Incomplete Coverage of Candidate Genes: A Poorly Considered Bias

Antonio Drago, Diana De Ronchi and Alessandro Serretti*

Institute of Psychiatry, University of Bologna, Italy

Abstract: Current genetic investigations are performed both on the basis of a rational and biologically based choice of candidate genes and through genome wide scans. Nonetheless, lack of replication is a common problem in psychiatric genetics as well as in other genetic fields. There are a number of reasons for this inconsistency, among them a well known but poorly considered issue is gene coverage. The aim of the present paper is to focus on this well known and defectively deemed bias, especially when a candidate gene approach is chosen. The rational and the technical feasibility of this proposal are discussed as well as a survey of current investigations. The known consistent methodology to fix this bias is also discussed.

INTRODUCTION

By identifying heritable risk factors, genetic association studies will reasonably help to clarify the biological basis of psychiatric disorders, to identify the most consistent prognostic factors, to simplify the therapeutic choices in every day clinic management. Anyway, despite a worldwide impressive scientific effort, there are not many conclusive results so far. This lack of consistency may be due to some methodological bias. There are roughly two main ways to investigate the associations between genetic variants and phenotypes or endophenotypes: genome wide analysis and candidate gene analysis. Going genome wide is probably the way of the future, provided that the haplotype strategy turns out to be really applicable and effective across the genome [1]: however, genome wide investigations focused on psychiatric disorders reported no definitive association results so far [2]. On the other hand, the candidate gene approach showed some good results and some genetic variations proved to be highly informative: for example, the insertion/deletion polymorphism in the promoter of the serotonin transporter has been independently demonstrated to be associated with some aspects of psychiatric disorders [3-5], and some clinical guidelines for the use of pharmacogenetic testing for CYP4502D6 and 2C19 mutations are available already [6]. The genome wide and candidate methods are not thought to be competitive: inductive and deductive information can be drawn with mutual advantage. Anyway, a large part of genetic research, both with genome wide and candidate gene methodologies, leaded to poorly replicated results. This might be due to a list of reasons, and some of them are methodological by nature: biases related to phenotype definitions (diagnosis, symptom clusters), sample size, selection bias, treatment lack of homogeneity, statistical biases and the little impact of a single gene over clinical phenotypes [7, 8]. Beyond this well documented complexity, one well known

but poorly considered bias is the incomplete coverage of genetic variations running within the gene which is under investigation. The aim of this paper is to consider if recent literature under estimated this last point. The reader will find a list of web resources helping to fix this bias at the end of the paper.

THE ROLE OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs)

Since the finding of the first physical map of the human genome in 1956, the suggestion of random DNA markers to build a sequence map in 1980 [9], the completion of the sequence of human genome in 2003 and the first study with a genome wide technique (more than 44 clinical trials published in the last two years), the research community has now the possibility to do both wide (entire genome) and hypothesis related (single nucleotide polymorphism) genetic association studies. Briefly, DNA variations can occur at different levels: duplications, insertions, deletions and transpositions. The most frequent changes involve single nucleotide substitutions, insertions and deletions. There are two basic classes of polymorphisms: SNP (single nucleotide polymorphism) and VNTR (variable number of tandem repeats). A list of SNPs is available from public databases (http://www.hapmap.org/downloads/index.html.en), and a part of these variations has a relevant prevalence within the general population: they are considered as "common" above the 5% frequency [10]. There is a variable number of SNPs for each gene: the longer the gene, the higher the number of running variations within its sequence. It is not easy to define the average number of one gene mutations, but it may be reasonably said that one single mutation is not representative of a gene's complete sequence. In fact, there might be other mutations within the same gene counteracting or enhancing the effect of the first one. We hypothesized that this simple, self – established statement is poorly considered by recent literature. To test this hypothesis, we surveyed the literature based on genetic association studies, inclusion criteria is listed below. A special attention might be paid to one of

*Address correspondence to this author at the Institute of Psychiatry, University of Bologna, Viale Carlo Pepoli 5, 40123 Bologna, Italy; Tel: +39 051 6584233; Fax: +39 051 521030; E-mail: alessandro.serretti@unibo.it

them: a sample size threshold (about one hundred patients) was chosen as inclusion criteria. This was meant to be closer to rational concerns about the studies' economic burden than to statistics: budget represents a relevant part in every study design, and a cut off of one hundred patients may identify studies which could afford a complete genetic analysis. Some words should be spent about this controversial topic. The definition of a statistically correct sample size for an association study appears to be quite a complex task: small sample sizes are not representative, while bigger sample sizes are associated with the risk of higher false positive rates, as recently demonstrated by Sullivan [11]: his *in silico* investigation showed dramatic rates of false positive results with a sample of 500 cases and 500 controls, examined for a set of 10 SNPs. Conflicting with this, genome wide analysis performed on thousands of patients only reported mild significant results: the recent genome wide investigation by Fanous and colleagues [12] (n=1383), for example, shed some light on the boundaries between schizophrenia and schizotypy, but the level of significance (p=0.04 and p=0.02) does not allow to discard the possibility of a first type error occurrence. Similar considerations might then be done for other interesting recent findings: in an relevant paper by Bulayeva and colleagues, the genome wide investigation of a set of genetic isolated schizophrenic pedigrees revealed different pattern of association with psychiatric phenotypes between younger and older pedigrees [13]. The study of genetic isolated pedigrees lowered the occurrence of population stratification factors, and the wide investigated sample sizes (hundreds) were a relevant point of the study, but the level of significance that was reported (p < 0.05) may not be considered sufficiently protective. Furthermore, it must be noticed that, even with a higher level of significance (p<0.002) at the genome-wide analysis, the subsequent SNP investigation does not necessarily confirm the results [14]. Many factors probably influenced the statistical power of these studies and therefore, they all should be considered for statistically determined correct sample size: the haplotype Linkage Disequilibrium (LD) which varies along chromosomes and within genes, the assumed genetic path associated with disorders (additive or multiplicative), the prevalence of the investigated disorder in the general population, the prevalence of the alleles associated with the disorder, and the definition of the relative risk for the disorder may represent some examples of these interactions as far as they are all influent toward the definition of a sample size tailored to a fixed statistical detection power [15].

So, even in the simplest case, when a candidate gene approach is chosen, the index sample size will vary according to the disorder, to the haplotype Linkage Disequilibrium average value in the candidate gene, to the associated expected Odds Ratio and so on. Thus, the correct identification of the exact average sample size for a rational analysis appears to be quite a complex task, and it is anyway beyond the aim of this paper. If the reader is interested in this topic, De La Vega recently proposed an online free software meant to deal with these issues [15], and the recently published HAPMAP project phase II focuses on these topics in deep too [10], reader's attention is then readdressed to these relevant publications. More simply, we choose to use the sample size to select appropriate studies for our survey on the basis

that, given the economical effort leading to the selection of about two hundred persons (cases and controls), it would be worth performing a completer SNP selection, avoiding the simple bias of an incomplete investigation, which otherwise would detract from the scientific and economical effort of the study. According to this advice, we listed some recent association studies published in the last three years (July 2004 – July 2007) in the field of psychiatric genetic research, and pointed out the investigated/known genetic variations ratio for each gene in the studies. Pubmed database was used with the following criteria:

1. Key words: gene, polymorphism, schizophrenic, bipolar, depressive, anxiety, obsessive, panic, PTSD, phobic.

2. Sample size = 90-100

3. Limits: Humans, Clinical Trial, Meta-Analysis, English.

Results are presented in Table **1**, SNP data are collected from NCBI database (http://www.ncbi.nlm.nih.gov/).

Table **1** reports the low investigated/known genetic variation ratio in almost all considered studies: the complete SNP coverage of investigated gene then represents quite a new topic in nowadays literature [16]. There are good and well known reasons to revalue this point: first of all, it is rational to assume that every SNP occurring in the genetic coding sequence is crucial as far as it can be associated with a change in the secondary mRNA structure: that is why silent substitutions cannot be considered *a priori* as devoid of interest if they run in exons. Moreover, SNPs within the coding sequence can be associated with different aminoacid sequence, leading to secondary, tertiary or quaternary possibly different protein structures, and, as a consequence or independently from it, altered function. Not all the DNA sequence is expressed in every cell [17], but a part of it being silent or playing a regulatory role, probably influencing the differentiation and specialisation processes [18-20]: it might be expected then that a number of SNPs will occur in non coding DNA, and may be in - or close to - a promoter, enhancer, silencer or other regulating sequence. Promoters are usually close to exons sequence (TATA box usually locate at -25bp; CAAT box at -80bp; CG box up to -950bp), but enhancers and silencers can be located at a considerable distance from the coding stream or inside introns. Other reasons to consider intronic polymorphisms relevant to genetic investigations, arise from the DNA accessibility to RNA polymerases: it is commonly accepted that this mechanism is at least partially limited by methylation and acetylation [21-26], and since cytosine has been proposed to be a site of methylation, its variation due to a change in the genetic sequence could have an influence on the gene expression. Consistently, there is evidence that the T102C silent variation influence HTR2A gene expression [27] probably through this mechanism. Finally, an intron variation could be completely silent but in strong Linkage Disequilibrium with a not yet known variation, possibly involving genes or sequences still far to be hypothesized as relevant to the topic of the study.

As Table **2** shows, nowadays investigations of non exon variations are limited to sequences located 1 kbp around exons. For the above mentioned considerations, scans should be wider. To complete Table **2** we used temporal criteria (the

Table 1.   Examples of Investigated/Known Variations Ratio

Author	Sample	Gene (s)	Variation (Investigated / Known)	Number of Tag SNPs	Number of LD Blocks	Main Association Result
[32]	1447	CREB 1	7 / 173	18	4	Positive association in men (suicide)
[33]	854	HTR2C; HTR1A	6 / 1100 (HTR2C); 3 / 16 (HTR1A)	18 (HTR2C) No haplotypes (HTR1A)	4 (HTR2C) No haplotypes (HTR1A)	No correlation
[34]	1913	SLC6A4	1 / 171	14	1	No correlation (MDD, positive correlation if stress associated; alcohol dependence)
[35]	1435 (meta analysis)	SLC6A4	1 / 171	14	1	Positive correlation (MDD, AD treatment)
[36]	1914	SLC6A4	1 / 171	14	1	No correlation (AD treatment)
[37]	1648 (meta analysis)	MTHFR	2 / 147	9	1	Positive correlation (depression, anxiety, psychosis)
[38]	9032 (meta analysis)	SLC6A4	1 / 171	14	1	Positive correlation (sucide; p = 0.0068)
[39]	3000 (meta analysis)	SLC6A4	1 / 171	14	1	Positive correlation (OCD)
[40]	195	GNbeta3	1 / 39	3	No haplotypes	Positive correlation (self mutilation)
[41]	222	HTR1A	2 / 16	No haplotypes	No haplotypes	Positive correlation (AD treatment, females)
[42]	4175 (298 MDD)	SLC6A4	1 / 171	14	1	No correlation
[43]	258	SLC6A4	1 / 171	14	1	Positive correlation (suicide)
[44]	196	DRD4	1 / 106	No haplotypes	No haplotypes	Positive correlation (neuroticism)
[45]	450	GLO1	1 / 152	11	2	Positive correlation (panic without agoraphobia)
[46]	937	GAL	4 / 30	3	1	Positive correlation (alcoholism)
[47]	755	DTNBP1	2 / 448	26	2	Positive correlation (negative symptoms of schizophrenia)
[48]	2376	BDNF	3 / 214	5	1	Positive correlation (MDD)
[49]	178	CLOCK	1 / 523	24	5	Positive correlation (AD treatment → insomnia)
[50]	273	ACE; ATR1	(1 + 1) / (259 + 422)	11 (ACE) 18 (ATR1)	4 (ACE) 2 (ATR1)	Positive correlation (AD treatment)
[51]	1512	COMT	1 / 293	14	3	Positive correlation (MDD)
[52]	753	GPR50	3 / 29	No haplotypes	No haplotypes	Positive correlation (DB)
[53]	1005	SLC6A4	1 / 171	14	1	Positive correlation (gene environment influence)
[54]	295	SLC6A4	1 / 171	14	1	No correlation (OCD)
[55]	159	GABBR1	5 / 204	21	6	Positive correlation (OCD)
[56]	287	BDNF	2 / 214	5	1	No correlation
[57]	273	21 candidate genes	90 polymorphisms (average = 3.5)	Different genes	Different genes	Positive association (PD)
[58]	230	NET	3 / 263	25	4	Positive correlation (PD without agoraphobia)

(Table 1) contd….

Author	Sample	Gene (s)	Variation (Investigated / Known)	Number of Tag SNPs	Number of LD Blocks	Main Association Result
[59]	373	SCL6A4; MAOA; TPH1; HTR1B	2 / 171 (SLC6A4); 1 / 165 (MAOA); 1 / 81 (TPH1); 1 / 25 (HTR1B)	14 (SLC6A4) 8 (MAOA) 2 (TPH1) 4 (HTR1B)	1 (SLC6A4) 1 (MAOA) None (TPH1) 1 (HTR1B)	No correlation
[60]	65 families	PIP5K2A	15 / 742	53	12	Positive association
[61]	433	FZD3	2 / 280	4	1	No correlation
[62]	944	BDNF	2 / 214	5	1	Positive association
[63]	896	SYN3	1 / 2742	257	45	No correlation
[64]	1153	RGS4	4 / 36	6	1	No correlation

Gene names are the officials ones according to NCBI database. To identify the Tag SNPs Haploviewer tagger program was used with default settings, CEU population was selected.

Table 2.   Papers Published in 2006–7 that Reported Positive Association Results for Intronic SNPs

Author	Disorder	Chromosome	Gene	Variation (Intron)	Distance from the Nearest Exon
[65]	Psychosis	22q13.33	MLC1	rs2235349; rs2076137	~ 100 bp and 65 bp
[66]	Psychosis	18p11.3	TGIF	D18S63	~ 9.2 kbp
[67]	PTSD	Xp11.23	MAO-B	rs1799836	~ 30 bp
[68]	BD	14q22.3	OTX2	rs28757218	~ 20 bp
[69]	Personality	6p12.3	TFAP2B	VNTR (intron 2)	~ 100 bp
[70]	Antipsychotic treatment	7q36.1	NOS3	VNTR (intron 4)	~ 200 bp
[71]	Lithium treatment response	17q11.2	SLC6A4	VNTR (intron2)	~ 100 bp
[72]	Antidepressant response	5p15.33	DAT1	VNTR (intron 8)	~ 200 bp
[73]	Antidepressant response	Xp11.23	MAO-B	rs1799836	~ 30 bp
[74]	Eating disorder spectrum	3p25.3	GHRL	rs35680	~ 800 bp
[75]	Personality	19q13.42	PRKCG	rs402691	1.2 kbp
[76]	Psychosis	22q11.21	COMT	rs737865; rs737864	~ 700 kb
[77]	Depressive disorder	17q11.2	SLC6A4	rs25531	~ 1.5 kbp
[78]	Psychosis	6p21.31	FKBP5	17081296	~ 2.5 kbp
[79]	Antipsychotic response	15q24.1	CYP1A2	rs2472304	~ 40 bp
[80]	BD	8p21	VMAT1	rs2279709	~ 220 bp
[81]	Psychosis	8p12	NRG1	rs6150532	~ 700 bp
[82]	BD	21q22.3	TRPM2	rs1618355	~ 15 bp
[83]	Suicide	11p15.1	TPH-1	rs684302; rs211105; rs1800532; rs7933505	~ 250 to 2000 bp
[78]	Psychosis	6p21.31	FKBP5	rs1360780	~ 1.5 kbp
[44]	Personality	7p15.1	CRHR2	rs2267717	~ 4.5 kbp
[84]	Suicide	11p15.1	TPH	Intron 7 SNPs	< 750 bp
[85]	Depressive disorder	11q23.2	HTR3A and HTR3B	rs2276307; rs2276308; rs3782025; rs2276302	From 60 to 150 bp

Gene names are the officials ones according to NCBI database.

last two years), and the following key words: intron, intronic, depressive, psychotic, schizophrenic, bipolar, personality, suicide, temperament, eating, anxiety, panic, obsessive, polymorphism, SNP. Pubmed served as database. As regard to SNPs, only studies with possible identification of genetic location are included. One study for each investigated variation is included.

Finally, 3' and 5' endings are also important regulatory regions, and it is reasonable to assume that modifications running in these genetic sequences, or sufficiently close to, might play a relevant role.

Some lines of evidence encourage a complete analysis of single genes' variations: Myer and colleagues recently found no significant difference in the promoter activity of HTR2A gene between the A- and G- allele of the -1438 locus when expressed with the major alleles at –1420 C/T and –783 A/G loci [28]. This was not consistent with some previous literature findings [29]; but when the minor allele G at -783 was found to be expressed with G-allele at -1438, the promoter activity was found to be significantly decreased. Consistently, a triallelic variant of the well known serotonin promoter polymorphism has been recently reported [30], and only the A allele carriers at the A/G SNP within 5-HTTLPR insertion polymorphism yield high mRNA levels, and the L(G) carriers actually behave like the low expressing short allele. This finding can explain some of the not replicated findings in literature. Moreover, previous studies which investigated only the long/short polymorphism should be reconsidered.

SNPs occur, on average, about less than 1.000 bases. Since there are about 3 billion chemical base pairs that make up human DNA and its 20.000 – 25.000 genes, about 3 - 5 million of SNPs might be expected. This is consistent with the recent phase II Hap Map project findings [10]. The analysis of such a number of variations is feasible from a long time: actual techniques (Illumina and Affymetrix) permit the analysis of 500,000 or more SNPs in a single test with accessible costs, and lists of TagSNPs which can cover the complete gene sequence are easily retrievable from public databases (http://www.hapmap.org/downloads/index. html.en) (http://www.ncbi.nlm.nih.gov/projects/mapview/ map_search.cgi?taxid=9606).

Moreover, online free software is available which identify is a list of significant SNPs to cover a gene's common variations (http://www.broad.mit.edu/mpg/haploview/download.php) (http://marketing.appliedbiosystems.com/mk/get/ SNPB_LANDING?_A=22924&_D=18392&_V=0).

As a conclusion, a more complete analysis of genes' variations does not represent a novelty in actual knowledge, and its rationality as a methodological strategy is expected to be self – established. Nonetheless, we report here that it is a poorly considered methodological point which could be easily fixed using free internet resources and laboratory extra work, likely affordable in studies able to perform genetic and clinic assessments of hundreds of patients. There is some evidence that every day genetic clinic use will be a cost effective or even cost-saving approach in general clinical practice [31]: in order to hasten this process, a complete coverage of single gene polymorphisms is probably needed.

REFERENCES

[1]   Licinio, J., Wong, M.L. Pharmacogenomics in psychiatry: genomic considerations. *Mol. Psychiatry* **2005**, *10*: 713.

[2]   Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **2007**, *447*: 661-78.

[3]   Cho, H.J., Meira-Lima, I., Cordeiro, Q., Michelon, L., Sham, P., Vallada, H., Collier, D.A. Population-based and family-based studies on the serotonin transporter gene polymorphisms and bipolar disorder: a systematic review and meta-analysis. *Mol. Psychiatry* **2005**, *10*: 771-81.

[4]   Sen, S., Burmeister, M., Ghosh, D. Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. *Am. J. Med. Genet.* **2004**, *127B*: 90-3.

[5]   Serretti, A., Mandelli, L., Lorenzi, C., Pirovano, A., Olgiati, P., Colombo, C., Smeraldi, E. Serotonin transporter gene influences the time course of improvement of "core" depressive and somatic anxiety symptoms during treatment with SSRIs for recurrent mood disorders. *Psychiatry Res.* **2007**, *149*: 185-93.

[6]   de Leon, J., Armstrong, S.C., Cozza, K.L. Clinical guidelines for psychiatrists for the use of pharmacogenetic testing for CYP450 2D6 and CYP450 2C19. *Psychosomatics* **2006**, *47*: 75-85.

[7]   Haslam, N. Bias in psychopathology research. *Curr. Opin. Psychiatry* **2006**, *19*: 625-30.

[8]   Serretti, A., Kato, M., Kennedy, J.L. Pharmacogenetic studies in depression: a proposal for methodologic guidelines. *Pharmacogenomics J.* **2007**.

[9]   Shortle, D., Koshland, D., Weinstock, G.M., Botstein, D. Segment-directed mutagenesis: construction *in vitro* of point mutations limited to a small predetermined region of a circular DNA molecule. *Proc. Natl. Acad. Sci. USA* **1980**, *77*: 5375-9.

[10] Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., Pasternak, S., Wheeler, D.A., Willis, T.D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S.B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R.C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M.M., Tsui, S.K., Xue, H., Wong, J.T., Galver, L.M., Fan, J.B., Gunderson, K., Murray, S.S., Oliphant, A.R., Chee, M.S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J.F., Phillips, M.S., Roumy, S., Sallée, C., Verner, A., Hudson, T.J., Kwok, P.Y., Cai, D., Koboldt, D.C., Miller, R.D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L.C., Mak, W., Song, Y.Q., Tam, P.K., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C.P., Delgado, M., Dermitzakis, E.T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B.E., Whittaker, P., Bentley, D.R., Daly, M.J., de, Bakker, P.I., Barrett, J., Chretien, Y.R., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Price, A., Purcell, S., Richter, D.J., Sabeti, P., Saxena, R., Schaffner, S.F., Sham, P.C., Varilly, P., Altshuler, D., Stein, L.D., Krishnan, L., Smith, A.V., Tello-Ruiz, M.K., Thorisson, G.A., Chakravarti, A., Chen, P.E., Cutler, D.J., Kashuk, C.S., Lin, S., Abecasis, G.R., Guan, W., Li, Y., Munro, H.M., Qin, Z.S., Thomas, D.J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L.R., Clarke, G., Evans, D.M., Morris, A.P., Weir, B.S., Tsunoda, T., Mullikin, J.C., Sherry, S.T., Feolo, M., Skol, A., Zhang, H., Zeng, C., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D.R., Suda, E., Rotimi, C.N., Adebamowo, C.A., Ajayi, I., Aniagwu, T., Marshall, P.A., Nkwodimmah, C., Royal, C.D., Leppert, M.F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I.F., Knoppers, B.M., Foster, M.W., Clayton, E.W., Watkin, J., Gibbs, R.A., Belmont, J.W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G.M., Wheeler, D.A., Yakub, I., Gabriel, S.B., Onofrio, R.C., Richter, D.J., Ziaugra, L., Birren, B.W., Daly, M.J., Altshuler, D., Wilson, R.K., Fulton, L.L., Rogers, J., Burton, J., Carter, N.P., Clee, C.M., Griffiths, M., Jones, M.C., McLay, K., Plumb, R.W., Ross, M.T., Sims, S.K., Willey,

D.L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J.C., L'Archevêque, P., Bellemare, G., Saeki, K., Wang, H., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A.L., Brooks, L.D., McEwen, J.E., Guyer, M.S., Wang, V.O., Peterson, J.L., Shi, M., Spiegel, J., Sung, L.M., Zacharia, L.F., Collins, F.S., Kennedy, K., Jamieson, R., Stewart, J.A second generation human haplotype map of over 3.1 million SNPs. *Nature* **2007**, *449*: 851-61.

[11] Sullivan, P.F. Spurious genetic associations. *Biol. Psychiatry* **2007**, *61*: 1121-6.

[12] Fanous, A.H., Neale, M.C., Gardner, C.O., Webb, B.T., Straub, R.E., O'Neill, F.A., Walsh, D., Riley, B.P., Kendler, K.S. Significant correlation in linkage signals from genome-wide scans of schizophrenia and schizotypy. *Mol. Psychiatry* **2007**, *12*: 958-65.

[13] Bulayeva, K.B., Glatt, S.J., Bulayev, O.A., Pavlova, T.A., Tsuang, M.T. Genome-wide linkage scan of schizophrenia: a cross-isolate study. *Genomics* **2007**, *89*: 167-77.

[14] Vazza, G., Bertolin, C., Scudellaro, E., Vettori, A., Boaretto, F., Rampinelli, S., De Sanctis, G., Perini, G., Peruzzi, P., Mostacciuolo, M.L. Genome-wide scan supports the existence of a susceptibility locus for schizophrenia and bipolar disorder on chromosome 15q26. *Mol. Psychiatry* **2007**, *12*: 87-93.

[15] De La Vega, F.M. Selecting single-nucleotide polymorphisms for association studies with SNPbrowser software. *Methods Mol. Biol.* **2007**, *376*: 177-93.

[16] Horne, B.D., Camp, N.J. Principal component analysis for selection of optimal SNP-sets that capture intragenic genetic variation. *Genet. Epidemiol.* **2004**, *26*: 11-21.

[17] Schena, M., Heller, R.A., Theriault, T.P., Konrad, K., Lachenmeier, E., Davis, R.W. Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol.* **1998**, *16*: 301-6.

[18] Halbleib, J.M., Saaf, A.M., Brown, P.O., Nelson, W.J. Transcriptional modulation of genes encoding structural characteristics of differentiating enterocytes during development of a polarized epithelium *in vitro*. *Mol. Biol. Cell* **2007**, *18*: 4261-78.

[19] Milanesi, L., Petrillo, M., Sepe, L., Boccia, A., D'Agostino, N., Passamano, M., Di Nardo, S., Tasco, G., Casadio, R., Paolella, G. Systematic analysis of human kinase genes: a large number of genes and alternative splicing events result in functional and structural diversity. *BMC Bioinformatics* **2005**, *6* Suppl 4: S20.

[20] Vinogradov, A.E., Anatskaya, O.V. Organismal complexity, cell differentiation and gene expression: human over mouse. *Nucleic Acids Res.* **2007**, *35*: 6350-6.

[21] Bannister, A.J., Miska, E.A. Regulation of gene expression by transcription factor acetylation. *Cell Mol. Life Sci.* **2000**, *57*: 1184-92.

[22] Fukuda, Y., Koga, M., Arai, M., Noguchi, E., Ohtsuki, T., Horiuchi, Y., Ishiguro, H., Niizato, K., Iritani, S., Itokawa, M., Arinami, T. Monoallelic and unequal allelic expression of the HTR2A gene in human brain and peripheral lymphocytes. *Biol. Psychiatry* **2006**, *60*: 1331-5.

[23] Hoffman, A.R., Hu, J.F. Directing DNA methylation to inhibit gene expression. *Cell Mol. Neurobiol.* **2006**, *26*: 425-38.

[24] Serman, A., Vlahovic, M., Serman, L., Bulic-Jakus, F. DNA methylation as a regulatory mechanism for gene expression in mammals. *Coll. Antropol.* **2006**, *30*: 665-71.

[25] Shilatifard, A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. *Annu. Rev. Biochem.* **2006**, *75*: 243-69.

[26] Verdone, L., Caserta, M., Di Mauro, E. Role of histone acetylation in the control of gene expression. *Biochem. Cell Biol.* **2005**, *83*: 344-53.

[27] Clifford, C.P., Nunez, D.J. 5HT 2a receptor T102C polymorphism and schizophrenia. *Lancet* **1996**, *347*: 1830.

[28] Myers, R.L., Airey, D.C., Manier, D.H., Shelton, R.C., Sanders-Bush, E. Polymorphisms in the Regulatory Region of the Human Serotonin 5-HT(2A) Receptor Gene (HTR2A) Influence Gene Expression. *Biol. Psychiatry* **2006**, *11*.

[29] Serretti, A., Drago, A., De Ronchi, D. HTR2A gene variants and psychiatric disorders: a review of current literature and selection of SNPs for future studies. *Curr. Med. Chem.* **2007**, *14*: 2053-69.

[30] Hu, X.Z., Rush, A.J., Charney, D., Wilson, A.F., Sorant, A.J., Papanicolaou, G.J., Fava, M., Trivedi, M.H., Wisniewski, S.R., Laje, G., Paddock, S., McMahon, F.J., Manji, H., Lipsky, R.H. Association between a functional serotonin transporter promoter polymorphism and citalopram treatment in adult outpatients with major depression. *Arch. Gen. Psychiatry* **2007**, *64*: 783-92.

[31] Rogowski, W. Genetic screening by DNA technology: a systematic review of health economic evidence. *Int. J. Technol. Assess Health Care* **2006**, *22*: 327-37.

[32] Perlis, R.H., Purcell, S., Fava, M., Fagerness, J., Rush, A.J., Trivedi, M.H., Smoller, J.W. Association between treatment-emergent suicidal ideation with citalopram and polymorphisms near cyclic adenosine monophosphate response element binding protein in the STAR*D study. *Arch. Gen. Psychiatry* **2007**, *64*: 689-97.

[33] Serretti, A., Mandelli, L., Giegling, I., Schneider, B., Hartmann, A.M., Schnabel, A., Maurer, K., Moller, H.J., Rujescu, D. HTR2C and HTR1A gene variants in German and Italian suicide attempters and completers. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2007**, *144*: 291-9.

[34] Dick, D.M., Plunkett, J., Hamlin, D., Nurnberger, J. Jr., Kuperman, S., Schuckit, M., Hesselbrock, V., Edenberg, H., Bierut, L. Association analyses of the serotonin transporter gene with lifetime depression and alcohol dependence in the Collaborative Study on the Genetics of Alcoholism (COGA) sample. *Psychiatr. Genet.* **2007**, *17*: 35-8.

[35] Serretti, A., Kato, M., De Ronchi, D., Kinoshita, T. Meta-analysis of serotonin transporter gene promoter polymorphism (5-HTTLPR) association with selective serotonin reuptake inhibitor efficacy in depressed patients. *Mol. Psychiatry* **2007**, *12*: 247-57.

[36] Kraft, J.B., Peters, E.J., Slager, S.L., Jenkins, G.D., Reinalda, M.S., McGrath, P.J., Hamilton, S.P. Analysis of association between the serotonin transporter and antidepressant response in a large clinical sample. *Biol. Psychiatry* **2007**, *61*: 734-42.

[37] Gilbody, S., Lewis, S., Lightfoot, T. Methylenetetrahydrofolate reductase (MTHFR) genetic polymorphisms and psychiatric disorders: a HuGE review. *Am. J. Epidemiol.* **2007**, *165*: 1-13.

[38] Li, D., He, L. Meta-analysis supports association between serotonin transporter (5-HTT) and suicidal behavior. *Mol. Psychiatry* **2007**, *12*: 47-54.

[39] Lin, P.Y. Meta-analysis of the association of serotonin transporter gene polymorphism with obsessive-compulsive disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2007**, *31*: 683-9.

[40] Joyce, P.R., McKenzie, J.M., Mulder, R.T., Luty, S.E., Sullivan, P.F., Miller, A.L., Kennedy, M.A. Genetic, developmental and personality correlates of self-mutilation in depressed patients. *Aust. N. Z. J. Psychiatry* **2006**, *40*: 225-9.

[41] Yu, Y.W., Tsai, S.J., Liou, Y.J., Hong, C.J., Chen, T.J. Association study of two serotonin 1A receptor gene polymorphisms and fluoxetine treatment response in Chinese major depressive disorders. *Eur. Neuropsychopharmacol.* **2006**, *16*: 498-503.

[42] Surtees, P.G., Wainwright, N.W., Willis-Owen, S.A., Luben, R., Day, N.E., Flint, J. Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. *Biol. Psychiatry* **2006**, *59*: 224-9.

[43] Lopez de Lara, C., Dumais, A., Rouleau, G., Lesage, A., Dumont, M., Chawky, N., Alda, M., Benkelfat, C., Turecki, G. STin2 variant and family history of suicide as significant predictors of suicide completion in major depression. *Biol. Psychiatry* **2006**, *59*: 114-20.

[44] Tochigi, M., Hibino, H., Otowa, T., Kato, C., Marui, T., Ohtani, T., Umekage, T., Kato, N., Sasaki, T. Association between dopamine D4 receptor (DRD4) exon III polymorphism and Neuroticism in the Japanese population. *Neurosci. Lett.* **2006**, *398*: 333-6.

[45] Politi, P., Minoretti, P., Falcone, C., Martinelli, V., Emanuele, E. Association analysis of the functional Ala111Glu polymorphism of the glyoxalase I gene in panic disorder. *Neurosci. Lett.* **2006**, *396*: 163-6.

[46] Belfer, I., Hipp, H., McKnight, C., Evans, C., Buzas, B., Bollettino, A., Albaugh, B., Virkkunen, M., Yuan, Q., Max, M.B., Goldman, D., Enoch, M.A. Association of galanin haplotypes with alcoholism and anxiety in two ethnically distinct populations. *Mol. Psychiatry* **2006**, *11*: 301-11.

[47] Fanous, A.H., van den Oord, E.J., Riley, B.P., Aggen, S.H., Neale, M.C., O'Neill, F.A., Walsh, D., Kendler, K.S. Relationship between a high-risk haplotype in the DTNBP1 (dysbindin) gene and clinical features of schizophrenia. *Am. J. Psychiatry* **2005**, *162*: 1824-32.

[48] Schumacher, J., Jamra, R.A., Becker, T., Ohlraun, S., Klopp, N., Binder, E.B., Schulze, T.G., Deschner, M., Schmal, C., Hofels, S., Zobel, A., Illig, T., Propping, P., Holsboer, F., Rietschel, M., Nothen, M.M., Cichon, S. Evidence for a relationship between genetic

variants at the brain-derived neurotrophic factor (BDNF) locus and major depression. *Biol. Psychiatry* **2005**, *58*: 307-14.

[49] Serretti, A., Cusin, C., Benedetti, F., Mandelli, L., Pirovano, A., Zanardi, R., Colombo, C., Smeraldi, E. Insomnia improvement during antidepressant treatment and CLOCK gene polymorphism. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2005**, *137*: 36-9.

[50] Bondy, B., Baghai, T.C., Zill, P., Schule, C., Eser, D., Deiml, T., Zwanzger, P., Ella, R., Rupprecht, R. Genetic variants in the angiotensin I-converting-enzyme (ACE) and angiotensin II receptor (AT1) gene and clinical outcome in depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2005**, *29*: 1094-9.

[51] Massat, I., Souery, D., Del-Favero, J., Nothen, M., Blackwood, D., Muir, W., Kaneva, R., Serretti, A., Lorenzi, C., Rietschel, M., Milanova, V., Papadimitriou, G.N., Dikeos, D., Van Broekhoven, C., Mendlewicz, J. Association between COMT (Val(158)Met) functional polymorphism and early onset in patients with major depressive disorder in a European multicenter genetic association study. *Mol. Psychiatry* **2005**, *10*: 598-605.

[52] Thomson, P.A., Wray, N.R., Thomson, A.M., Dunbar, D.R., Grassie, M.A., Condie, A., Walker, M.T., Smith, D.J., Pulford, D.J., Muir, W., Blackwood, D.H., Porteous, D.J. Sex-specific association between bipolar affective disorder in women and GPR50, an X-linked orphan G protein-coupled receptor. *Mol. Psychiatry* **2005**, *10*: 470-8.

[53] Grabe, H.J., Lange, M., Wolff, B., Volzke, H., Lucht, M., Freyberger, H.J., John, U., Cascorbi, I. Mental and physical distress is modulated by a polymorphism in the 5-HT transporter gene interacting with social stressors and chronic disease burden. *Mol. Psychiatry* **2005**, *10*: 220-4.

[54] Kim, S.J., Lee, H.S., Kim, C.H. Obsessive-compulsive disorder, factor-analyzed symptom dimensions and serotonin transporter polymorphism. *Neuropsychobiology* **2005**, *52*: 176-82.

[55] Zai, G., Arnold, P., Burroughs, E., Barr, C.L., Richter, M.A., Kennedy, J.L. Evidence for the gamma-amino-butyric acid type B receptor 1 (GABBR1) gene as a susceptibility factor in obsessive-compulsive disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2005**, *134*: 25-9.

[56] Shimizu, E., Hashimoto, K., Koizumi, H., Kobayashi, K., Itoh, K., Mitsumori, M., Ohgake, S., Okamura, N., Koike, K., Matsuzawa, D., Zhang, L., Kumakiri, C., Nakazato, M., Komatsu, N., Iyo, M. No association of the brain-derived neurotrophic factor (BDNF) gene polymorphisms with panic disorder. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2005**, *29*: 708-12.

[57] Maron, E., Nikopensius, T., Koks, S., Altmae, S., Heinaste, E., Vabrit, K., Tammekivi, V., Hallast, P., Koido, K., Kurg, A., Metspalu, A., Vasar, E., Vasar, V., Shlik, J. Association study of 90 candidate gene polymorphisms in panic disorder. *Psychiatr. Genet.* **2005**, *15*: 17-24.

[58] Lee, Y.J., Hohoff, C., Domschke, K., Sand, P., Kuhlenbaumer, G., Schirmacher, A., Freitag, C.M., Meyer, J., Stober, G., Franke, P., Nothen, M.M., Fritze, J., Fimmers, R., Garritsen, H.S., Stogbauer, F., Deckert, J. Norepinephrine transporter (NET) promoter and 5'-UTR polymorphisms: association analysis in panic disorder. *Neurosci. Lett.* **2005**, *377*: 40-3.

[59] Maron, E., Lang, A., Tasa, G., Liivlaid, L., Toru, I., Must, A., Vasar, V., Shlik, J. Associations between serotonin-related gene polymorphisms and panic disorder. *Int. J. Neuropsychopharmacol.* **2005**, *8*: 261-6.

[60] Schwab, S.G., Knapp, M., Sklar, P., Eckstein, G.N., Sewekow, C., Borrmann-Hassenbach, M., Albus, M., Becker, T., Hallmayer, J.F., Lerer, B., Maier, W., Wildenauer, D.B. Evidence for association of DNA sequence variants in the phosphatidylinositol-4-phosphate 5-kinase IIalpha gene (PIP5K2A) with schizophrenia. *Mol. Psychiatry* **2006**, *11*: 837-46.

[61] Jeong, S.H., Joo, E.J., Ahn, Y.M., Lee, K.Y., Kim, Y.S. Investigation of genetic association between human Frizzled homolog 3 gene (FZD3) and schizophrenia: results in a Korean population and evidence from meta-analysis. *Psychiatry Res.* **2006**, *143*: 1-11.

[62] Neves-Pereira, M., Cheung, J.K., Pasdar, A., Zhang, F., Breen, G., Yates, P., Sinclair, M., Crombie, C., Walker, N., St Clair, D.M. BDNF gene is a risk factor for schizophrenia in a Scottish population. *Mol. Psychiatry* **2005**, *10*: 208-12.

[63] Lachman, H.M., Stopkova, P., Papolos, D.F., Pedrosa, E., Margolis, B., Aghalar, M.R., Saito, T. Analysis of synapsin III-196 promoter mutation in schizophrenia and bipolar disorder. *Neuropsychobiology* **2006**, *53*: 57-62.

[64] Cordeiro, Q., Talkowski, M.E., Chowdari, K.V., Wood, J., Nimgaonkar, V., Vallada, H. Association and linkage analysis of RGS4 polymorphisms with schizophrenia and bipolar disorder in Brazil. *Genes Brain Behav.* **2005**, *4*: 45-50.

[65] Selch, S., Strobel, A., Haderlein, J., Meyer, J., Jacob, C.P., Schmitt, A., Lesch, K.P., Reif, A. MLC1 polymorphisms are specifically associated with periodic catatonia, a subgroup of chronic schizophrenia. *Biol. Psychiatry* **2007**, *61*: 1211-4.

[66] Chavarria-Siles, I., Walss-Bass, C., Quezada, P., Dassori, A., Contreras, S., Medina, R., Ramirez, M., Armas, R., Salazar, R., Leach, R.J., Raventos, H., Escamilla, M.A. TGFB-induced factor (TGIF): a candidate gene for psychosis on chromosome 18p. *Mol. Psychiatry* **2007**, *12*: 1033-41.

[67] Pivac, N., Knezevic, J., Kozaric-Kovacic, D., Dezeljin, M., Mustapic, M., Rak, D., Matijevic, T., Pavelic, J., Muck-Seler, D. Monoamine oxidase (MAO) intron 13 polymorphism and platelet MAO-B activity in combat-related posttraumatic stress disorder. *J. Affect. Disord.* **2007**, *103*: 131-8.

[68] Sabunciyan, S., Yolken, R., Ragan, C.M., Potash, J.B., Nimgaonkar, V.L., Dickerson, F., Llenos, I.C., Weis, S. Polymorphisms in the homeobox gene OTX2 may be a risk factor for bipolar disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2007**, *144*: 1083-6.

[69] Nilsson, K.W., Damberg, M., Ohrvik, J., Leppert, J., Lindstrom, L., Anckarsater, H., Oreland, L. Genes encoding for AP-2beta and the Serotonin Transporter are associated with the Personality Character Spiritual Acceptance. *Neurosci. Lett.* **2007**, *411*: 233-7.

[70] Liou, Y.J., Lai, I.C., Lin, M.W., Bai, Y.M., Lin, C.C., Liao, D.L., Chen, J.Y., Lin, C.Y., Wang, Y.C. Haplotype analysis of endothelial nitric oxide synthase (NOS3) genetic variants and tardive dyskinesia in patients with schizophrenia. *Pharmacogenet. Genomics* **2006**, *16*: 151-7.

[71] Roberts, J., Scott, A.C., Howard, M.R., Breen, G., Bubb, V.J., Klenova, E., Quinn, J.P. Differential regulation of the serotonin transporter gene by lithium is mediated by transcription factors, CCCTC binding protein and Y-box binding protein 1, through the polymorphic intron 2 variable number tandem repeat. *J. Neurosci.* **2007**, *27*: 2793-801.

[72] O'Gara, C., Stapleton, J., Sutherland, G., Guindalini, C., Neale, B., Breen, G., Ball, D. Dopamine transporter polymorphisms are associated with short-term response to smoking cessation treatment. *Pharmacogenet Genomics* **2007**, *17*: 61-7.

[73] Tadic, A., Rujescu, D., Muller, M.J., Kohnen, R., Stassen, H.H., Dahmen, N., Szegedi, A. A monoamine oxidase B gene variant and short-term antidepressant treatment response. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2007**, *31*: 1370-7.

[74] Ando, T., Ichimaru, Y., Konjiki, F., Shoji, M., Komaki, G. Variations in the preproghrelin gene correlate with higher body mass index, fat mass, and body dissatisfaction in young Japanese women. *Am. J. Clin. Nutr.* **2007**, *86*: 25-32.

[75] Schlaepfer, I.R., Clegg, H.V., Corley, R.P., Crowley, T.J., Hewitt, J.K., Hopfer, C.J., Krauter, K., Lessem, J., Rhee, S.H., Stallings, M.C., Wehner, J.M., Young, S.E., Ehringer, M.A. The human protein kinase C gamma gene (PRKCG) as a susceptibility locus for behavioral disinhibition. *Addict. Biol.* **2007**, *12*: 200-9.

[76] Borroni, B., Grassi, M., Costanzi, C., Zanetti, M., Archetti, S., Franzoni, S., Caimi, L., Padovani, A. Haplotypes in cathechol-O-methyltransferase gene confer increased risk for psychosis in Alzheimer disease. *Neurobiol. Aging* **2007**, *28*: 1231-8.

[77] Dannlowski, U., Ohrmann, P., Bauer, J., Kugel, H., Baune, B.T., Hohoff, C., Kersting, A., Arolt, V., Heindel, W., Deckert, J., Suslow, T. Serotonergic genes modulate amygdala activity in major depression. *Genes Brain Behav.* **2006**.

[78] Gawlik, M., Moller-Ehrlich, K., Mende, M., Jovnerovski, M., Jung, S., Jabs, B., Knapp, M., Stoeber, G. Is FKBP5 a genetic marker of affective psychosis? A case control study and analysis of disease related traits. *BMC Psychiatry* **2006**, *6*: 52.

[79] Tiwari, A.K., Deshpande, S.N., Lerer, B., Nimgaonkar, V.L., Thelma, B.K. Genetic susceptibility to Tardive Dyskinesia in chronic schizophrenia subjects: V. Association of CYP1A2 1545 C>T polymorphism. *Pharmacogenomics J.* **2006**.

[80] Lohoff, F.W., Dahl, J.P., Ferraro, T.N., Arnold, S.E., Gallinat, J., Sander, T., Berrettini, W.H. Variations in the vesicular monoamine transporter 1 gene (VMAT1/SLC18A1) are associated with bipolar i disorder. *Neuropsychopharmacology* **2006**, *31*: 2739-47.

[81] Lachman, H.M., Pedrosa, E., Nolan, K.A., Glass, M., Ye, K., Saito, T. Analysis of polymorphisms in AT-rich domains of neuregulin 1

gene in schizophrenia. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2006,** *141*: 102-9.

[82] Xu, C., Macciardi, F., Li, P.P., Yoon, I.S., Cooke, R.G., Hughes, B., Parikh, S.V., McIntyre, R.S., Kennedy, J.L., Warsh, J.J. Association of the putative susceptibility gene, transient receptor potential protein melastatin type 2, with bipolar disorder. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **2006,** *141*: 36-43.

[83] Zaboli, G., Gizatullin, R., Nilsonne, A., Wilczek, A., Jonsson, E.G., Ahnemark, E., Asberg, M., Leopardi, R. Tryptophan hydroxylase-1 gene variants associate with a group of suicidal borderline women. *Neuropsychopharmacology* **2006,** *31*: 1982-90.

[84] Viana, M.M., De Marco, L.A., Boson, W.L., Romano-Silva, M.A., Correa, H. Investigation of A218C tryptophan hydroxylase polymorphism: association with familial suicide behavior and proband's suicide attempt characteristics. *Genes Brain Behav.* **2006,** *5*: 340-5.

[85] Yamada, K., Hattori, E., Iwayama, Y., Ohnishi, T., Ohba, H., Toyota, T., Takao, H., Minabe, Y., Nakatani, N., Higuchi, T., Detera-Wadleigh, S.D., Yoshikawa, T. Distinguishable haplotype blocks in the HTR3A and HTR3B region in the Japanese reveal evidence of association of HTR3B with female major depression. *Biol. Psychiatry* **2006,** *60*: 192-201.