Investigation of Copy Number Variations and micro-RNA Binding Site Polymorphisms

Doctoral Theses

Réka Kovács-Nagy MD

Semmelweis University
Doctoral School of Molecular Medicine

Supervisor: Zsolt Rónai MD assistant professor, Ph.D.

Official reviewers: Attila Patócs MD associate professor, Ph.D.
Gábor Speer MD specialist for internal medicine, Ph.D.

Head of the Final Examination Committee: Ilona Kovalszky professor, D.Sc.
Members of the Final Examination Committee: György Hegyi professor emeritus, D.Sc.
László Tretter professor, D.Sc.

Budapest
2013

Introduction

Both genetic and environmental factors play a crucial role in the development of complex traits and disorders. One of the most common approaches for the analysis of the inheritance of these phenotypes is the investigation of genetic polymorphisms, such as single nucleotide polymorphisms (SNP) as putative genetic components. SNPs localized in the 3' untranslated region (UTR) of candidate genes are of molecular importance, as they might alter the binding site of micro-RNAs thus influencing micro-RNA-based regulation (miRSNPs). An other group of polymorphisms that has recently been more thoroughly investigated is a length variation of sub-chromosomal level referred to as copy number variation (CNV).

As the more detailed study of both CNVs and miRSNPs has recently been initiated, several technical issues emerged regarding their analysis. One of the major aim of our work was the elaboration of reliable and efficient techniques for the analysis of CNVs and miRSNPs. The designed new techniques were applied for the molecular and association analysis of putative miRSNPs of the *wolframin* gene (*WFS1*) and those of the *SNAP-25* gene, as well as for the study of a CNV involving the *glycogene synthase kinase B* (*GSKB*) gene.

The base of the PhD work is an interdisciplinary collaboration. The psychological and psychiatric characterization of the participants was carried out by psychologist and psychiatrist coworker specialists. Thus, in these theses more emphasis was put on the molecular and genetic aspects of the study, information about the phenotype analysis has only been summarized briefly.

Aims

Two novel polymorphism types were analyzed. CNVs have only been described recently, one of our targets belonged into this group affecting the *GSK3B* gene region. The other group is of molecular functional interest: miR-SNPs putatively influencing the expression of *SNAP-25* and *WFS1* genes were also investigated. The aims of the PhD theses were as follows:

1. Previous studies have suggested that a CNV polymorphism on chromosome 3 involving the *Nr1I2* and *GSK3B* genes were in association with unipolar depression, results were however controversial. Our aim was the more accurate characterization of the exact chromosomal localization as well as the frequency of the **CNV affecting the *GSK3B* gene in patients with uni- and with bipolar depression as well as in a healthy control group.**

2. As the SNAP-25 protein plays a role in exocytosis, polymorphic variants of the gene have been suggested to be putative risk factors of several complex neuro-psychiatric and psychological phenotypes. Consequently **SNPs were aimed to be identified in the 3' UTR of the *SNAP-25* gene that might influence microRNA based regulation. Moreover we wanted to elaborate molecular genetic methods for the efficient and reliable genotype analysis of these variants.**

3. Although the monogenic disease (Wolfram-syndrome) caused by the loss-of-function mutations in the *WFS1* gene has been characterized in detail, less data are available about the role of the more common polymorphisms of the gene. Thus **we aimed the identification of miR-SNPs in the *WFS1* gene by *in silico* approaches, as well as the molecular and genetic analysis.** Moreover we decided to extend this study for the investigation of the putative association between these sites and psychological traits (aggression, impulsivity, anxiety, depression).

Methods

432 patients suffering of uni- or bipolar depression as well as 801 control patients took part in the study. Diagnosis of the uni- and bipolar depression was decided based on the criteria of DSM (Diagnostic and Statistical Manual of Mental Disorders) IV in the Clinic of Psychiatry and Psychotherapy and in the Clinical and Research Mental Hygiene Department of the "Kútvölgyi" Clinic Center of the Semmelweis University. Participants signed informed consent, the study was approved by the Scientific and Research Ethics Committee of the Medical Research Council. The Hospital Anxiety and Depression Scale (HADS), the Buss–Perry Aggression Scale and the Barratt Impulsivity Scale were used for the psychological characterization of the participants.

DNA sampling was carried out by non-invasive method collecting buccal smear using a swab. DNA purification was initiated by the precipitation of proteins using "salting-out" approach, which was followed by the precipitation of DNA by ethanol and isopropanol using standard protocols. CNV analysis and the genotype determination of the miRSNPs was carried out using fluorescent sequence specific TaqMan® probes in combination with real-time polymerase chain reaction technique. Some experiments were carried out by the high throughput TaqMan® OpenArray® genotyping system, which works using the same chemistry in a miniaturized and highly efficient manner. Functional analysis of the microRNA binding sites was carried out using pGL3 and pMIR Report vectors, subcloning the whole 3' UTR region of the genes of interest behind the gene of luciferase. DNA-constructs carrying the different allelic variants of the analyzed SNPs were generated by site directed mutagenesis, transfection experiments were carried out using HEK293T cell line. Relative luciferase activity was determined by normalizing the measured luminometric data by β-galactosidase transfection control.

Abbreviations

SNP: single nucleotide polymorphism, miRSNP: SNPs altering the binding site of microRNAs, CNV: copy number variation, GSK3B: glycogene synthase kinase 3B, WFS1: wolframin gene, SNAP-25: synaptosomal-associated protein, 25kDa

Results

Case–control analysis of depression and the CNV polymorphism affecting the GSK3B

Recent studies have demonstrated that a CNV polymorphism including the *GSK3B* gene might be the genetic component of uni- and bipolar depression, results are however controversial. 410 patients suffering of depression and 410 health controls were included in the study and **statistically significant association (p = 0.00001) was demonstrated between *GSK3B* copy number variations and bipolar depression.** Furthermore we observed, that **amplification of the 3' region of the *GSK3B* gene is combined with the deletion (single copy only) of the 5' end of the gene in most cases.**

Genotype analysis of miRSNPs in the SNAP-25 gene

In silico analysis of the *SNAP-25* gene revealed **two SNPs (rs3746544 and rs1051312), which might influence the binding of microRNAs. Efficient and reliable technique was elaborated for the genotype and haplotype analysis of these two polymorphic loci**. Haplotype of more than 1000 individuals was determined, and it was observed that the G–C haplotype cannot be found in the investigated Hungarian population.

Genotype analysis of miRSNs in the WFS1 gene

Two miRSNPs **(rs1046322 and rs9457)** were identified in the *WFS1* gene by *in silico* analysis. **Two independent methods were elaborated for the genotype analysis** of these two loci, which were based on the application of PCR–RFLP and real-time PCR, and genotype and allele-frequencies were determined in a health control population.

Investigation of WSF1 gene miRSNP variants and aggression

DNA samples of 801 individuals were analyzed, 17 SNPs, including the 2 miRSNPs of the *WFS1* gene were investigated. All genotype frequencies corresponded to the Hardy–Weinberg-equilibrium. 4 scales of 3 questionnaires (impulsivity, aggression,

anxiety and depression) were used. Association study of the investigated phenotypes and analyzed SNPs revealed statistically **significant (p = 0.0005) relationship between aggression measured by self reporting questionnaire and one of the miRSNPs (rs1046322)**. This results remained statistically significant even after using Bonferroni-correction (17 × 4 tests) for multiple testing.

Functional analysis of a miRSNP in the WFS1 gene

In silico data suggested that the rs1046322 polymorphism altered the binding site of miR-668: G allele resulted in a perfect binding site for the seed sequence of the miRNA, whereas A allele was expected to weaken binding efficiency. This assumption was planned to be confirmed by molecular methods applying luciferase reporter system. Therefore the whole 3' UTR sequence (or a control sequence for reference) of the WFS1 gene was cloned behind the luciferase gene, and site directed mutagensis was used to generate the DNA construct with the other allelic variant. DNA-constructs were transfected in HEK293T cells by co-transfecting miR-668. Our results suggested that **miR-668 can bind to the 3' UTR of the WFS1 gene and this interaction results in the inhibition of the translation of the reporter protein. Moreover we proved that the inhibitory effect of the A variant is lower than that of the G form.** Our data suggest the functional role of rs1046322 SNP in the regulation of the *WFS1* gene.

Conclusions

CNV polymorphisms and miRSNPs of several candidate genes were investigated in control samples and subjects suffering from uni- and bipolar depression. Recent studies showed that *GSK3B* is the candidate gene of uni- and bipolar depression, results about the CNV affecting the copy number of the gene are however controversial. Significant association was found between the CNV variants of *GSK3* bipolar depression in our study. Interestingly the deletion of the 5' region of the gene could be observed in combination with the amplification of the 3' part, which in some cases

reached extremely high copy number. This interesting pattern might be explained by the assumption, that the deletion of the 5' region results in lower gene expression, and it is compensated by the higher number of miRNA binding sites in the 3' region, increasing the binding site to miRNA ratio resulting in lower inhibiting effect if miRNAs.

MiRSNPs were investigated in the *SNAP-25* and *WFS1* genes. Two polymorphic loci were identified in both genes by *in silico* methods, respectively. The two SNPs (rs3746544 and rs1051312) in the *SNAP-25* gene were on close proximity of each other (there were only 3 basepairs between them), and two miRNAs were shown to be influenced by their genotypes. Therefore an efficient, real-time PCR-based method was elaborated for the direct haplotype determination of the two polymorphisms, which was applied to analyze more than 1000 persons. We observed that the G–C haplotype does not occur in our population.

Loss of function mutations of the *WFS1* gene are known to cause neural degeneration accompanied by psychiatric disorders, the syndrome is inherited in a monogenic manner. Moreover healthy persons carrying the mutations in heterozygote form suffer from psychiatric problems rather frequently. Consequently it was suggested that polymorphisms in the *WFS1* gene might be in the genetic background of psychological traits related to aggression and impulsivity. However, no unambiguous results have been published so far. Our results confirmed the connection between rs1046322 SNP of the *WFS1* gene and aggression, this results was statistically significant even after Bonferroni-correction for multiple testing. Moreover our molecular analyses demonstrated that this site influences the binding efficiency of miR-668 fulfilling the criteria of miR-SNP.

Functional analysis of polymorphisms associated with diseases is crucial regarding the identification of the patomechanism of the given illness. These data might contribute to the elaboration of efficient prevention and therapeutic protocols.

List of publications

Publications related to the theses

1. **Kovacs-Nagy R**, Elek Z, Szekely A, Nanasi T, Sasvari-Szekely M, Ronai Z. (2013) Association of aggression with a novel microRNA binding site polymorphism in the Wolframin gene. Am J Med Genet B Neuropsychiatr Genet, 162B, (4): 404-12 (IF: 3.705)
2. **Kovacs-Nagy R**, Hu J, Ronai Z, Sasvari-Szekely M. SNAP-25. (2009) A novel candidate gene in psychiatric genetics. Neuropsychopharmacol Hung, XI, (2): 89-94.
3. **Kovacs-Nagy R**, Sarkozy P, Hu J, Guttman A, Sasvari-Szekely M, Ronai Z. (2011) Haplotyping of putative microRNA binding sites in the SNAP-25 gene. Electrophoresis, 2011, 32, (15): 2013-20. (IF: 3.303)

Publications unrelated of the theses

1. Nagy G, **Nagy R**, Székely A, Sasvári-Székely M, Somogyi A. "Investigation of T1645C polymorphism of the KCNA3 gene in diabetes". *Article in Hungarian* Magyar Belorvosi Archivum, 2010, 63: 99-103.
2. Nagy G, **Kovacs-Nagy R**, Kereszturi E, Somogyi A, Szekely A, Nemeth N, Hosszufalusi N, Panczel P, Ronai Z, Sasvari-Szekely M. Association of hypoxia inducible factor-1 alpha gene polymorphism with both type 1 and type 2 diabetes in a Caucasian (Hungarian) sample. BMC Med Genet. 2009:19(10). 79. (IF: 2,84)
3. Kotyuk E, **Kovács-Nagy R**, Faludi G, Urbán R, Rónai Z, Sasvári-Székely M és Székely A. "Association of nicotine dependence and –521CT polymorphism of the dopamine D4 receptor in a patient group with major depression" *Article in Hungarian* Neuropsychopharmacol Hung. 2009, XI/2, 59-67.
4. Szekely A, **Kovacs-Nagy R**, Bányai ÉI, Gősi-Greguss AC, Varga K, Halmai Z, Ronai Z, Sasvari-Szekely M Association Between Hypnotizability and the Catechol-O-

Methyltransferase (COMT) Polymorphism. Int J Clin Exp Hypn. 2010, 58(3): 301-315. (IF: 2.068)
5. Elek Z, Brauswetter D, **Kovács-Nagy R**, Nagy G, Székely A, Somogyi A, Rónai Z, Sasvári-Székely M. "Genetic variability of micro-RNA binding sites and diabetes mellitus". *Article in Hungarian* Magyar Belorvosi Archivum, 2010:2, 99-103.
6. **Kovács-Nagy R**, Nagy G, Somogyi A, Székely A, Sasvári-Székely M, Rónai Z. "Putative novel genetic risk factor of type 2 diabetes mellitus" *Article in Hungarian* Magyar Belorvosi Archivum 2010:2, 91-98.
7. Tiszlavicz Z, Szabolcs A, Takács T, Farkas G, **Kovács-Nagy R**, Szántai E, Sasvári-Székely M, Mándi Y. Polymorphisms of beta defensins are associated with the risk of severe acute pancreatitis. Pancreatology. 2010:10(4), 483-90. (IF: 2.195)

Acknowledgement

I would like to thank my supervisor, Zsolt Rónai, the head of the Molecular Genetic Laboratory, Mária Sasvári and the head of the Pathobiochemsitry PhD School for supporting my research work. I thank all my co-workers, especially Eszter Szántai, Diána Brauswetter, Zsuzsanna Elek for their help and support.

I also thank Anna Székely for the collaboration and for help help in the statistical analyses.

I also thank for the work of the psychiatrists and psychologists who worked in these projects, and the participants for doing the tests and providing DNA sample.