotypes, which limits the current findings both in the direction of small overall impact and of a possible false-positive finding. In fact, it has been reported that an astonishing number of positive findings in genetic research could be false positive [60]. Therefore our marginal and at times vanishing results could very likely be false-positive findings. However, the power of our sample was sufficient to detect quite small effect sizes. Further, even if the pharmacogenetic design was not standardized, it is closer to everyday clinical practice; in this case, the lack of positive findings could be a false-negative result instead.
In conclusion, the results we report somewhat replicate the positive finding of an association between HSPA1B variation and SZ, and the identification of influencing haplotypes for SZ symptomatology. More studies investigating all the genetic variants, including copy number variations, and limiting the stratification bias by a family-based association approach are warranted.

Acknowledgments

This work was supported by a grant from the Medical Research Center, Korea Science and Engineering Foundation, Republic of Korea (R13-2002-005-04001-0), and the Catholic University of Korea College of Medicine Alumni Award of the year 2008.

References