

# Trends in meta-analysis of genetic association studies

Elias Zintzaras · Joseph Lau

Received: 27 August 2007 / Accepted: 5 November 2007 / Published online: 12 December 2007  
© The Japan Society of Human Genetics and Springer 2007

**Abstract** The number of published genetic association studies (GASs) is increasing tremendously due to the availability of mapped single-nucleotide polymorphisms (SNPs) and advances in genotyping technologies. A search in HUGENet illustrates the rapid accumulation of evidence for major diseases. Recently, there has been a lot of activity regarding genome-wide association studies (GWASs), and a growing number of forthcoming studies is expected. GASs and GWASs are usually underpowered to detect significant associations, and the varying quality of reporting publications befuddles researchers. A meta-analysis can increase power and provide standards of reporting results. However, the conduct of a meta-analysis of GASs faces a major obstacle, which is the structure and diversity of stored information in databases. Similar problems are expected for GWASs, though the data are not yet publicly available. The development of a Web-based system for the detailed and structured recording of GAS or GWAS data, accompanied by an estimation of the overall genetic risk effects, would enable scientists to keep track of evidence for gene–disease associations.

**Keywords** Meta-analysis · Epidemiology · Genetics · Genomics · Polymorphism · Association · Database · Quality

---

E. Zintzaras (✉)  
Department of Biomathematics, University of Thessaly School of Medicine, Papakyriazi 22, Larissa 41222, Greece  
e-mail: zintza@med.uth.gr

E. Zintzaras · J. Lau  
Center for Clinical Evidence Synthesis, Institute for Clinical Research and Health Policy Studies, Tufts—New England Medical Center, Tufts University School of Medicine, Boston, MA, USA

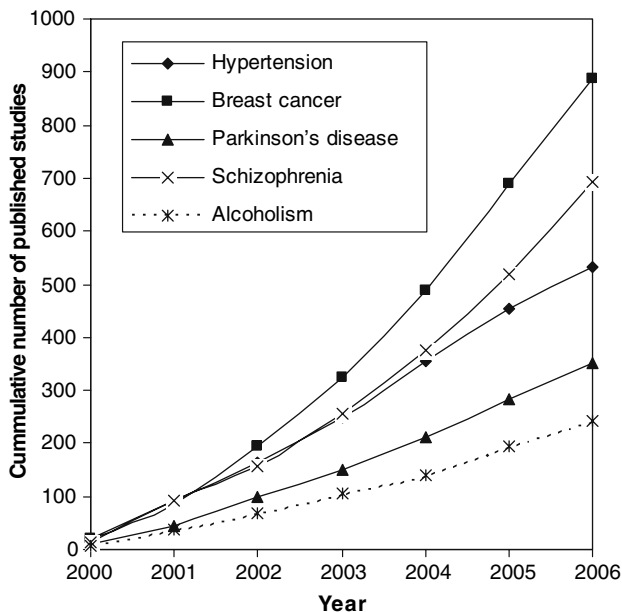
## Introduction

Genetic association studies (GASs), or candidate–gene studies, assess the association between disease status and genetic variants (gene polymorphisms) in a population. They have been particularly popular for investigating complex diseases and the number of papers on GASs has increased tremendously. This trend is expected to accelerate because of the rapidly increasing availability of mapped single-nucleotide polymorphisms (SNPs) and advances in genotyping technologies (Donahue and Allen 2005). Dealing with all of the accumulated evidence, especially when studies frequently produce controversial or inconclusive results, is a major challenge. Meta-analysis provides a tool to estimate population-wide genetic risk effects for pertinent gene–disease associations, therefore, helping to resolve contradictory results and to decrease the uncertainty of the estimated risk. Since 2002, the molecular research of genotype–phenotype associations has entered the genome-wide era, as a consequence of the completion of the Human Genome and HapMap projects and the development of ultra-high-volume genotyping chip technologies. In contrast to GASs, the genome-wide association studies (GWASs) are hypothesis-free (i.e., unbiased) and involve a massive scan of the genome with a dense set of SNPs (up to 500,000) in a single experiment, in search of causal variants. Thus, GWASs represent a comprehensive option to GASs when there is a lack of evidence regarding the function or the location of the causal genes. However, the vast amount of data produced by GWASs represent a major methodological challenge both in their primary analysis and in their meta-analysis (Thomas and Witte 2002; Hirschhorn and Daly 2005; Zintzaras and Lau 2007).

## Accumulation of evidence

A search in HuGENet (<http://www.cdc.gov/genomics/hugenet/default.htm>) conducted for the period January 2000 to December 2006 found 535, 887, 351, 694, and 249 GASs published in hypertension (Zintzaras et al. 2006a, 2006b), schizophrenia (Zintzaras 2006a, 2006b, 2007), Parkinson's disease (Zintzaras and Hadjigeorgiou 2004, 2005), breast cancer (Zintzaras 2006a, 2006b), and alcoholism (Zintzaras et al. 2006a, 2006b), respectively (see Fig. 1). These five topics are among many other common complex diseases with a major impact in public health, and whose underlining pathogenetic mechanisms are not clearly understood. In breast cancer and Parkinson's disease, the number of studies published each year after 2002 is consistently above 100 and 50 studies, respectively, and in schizophrenia, there is an upward trend, with an average growth rate of 75% per year. In hypertension and alcoholism, the annual number of published studies on average is 76 and 35, respectively. Therefore, the same or even a larger number of publications of GASs in these topics in the coming years can be expected, and the pattern of published GASs is likely to be similar for other topics.

A search in PubMed for GWASs (August 2007) has yielded 36 articles describing 42 studies for various multifactorial diseases, such as coronary artery disease, Crohn's disease, Parkinson's disease, age-related macular degeneration, and Alzheimer's disease. Moreover, an increasing number of funded NIH and European initiatives



**Fig. 1** Cumulative frequency of published genetic association studies in HuGENet for five main complex diseases

have been undertaken in this direction during the last few years, leading to a growing number of forthcoming and promising studies (Thomas and Witte 2002).

## Validity of genotype–phenotype associations

The published literature has demonstrated a plethora of questionable gene–disease associations, based on both the candidate gene and the genome-wide approach, the replication of which has often failed in independent studies (NCI-NHGRI Working Group on Replication in Association Studies 2007). Although replication is essential for establishing the credibility of a genotype–phenotype association, a number of methodological and design issues in the initial studies are hampering research.

Small sample size is a frequent problem and can result in insufficient power to detect minor contributing roles of one or more alleles. The most realistic genetic association between a polymorphic locus and a disease has been claimed to yield an odds ratio of between 1.1 and 1.5. Therefore, to achieve a satisfactory power (>80%) to identify a modest genetic effect (odds ratio 1.2) of a polymorphism present in 10% of individuals, a sample size of 10,000 subjects or more would be needed for a GAS. Likewise, for GWASs, testing for 500,000 SNP associations in a case–control study at a 5% significant level, a Bonferroni correction would require significance at  $p = 0.05/500,000 = 10^{-7}$ . Then, to attain a 95% power for OR = 1.2 and minor allele frequency of 10%, 15,000 case–control pairs are required in a single study, leading to a significant financial burden (Thomas and Witte 2002; Wang et al. 2005). The meta-analysis of multiple studies clearly has a role in offering an analysis with the potential for higher probability to detect significant results (Munafò and Flint 2004).

Population stratification can be a confounding factor in gene–disease associations and arises when differences in the genetic structure of the underlying population are not taken into account (Zintzaras and Sakelaridis 2007). Then, the cases and controls are not matched for their genetic background, which can lead to biased or spurious results (Cardon and Palmer 2003). Moreover, the quality design issues of individual studies, such as the definition of a phenotype, validity of the genotyping method, and heterogeneity in exposure to environmental challenges, can increase the risk of biases (Zintzaras and Stefanidis 2005; Zintzaras et al. 2007). A meta-analysis provides a robust tool to investigate discrepant results, to decrease the uncertainty of the effect size of estimated risk, and to explore the heterogeneity between studies (Zintzaras and Ioannidis 2005; Zintzaras and Lau 2007).

## The role of meta-analysis

The number of meta-analyses appearing in HuGENet is 9, 36, 10, 19, and 7, for hypertension, schizophrenia, Parkinson's disease, breast cancer, and alcoholism, respectively. These meta-analyses can provide answers to the question of whether there is evidence of an association between gene polymorphism and disease. Table 1 shows the meta-analyses' summary risk effects (odds ratios) for investigating associations between various gene polymorphisms and the five diseases of interest. The meta-analyses were based on all of the available studies at the time of performing the meta-analysis. In total, 96 gene polymorphisms were examined, and the number of studies included in the meta-analyses ranged from 2 to 48. Significant associations under any genetic contrast were found for four polymorphisms in hypertension, 26 in schizophrenia, four in Parkinson's disease, 11 in breast cancer, and six in alcoholism. An association was considered to be significant when the  $p$ -value was less than 0.05 or the 95% confidence interval of the odds ratio did not include 1.0. So far (PubMed accessed August 2007), one meta-analysis in the field of GWASs for Parkinson's disease has been published, synthesizing only three studies (Evangelou et al. 2007), but there is still a need for establishing a proper methodology for the meta-analysis of genome-wide rich data (Zintzaras and Ioannidis 2007).

The methodological issues relating to meta-analyses have been previously described in detail (Munafò and Flint 2004; Zintzaras and Lau 2007), and are beyond the scope of this paper. However, the benefit of getting an assessment of the overall risk effects from a meta-analysis is obvious. As evidence is accumulating rapidly, the updating of genetic risk effects can provide information on whether an association is real, or that more evidence is needed in order to draw reliable conclusions on the association (Neale and Sham 2004). However, a meta-analysis requires a large amount of labor and effort, since it involves systematic searches in databases, article retrieval, data extraction, data entry, and data analysis (Hirschhorn et al. 2002).

## Obstacles in meta-analysis

The major obstacle in undertaking a meta-analysis of GASs is the structure and diversity of stored information in databases, such as the Genetic Association Database (<http://geneticassociationdb.nih.gov/cgi-bin/index.cgi>), PubMed (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>), EMBASE (<http://www.embase.com/>), HuGENet, and OMIM (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>). The information provided in the databases is not structured and standardized appropriately

for meta-analysis use, and they are not comprehensive. These databases, though important for promoting research, do not claim to be and are not a substitute for the meta-analysis of GASs. Thus, gathering and meta-analyzing data related to an SNP association with a disease/disorder is not straightforward. The minimum information that the articles should provide for meta-analysis is the genotype frequencies, the 'race' of samples, the study design, the demographic characteristics, and possible effect modifiers or confounders. In addition, the literature is filled with alternative, idiosyncratic, and arbitral gene and SNP names and symbols, making cross-comparison and meta-analysis difficult [e.g., an SNP can be found with more than one arbitral name, designating amino acid or nucleotide substitutions or else, and rarely with the official dbSNP ID rs-number (<http://www.ncbi.nlm.nih.gov/projects/SNP/>)] (Kitsios and Zintzaras 2007; Zdoukopoulos and Zintzaras 2007). Much of the information that the biological researcher is interested in is available in public reference databases and in the millions of articles of the scientific research literature, mostly accessible via the Internet. It is estimated that about 80% of biological data are in text form, and even the abstracts are written in free text utilizing a complex biological vocabulary, which may vary significantly in different areas of research. Thus, despite their wide availability, these data are not generally machine-searchable. Consequently, no promising or significant increase in the efficiency of data integration may be expected, for example, from the automation of PubMed data retrieval (Teufel et al. 2006; Lacroix 2002; Colhoun 2003). Regarding GWASs, no publicly available database is currently storing the enormous amount of accumulating data.

## Improving the quality of reporting genotype–phenotype association

The situation of GASs is similar to the problem clinical medicine faced more than 10 years ago, where the large number of randomized control trials of varying quality befuddled clinicians (Lau et al. 1992). Then, concerted efforts by the medical community were initiated to improve the quality of reporting research publications. For example, the CONSORT statement (Begg et al. 1996) was published with a view to improving the quality of the reporting of randomized controlled trials, the QUOROM statement (Moher et al. 1999) was focused on the reporting of meta-analyses, and the STARD initiative (Bossuyt et al. 2003) was developed to improve the reporting of studies of diagnostic accuracy. The Cochrane Collaboration was created in response to the need for collecting, synthesizing, and disseminating the effect of health care in prevention. This is a successful model that should be emulated by other

**Table 1** Results of 81 meta-analyses of genetic association studies shown in HuGENet. The conclusive pooled odds ratio (OR) and the corresponding 95% confidence interval (CI) as provided by individual meta-analyses are shown. When the OR was not applicable, the *p*-value for testing the association is shown

Disease	Author (year)	No. of studies	Gene polymorphism; contrast	OR (95%CI) or <i>p</i> -value	
Hypertension	Sookoian (2007)	42	CYP11B2 C-344T; CC vs. TT	0.83 (0.76–0.91)	
	Zintzaras (2006)	16	eNOS G894T; G vs. T	1.00 (0.95–1.06)	
		20	eNOS G894T; (GG + GT) vs. TT	0.98 (0.92–1.05)	
		15	eNOS G894T; GG vs. (GT + TT)	0.97 (0.84–1.12)	
		12	eNOS 4a/b; b vs. a	0.85 (0.74–0.98)	
		15	eNOS 4a/b; (bb + ba) vs. aa	0.78 (0.68–0.90)	
		11	eNOS 4a/b; bb vs. (ba + aa)	1.38 (0.85–2.24)	
		6	eNOS T786C; T vs. C	0.92 (0.82–1.02)	
		7	eNOS T786C; (TT + TC) vs. CC	0.93 (0.82–1.05)	
		6	eNOS T786C; TT vs. (TC + CC)	0.87 (0.66–1.16)	
		3	eNOS G23T; G vs. T	0.88 (0.72–1.08)	
		3	eNOS G23T; (GG + GT) vs. TT	0.84 (0.64–1.11)	
		3	eNOS G23T; GG vs. (GT + TT)	0.88 (0.58–1.35)	
		Andersen (2006)	11	GNB3 C825T; TT vs. (TC + CC)	1.08 (1.01–1.17)
			11	GNB3 C825T; TC vs. CC	1.07 (1.01–1.13)
	Mondry (2005)	25	AGT M235T; TT vs. MM	1.21(1.11–1.32)	
	Sethi (2003)	22	AGT M235T; MT vs. MM	1.08 (1.01–1.15)	
		22	AGT M235T; TT vs. MM	1.19 (1.10–1.30)	
		12	AGT M235T; TT vs. MM	1.29 (0.96–1.74)	
		12	AGT M235T; MT vs. MM	1.60 (1.19–2.15)	
		6	AGT M235T; TT vs. MM	0.94 (0.45–1.96)	
		6	AGT M235T; MT vs. MM	1.16 (0.56–2.56)	
		Chen (2003)	10	AGT M235T; TT vs. (TM + MM)	1.76 (1.44– 2.16)
			10	AGT M235T; T vs. M	1.54 (1.31– 1.81)
	Qu (2001)	18	ACE I/D; DD vs. DI + II	1.37 (1.15–1.63)	
	Kato (1999)	6	AGT M235T; T vs. M	1.22 (1.05–1.42)	
	Staessen (1999)	32	AGT M235T; TT vs. MM	1.31 <i>p</i> < 0.01	
		32	AGT M235T; MT vs. MM	1.11 <i>p</i> = 0.03	
		28	ACE I/D; DD vs. (DI + II)	1.10 (0.95–1.27)	
		28	ACE I/D; DI vs. II	1.06 (0.93–1.22)	
Schizophrenia		Dawei (2007)	6	GRIN2B 2664C/T; C vs. T	0.92 (0.81–1.04)
			6	GRIN2B 2664C/T; (CC + CT) vs. TT	0.95 (0.76–1.20)
	6		GRIN2B 2664C/T; CC vs. (CT + TT)	0.85 (0.70–1.03)	
	6		GRINB2 366C/G; C vs. G	0.96 (0.83–1.10)	
	6		GRINB2 366C/G; (CC + CG) vs. GG	1.00 (0.66–1.51)	
	6		GRINB2 366C/G; CC vs. (CG + GG)	0.89 (0.72–1.11)	
	Gilbody (2006)	12	MTHFR C677T; T vs. C	1.17 (1.08–1.26)	
		12	MTHFR C677T; TT vs. CC	1.44 (1.21–1.70)	
		12	MTHFR C677T; TC vs. CC	1.07 (0.96–1.20)	
	Guo (2006)	7	RGS4 rs10917670; G vs. A	1.03 (0.92–1.16)	
7		RGS4 rs951436; G vs. T	1.00 (0.89–1.12)		
7		RGS4 rs951439; A vs. G	1.02 (0.90–1.15)		
7		RGS4 rs2661319; G vs. A	1.02 (0.95–1.10)		
Detera-Waldeigh (2006)		2	G72/G30 rs746187; T vs. C	<i>p</i> = 0.60	
	3	G72/G30 rs3916964; T vs. C	<i>p</i> = 0.11		
	4	G72/G30 rs3916965; G vs. A	<i>p</i> < 0.01		

**Table 1** continued

Disease	Author (year)	No. of studies	Gene polymorphism; contrast	OR (95%CI) or <i>p</i> -value
		2	G72/G30 rs3916966; C vs. A	<i>p</i> = 0.07
		4	G72/G30 rs3916967; G vs. A	<i>p</i> = 0.07
		5	G72/G30 rs2391191; G vs. A	<i>p</i> < 0.01
		2	G72/G30 rs3918341; G vs. A	<i>p</i> = 0.60
		3	G72/G30 rs778294; G vs. A	<i>p</i> = 0.94
		3	G72/G30 rs3916971; T vs. C	<i>p</i> = 0.09
		5	G72/G30 rs778293; G vs. A	<i>p</i> < 0.01
		5	G72/G30 rs3918342; T vs. C	<i>p</i> < 0.01
		3	G72/G30 rs1421292; T vs. A	<i>p</i> < 0.01
	Talkowski (2006)	13	RGS4 rs10917670; G vs. A	1.26 (0.97–1.64)
		13	RGS4 rs951436; G vs. T	1.20 (1.04–1.39)
		13	RGS4 rs951439; A vs. G	1.25 (1.06–1.48)
		13	RGS4 rs2661319; G vs. A	1.171 (1.01–1.36)
	Jonsson (2006)	6	BDNF 270 C/T; T vs. C	1.35 (1.03–1.77)
		11	BDNF Val66Met; Val vs. Met	1.02 (0.93–1.11)
		11	BDNF Val66Met; ValVal vs. MetVal	1.16 (1.04–1.30)
	Jonsson (2006)	5	XBP1 116C/G; GG vs. (GC + CC)	1.04 (0.91–1.19)
		5	XBP1 116C/G; G vs. C	1.08 (0.98–1.19)
		5	XBP1 116C/G; (GG + GC) vs. CC	1.26 (1.01–1.57)
	Xu (2006)	17	APOE e4; e4-alleles vs. others	1.08 (0.88–1.33)
		17	APOE e4; e4-carriers vs. non-carriers	1.09 (0.87–1.36)
	Jeong (2006)	4	FZD3 rs960914; C vs. T	1.079 (0.88–1.32)
		6	FZD3 rs2241802; A vs. G	1.142 (0.95–1.38)
	Munafò (2006)	19	NRG1 SNP8NRG221533; T vs. C	1.04 (0.99–1.10)
	Li (2006)	3	NRG1 SNP8NRG221132; G vs. A	1.27 (1.06–1.51)
		12	NRG1 SNP8NRG221533; T vs. C	1.09 (1.02–1.16)
		8	SNP8NRG241930; G vs. A	1.18 (1.08–1.29)
		8	SNP8NRG243177; T vs. C	1.1 (1.02–1.19)
	Zintzaras (2006)	10	MTHFR C677T; T vs. C	1.13 (1.04–1.23)
		10	MTHFR C677T; TT vs. CC	1.34 (1.12–1.60)
		10	MTHFR C677T; TT + TC vs. CC	1.21 (1.00–1.25)
		10	MTHFR C677T; TT vs. TC + CC	1.32 (1.12–1.56)
		4	MTHFR A1298C; C vs. A	1.16 (1.03–1.31)
		4	MTHFR A1298C; CC vs. AA	1.37 (1.03–1.82)
		4	MTHFR A1298C; CC + AC vs. AA	1.19 (1.02–1.40)
		4	MTHFR A1298C; CC vs. AC + AA	1.27 (0.97–1.67)
	Glatt (2006)	27	DRD2 Ser311Cys; Cys/Ser vs. Ser/Ser	1.38 (1.11–1.72)
		27	DRD2 Ser311Cys; Cys/Cys vs. Cys/Ser	1.03 (0.35–3.05)
	Muntjewerff (2006)	8	MTHFR C677T; TT vs. TC + CC	1.70 (1.27–2.29)
	Li (2006)	48	5-HT2A T102C; T vs. C	0.94 (0.87–1.02)
		42	5-HT2A T102C; TT + TC vs. CC	0.91 (0.8–1.03)
		44	5-HT2A T102C; TT vs. TC + CC	0.97 (0.9–1.05)
	Fan (2005)	19	SLC6A4 S/L; S vs. L	0.99 (0.92–1.07)
		12	SLC6A4 in2 VNTR; allele 12 vs. others	1.24 (1.11–1.38)
	Munafò (2005)	18	COMT Val158Met; Met vs. Val	1.06 (0.99–1.13)
	Lewis (2005)	6	MTHFR C677T; TT vs. TC + CC	1.48 (1.18–1.86)
		6	MTHFR C677T; TC vs. CC	1.04 (0.87–1.25)
	Glatt (2005)	15	NOTCH4 (TAA)n8; n8 vs. other alleles	1.02 (0.87–1.21)

**Table 1** continued

Disease	Author (year)	No. of studies	Gene polymorphism; contrast	OR (95%CI) or <i>p</i> -value	
Parkinson's disease	Gamma (2005)	10	NOTCH4 (CTG)n10; n10 vs. other alleles	0.98 (0.89–1.08)	
		6	NOTCH4 MspI (1); T vs. C	0.99 (0.87–1.13)	
		9	NOTCH4 MspI (2); G vs. A	0.98 (0.89–1.08)	
		6	SLC6A3 (40 bp)n; n9 vs. other alleles	0.90 (0.70–1.10)	
		6	SLC6A3 (40 bp)n; n10 vs. other alleles	1.00 (0.80–1.30)	
		Fan (2005)	23	COMT Val158Met; Met vs. Val	1.03 (0.94–1.11)
		Lin (2004)	8	NT3 (CA)n; A3 vs. other alleles	1.16 (0.95–1.43)
			9	CNTF G > A; A vs. G	1.05 (0.90–1.23)
		Duan (2004)	10	TNFA G308A; A vs. G	1.00 ( <i>p</i> < 0.01)
		Glatt (2004)	10	DRD2 141C ins/del; ins vs. del	1.10 (0.80–1.40)
	Abdolmaleky (2004)	31	5HT2a receptor T102C; C vs. T	1.10 (1.00–1.20)	
		31	5HT2a receptor T102C; CC vs. TC + TT	1.30 (1.00–1.60)	
	Glatt (2003)	24	DRD2 Cys311Ser; Cys vs. Ser	1.30 ( <i>p</i> < 0.01)	
	Uchida (2003)	3	SIGMAR1 Gln2Pro; Pro vs. Gln	<i>p</i> = 0.06	
	Glatt (2003)	17	DRD4 exon 3 VNTR; A2 vs. other	0.90 (0.80–1.00)	
		17	DRD4 exon 3 VNTR; A4 vs. other	1.00 (0.90–1.10)	
		17	DRD4 exon 3 VNTR; A7 vs. other	1.10 (0.90–1.30)	
	Glatt (2003)	11	KCNN3 CAG-Repeat; S vs. L	1.20 (1.00–1.50)	
	Jonsson (2003)	25	DRD2 Ser311Cys; Cys vs. Ser	1.43 (1.16–1.78)	
	Schurhoff (2003)	15	APOE e2; e2-carriers vs. others	(0.77–1.10)	
		15	APOE e3; e3-carriers vs. others	(0.88–1.11)	
		15	APOE e4; e4-carriers vs. non carriers	(0.92–1.22)	
	Jonsson (2003)	20	DRD4 exon 3 VNTR; L vs. S	0.95 (0.85–1.07)	
		7	DRD4 exon 1 VNTR; L vs. S	1.03 (0.74–1.42)	
		3	DRD4 521C/T; C vs. T	1.22 (1.04–1.42)	
	Glatt (2003)	14	COMT Val158Met; Met vs. Val	1.00 (0.90–1.10)	
	Jonsson (2003)	8	DRD3 Ser9Gly; SerSer vs. SerGly + Glygly	1.10 (1.01–1.20)	
	Lung (2002)	12	DRD4 exon 3 VNTR; L vs. S	6.24 ( <i>p</i> < 0.01)	
	Philibert (2006)	4 (male)	HOPA 12 bp(ins); 12 bp(ins) vs. Wt	1.89 (1.03–3.47)	
		4 (female)	HOPA 12 bp(ins); 12 bp(ins) vs. Wt	1.93 (1.01–3.67)	
	Borlak (2006)	5	NAT2 G509A; A vs. G	0.94 (0.79–1.12)	
		5	NAT2 G509A; AA vs. (AG + GG)	0.94 (0.78–1.43)	
	Healy (2006)	8	UCHL-1 S18Y; (YY + SY) vs. SS	1.01 (0.76–1.35)	
		8	UCHL-1 S18Y; YY vs. (SY+SS)	0.96 (0.86–1.08)	
		8	UCHL-1 S18Y; YY vs. SS	1.00 (0.74–1.33)	
	Zintzaras (2005)	6	BDNF G196A; A vs. G	1.00 (0.90–1.12)	
		6	BDNF G196A; AA vs. GG	0.91 (0.58–1.41)	
		6	BDNF G196A; AA vs. (AG + GG)	1.08 (0.89–1.31)	
		6	BDNF G196A; (AG + AA) vs. GG	0.94 (0.80–1.12)	
		7	NACP-Rep1 (0, +1); 0 vs. +1	0.79 (0.70–0.89)	
Zhang (2005)	14	tau gene haplotype h1 ; H1H1 vs. others	1.42 (1.23–1.65)		
	14	tau gene haplotype h1 ; H1-carriers vs. others	1.52 (1.12–2.04)		
Zintzaras (2005)	4	PON1-55M/L; M vs. L	1.32 (1.10–1.59)		
	4	PON1-55M/L; (MM + LM) vs. LL	1.45 (1.13–1.86)		
	4	PON1-55M/L; MM vs. LL	1.64 (1.10–2.46)		
	4	PON1-55M/L; MM vs. (LM + LL)	1.34 (0.92–1.94)		
	5	PON1-192Q/R; R vs. Q	1.09 (0.93–1.26)		
	5	PON1-192Q/R; RR vs. (QR + QQ)	1.15 (0.90–1.49)		

**Table 1** continued

Disease	Author (year)	No. of studies	Gene polymorphism; contrast	OR (95%CI) or <i>p</i> -value	
Breast cancer	Huang (2004)	5	PON1-192Q/R; RR vs. QQ	1.11 (0.78–1.56)	
		5	PON1-192Q/R; (RR + QR) vs. QQ	1.08 (0.86–1.36)	
		22	APOE e2; e2-carriers vs. others	1.2 (1.02–1.42)	
		22	APOE e3; e3-carriers vs. others	0.83 (0.63–1.09)	
		22	APOE e4; e4-carriers vs. non carriers	0.99 (0.87–1.14)	
	Persad (2003)	9	CYP2D6; 4*A vs. 4*G	0.84 (0.66–1.08)	
	Christensen (1998)	15	CYP2D6; 4*A vs. 4*G	1.48 (1.10–1.99)	
	Rostami (1998)	18	CYP2D6; 4*A vs. 4*G	1.32 (0.98–1.78)	
	Masson (2005)	17	CYP1A1 3801T→C; C/C vs. TT	0.97 (0.52–1.80)	
		17	CYP1A1 3801T→C; TC vs. TT	0.91 (0.70–1.19)	
		17	CYP1A1 Ile462Val; ValVal vs. IleIle	1.04 (0.63–1.74)	
		17	CYP1A1 Ile462Val; IleVal vs. IleIle	0.92 (0.76–1.10)	
		17	CYP1A1 Thr461Asp; AspAsp vs. ThrThr	0.95 (0.20–4.49)	
		17	CYP1A1 Thr461Asp; TheAsp vs. ThrThr	1.12 (0.87–1.43)	
		Zintzaras (2006)	18	MTHFR C677T; T vs. C	1.02 (0.95–1.10)
			18	MTHFR C677T; TT vs. CC	1.07 (0.90–1.27)
			18	MTHFR C677T; TT vs. (TC + CC)	1.06 (0.91–1.24)
			18	MTHFR C677T; (TT + TC) vs. CC	1.02 (0.94–1.11)
	18		MTHFR A1298C; A vs. C	0.98 (0.89–1.07)	
	18		MTHFR A1298C; AA vs. CC	1.04 (0.81–1.33)	
	18		MTHFR A1298C; AA vs. (AC + CC)	0.95 (0.87–1.03)	
	18		MTHFR A1298C; (AA + AC) vs. CC	1.06 (0.84–1.35)	
	Terry (2006)	14	NAT2 rapid acetylators	1.20 (1.00–1.50)	
		14	NAT2 slow acetylators	1.50 (1.20–1.80)	
		14	GSTM1 D/I; D/D vs. (D/I + I/I)	1.50 (1.10–2.10)	
	Wen (2005)	4	CYP1B2 Ala119Ser; Ala/Ser vs. Ala/Ala	1.07 (0.91–1.25)	
		4	CYP1B2 Ala119Ser; Ser/Ser vs. Ala/Ala	0.98 (0.81–1.19)	
		9	CYP1B2 Leu432Val; Leu/Val vs. Leu	1.06 (0.90–1.26)	
		9	CYP1B2 Leu432Val; Val/Val vs. Leu/Leu	1.01 (0.88–1.15)	
		3	CYP1B2 Asn453Ser; Asn/Ser vs. Asn/Asn	0.91 (0.79–1.04)	
		3	CYP1B2 Asn453Ser; Ser/Ser vs. Asn/Asn	0.85 (0.54–1.34)	
		14	COMT Val158Met; Val/Met vs. Val/Val	0.96 (0.83–1.12)	
		14	COMT Val158Met; Met/Met vs. Val/Val	0.89 (0.74–1.07)	
	Sull (2004)	30	GSTM1; D/D vs. (D/I + I/I)	1.03 (0.94–1.12)	
	Egan (2004)	19	GSTM1; D/D vs. (D/I + I/I)	1.05 (0.98–1.13)	
		15	GSTT1; D/D vs. (D/I + I/I)	1.11 (1.01–1.22)	
	Ye (2002)	10	GSTP1 Ile105Val; Val/Val vs. Ile/Ile	1.04 (0.87–1.25)	
		15	CYP17 MspA1; (A1/A2 + A2/A2) vs. A1/A1	0.98 (0.89–1.07)	
		15	CYP17 MspA2; A2/A2 vs. A1/A1	1.05 (0.87–1.21)	
	Han (2006)	48	XRCC3 Thr(241)Met; Met/Met vs. Thr/Thr	1.14 (1.06–1.23)	
	Garcia-Closas (2006)	11	XRCC3 Thr(241)Met; Thr/Met vs. Thr/Thr	1.03 (0.97–1.09)	
		11	XRCC3 Thr(241)Met; Met/Met vs. Thr/Thr	1.16 (1.04–1.30)	
10		BRCA2 N372H; His/His vs. Asn/Asn	1.13 (1.01–1.28)		
10		XRCC2 R188H; His/His vs. Arg/Arg	1.06 (0.59–1.91)		
Kaklamani (2003)	7	TGFBR1, *6A-carriers vs. others	1.48 (1.11–1.96)		
Hung (2005)	9	XRCC1 Arg194Trp; Arg/Trp vs. Arg/Arg	0.94 (0.77–1.15)		
	5	XRCC1 Arg194Trp; Trp/Trp vs. Arg/Arg	1.07 (0.64–1.84)		
	9	XRCC1 Arg194Trp; (Arg/Trp + Trp/Trp) vs. Arg/Arg	0.95 (0.78–1.16)		

**Table 1** continued

Disease	Author (year)	No. of studies	Gene polymorphism; contrast	OR (95%CI) or <i>p</i> -value	
Alcoholism	Paracchini (2006)	11	XRCC1 Arg399Gln; Arg/Gln vs. Arg/Arg	1.06 (0.94–1.20)	
		11	XRCC1 Arg399Gln; Gln/Gln vs. Arg/Arg	1.17 (0.98–1.39)	
		11	XRCC1 Arg399Gln; (Arg/Gln + Gln/Gln) vs. Arg/rg	1.07 (0.96–1.20)	
		13	CYP1B1 Leu432Val; Val/Val vs. Leu/Leu	1.00 (0.60–1.60)	
		13	CYP1B1 Leu432Val; Val/Leu vs. Leu/Leu	1.00 (0.80–1.20)	
		13	CYP1B1 Leu432Val; (Val/Val + Val/Leu) vs. Leu/Leu	1.00 (0.80–1.20)	
		Spinola (2004)	3	L-myc EcoRI; S/S vs. L/L	1.90 (1.10–3.10)
		Cox (2006)	5	AURKA F31I; I/I vs. F/F	1.28 (1.08–1.53)
		Manuguerra (2006)	3	XPD/ERCC2 Asp312Asn (C/A); AA vs. GG	0.87 (0.34–2.23)
			3	XPD/ERCC2 Asp312Asn (C/A); (GA + AA) vs. GG	1.08 (0.44–2.65)
			3	XPD/ERCC2 Asp312Asn (C/A); GA vs. GG	1.12 (0.46–2.74)
			4	XPD/ERCC2 Lys751Gln (A/C); CA vs. AA	1.15 (1.00–1.32)
	4		XPD/ERCC2 Lys751Gln (A/C); CC vs. AA	1.22 (1.00–1.48)	
	4		XPD/ERCC2 Lys751Gln (A/C); (CA + CC) vs. AA	1.17 (1.02–1.33)	
	3		XRCC3 Thr241Met (C/T); CT vs. CC	0.91 (0.80–1.03)	
	3		XRCC3 Thr241Met (C/T); TT vs. CC	1.11 (0.94–1.33)	
	Fletcher (2006)	3	XRCC3 Thr241Met (C/T); (CT + TT) vs. CC	0.95 (0.85–1.07)	
		6	STK15 F31I; Ile/Ile vs. Phe/Phe	1.19 (0.96–1.48)	
	Zhang (2006)	14	XRCC1 Q399R; cAG vs. GG	1.10 (1.00–1.20)	
		14	XRCC1 Q399R; AG vs. GG	1.10 (1.00–1.10)	
	Pooley (2006)	11	XRCC1 R194W; CT/TT vs. CC	1.00 (0.80–1.20)	
		4	PGR-12 V660L; VL vs. VV	1.13 (1.03–1.24)	
	Wen (2005)	4	PGR-12 V660L; LL vs. VV	1.30 (0.98–1.73)	
		4	IGF-I; (CA)17 or (CA)19-carriers vs. others	1.22 (1.06–1.41)	
	Gorwood (2004)	16	5-HTT (S/L); S vs. L	1.31	
		6	ADH1 (R/H); RR vs. RH	2.11 (1.32–3.44)	
	Whitfield (2002)	4	ADH1 (R/H); RH vs. HH	1.36 (1.07–1.72)	
		4	ADH1 (R/H); RH vs. HH	1.70 (1.34–2.17)	
	Noble (2003)	21	DRD2 (Taq1); A1 vs. A2	1.53 (1.31–1.77)	
	Luczak (2006)	22	ALDH2 (*1/*2); *1/*2 vs. *1/*1	0.22 (0.16–0.30)	
		22	ALDH2 (*1/*2); *2/*2 vs. *1/*1	0.12 (0.06–0.24)	
		22	ALDH2 (*1/*2); *2/*2 vs. *1/*2	0.49 (0.23–1.05)	
	Feinn (2005)	19	5-HTT (S/L); S vs. L	1.18 (1.03–1.33)	
	Zhu (2003)	4	NPY Leu7Pro; Pro vs. Leu	1.09 (0.81–1.33)	
	Zintzaras (2006)	33	ADH2 (*1/*2); 1 vs. 2	1.89 (1.56–2.28)	
		33	ADH2 (*1/*2); *1/*1 vs. (*1/*2 + *2/*2)	2.62 (1.88–3.66)	
		33	ADH2 (*1/*2); (*1/*1 + *1/*2) vs. (*2/*2)	1.87 (1.54–2.28)	
		21	CYP2E1 (*1/*2); 1 vs. 2	0.92 (0.82–1.04)	
		24	ADH3 (*1/*2); 1 vs. 2	1.32 (1.12–1.57)	
		24	ADH3 (*1/*2); *1/*1 vs. (*1/*2 + *2/*2)	1.35 (1.07–1.71)	
		24	ADH3 (*1/*2); (*1/*1 + *1/*2) vs. (*2/*2)	1.35 (1.08–1.67)	
		23	ALDH2 (*1/*2); *1 vs. *2	4.35 (3.04–6.23)	
23		ALDH2 (*1/*2); *1/*1 vs. (*1/*2 + *2/*2)	4.16 (2.94–5.89)		
23		ALDH2 (*1/*2); (*1/*1 + *1/*2) vs. (*2/*2)	39.77 (12.97–122)		



scientific disciplines that have similar needs to synthesize evidence.

The need for data sharing has been highlighted by the Genetic Association Information Network (GAIN) initiative, in an effort to facilitate the subsequent and joint analysis of GWASs data (Thomas and Witte 2002). All genotypes will be made public as they are generated and checked for quality. However, the sheer scale of GWAS data will pose significant practical challenges regarding the form of the stored statistical results and their suitability for meta-analysis.

## Conclusion

Given the rapid accumulation of evidence regarding gene–disease associations, a Web-based system for data storage and automated meta-analysis of genetic association studies (GASs) and genome-wide association studies (GWASs) results is essential to: (1) keep track of the evidence for gene–disease associations, (2) improve the quality of reporting, and (3) reduce effort and labor for performing meta-analysis.

## References

- Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF (1996) Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 276:637–639
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG; Standards for Reporting of Diagnostic Accuracy (2003) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Clin Chem* 49:7–18
- Cardon LR, Palmer LJ (2003) Population stratification and spurious allelic association. *Lancet* 361:598–604
- Colhoun HM, McKeigue PM, Davey Smith G (2003) Problems of reporting genetic associations with complex outcomes. *Lancet* 361:865–872
- Donahue MP, Allen AS (2005) Genetic association studies in cardiology. *Am Heart J* 149:964–970
- Evangelou E, Maraganore DM, Ioannidis JP (2007) Meta-analysis in genome-wide association datasets: strategies and application in Parkinson disease. *PLoS ONE* 2:e196
- Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6:95–108
- Hirschhorn JN, Lohmueller K, Byrne K, Hirschhorn K (2002) A comprehensive review of genetic association studies. *Genet Med* 4:45–61
- Kitsios G, Zintzaras E (2007) Genetic variation associated with ischemic heart failure: a HuGE review and meta-analysis. *Am J Epidemiol* 166:619–633
- Lacroix Z (2002) Biological data integration: wrapping data and tools. *IEEE Trans Inf Technol Biomed* 6:123–128
- Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC (1992) Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 327:248–254
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 354:1896–1900
- Munafò MR, Flint J (2004) Meta-analysis of genetic association studies. *Trends Genet* 20:439–444
- NCI-NHGRI Working Group on Replication in Association Studies (2007) Replicating genotype–phenotype associations. *Nature* 447:655–660
- Neale BM, Sham PC (2004) The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 75:353–362
- Teufel A, Krupp M, Weinmann A, Galle PR (2006) Current bioinformatics tools in genomic biomedical research. *Int J Mol Med* 17:967–973
- Thomas DC, Witte JS (2002) Point: population stratification: a problem for case–control studies of candidate–gene associations? *Cancer Epidemiol Biomarkers Prev* 11:505–512
- Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. *Nat Rev Genet* 6:109–118
- Zdoukopoulos N, Zintzaras E (2007) Genetic markers in placental abruption: a HuGE review and meta-analysis. *Epidemiology* (in press)
- Zintzaras E (2006a) C677T and A1298C methylenetetrahydrofolate reductase gene polymorphisms in schizophrenia, bipolar disorder and depression: a meta-analysis of genetic association studies. *Psychiatr Genet* 16:105–115
- Zintzaras E (2006b) Methylenetetrahydrofolate reductase gene and susceptibility to breast cancer: a meta-analysis. *Clin Genet* 69:327–336
- Zintzaras E (2007) Brain-derived neurotrophic factor gene polymorphisms and schizophrenia: a meta-analysis. *Psychiatr Genet* 17:69–75
- Zintzaras E, Hadjigeorgiou GM (2004) Association of paraoxonase 1 gene polymorphisms with risk of Parkinson’s disease: a meta-analysis. *J Hum Genet* 49:474–481
- Zintzaras E, Hadjigeorgiou GM (2005) The role of G196A polymorphism in the brain-derived neurotrophic factor gene in the cause of Parkinson’s disease: a meta-analysis. *J Hum Genet* 50:560–566
- Zintzaras E, Ioannidis JP (2005) Heterogeneity testing in meta-analysis of genome searches. *Genet Epidemiol* 28:123–137
- Zintzaras E, Ioannidis JPA (2007) Meta-analysis for ranked discovery datasets: theoretical framework and empirical demonstration for microarrays. *Comput Biol Chem*. doi: 10.1016/j.compbiolchem.2007.09.003
- Zintzaras E, Lau J (2007) Synthesizing genetic association studies: methodological and statistical considerations. *J Clin Epidemiol* (in press)
- Zintzaras E, Stefanidis I (2005) Association between the GLUT1 gene polymorphism and the risk of diabetic nephropathy: a meta-analysis. *J Hum Genet* 50:84–91
- Zintzaras E, Sakelaridis N (2007) Is 472G/A catechol-O-methyltransferase gene polymorphism related to panic disorder? *Psychiatr Genet* 17:267–273
- Zintzaras E, Kitsios G, Stefanidis I (2006a) Endothelial NO synthase gene polymorphisms and hypertension: a meta-analysis. *Hypertension* 48:700–710
- Zintzaras E, Stefanidis I, Santos M, Vidal F (2006b) Do alcohol-metabolizing enzyme gene polymorphisms increase the risk of alcoholism and alcoholic liver disease? *Hepatology* 43:352–361
- Zintzaras E, Uhlig K, Koukoulis GN, Papanasiou AA, Stefanidis I (2007) Methylenetetrahydrofolate reductase gene polymorphism as a risk factor for diabetic nephropathy: a meta-analysis. *J Hum Genet* 52:881–890