Original Article

Association Analysis of CMYA5 rs4704591 Polymorphism with Rheumatoid Arthritis in Caucasians

Assoziationsanalyse des CMYA5 rs4704591-Polymorphismus mit rheumatoider Arthritis bei Kaukasiern

Authors
Mansour Zamanpoor[1, 2, 4], Natsha Anne Austin[1], Hamid Ghaedi[2], Nadine H. Nograles[3], Angela E. Brown[4], Andrew D. Wilson[4, 5], Tony R. Merriman[1], Ian M. Morison[6], Mir Davood Omrani[2]

Affiliations
1 Biochemistry, University of Otago, Dunedin, New Zealand
2 Medical Genetics, Shahid Beheshti University of Medical Sciences, Tehran, Iran (the Islamic Republic of)
3 Biomedical Sciences, Newcastle University Medicine Malaysia, Nusajaya, Malaysia
4 Wellington Regional Genetics Laboratory, Wellington Regional Hospital, Wellington, New Zealand
5 Department of Pathology and Molecular Medicine, University of Otago, Wellington, New Zealand
6 Pathology, University of Otago, Dunedin, New Zealand

Correspondence
Mansour Zamanpoor
Biochemistry
University of Otago
Dunedin 9016
New Zealand
Tel.: +6434797863, Fax: +6434797866
mansour.zamanpoor@gmail.com

ABSTRACT

Introduction Single nucleotide polymorphisms (SNPs) in the Cardiomyopathy-Associated Protein 5 (CMYA5) gene have been associated with rheumatoid arthritis (RA) in genome-wide association studies. In this study, we aimed to replicate the association between CMYA5 gene polymorphisms and RA in independent Caucasian case-control cohorts and perform a meta-analysis to investigate the association of CMYA5 gene polymorphisms with RA in Caucasian populations.

Methods We analysed 2731 RA cases and 1783 healthy controls in 4 independent Caucasian sample sets. rs4704591 in CMYA5 gene was genotyped using the TaqMan SNP genotyping assay. Meta-analysis was conducted across Caucasian cohorts.

Results Our analysis showed no evidence for association of rs4704591 with RA in the replication sample sets (p = 0.941, OR = 0.997). Meta analysis showed a weak association between the minor allele of the CMYA5 rs4704591 variant (C) and RA in the combined RA cohorts (p = 0.023, OR = 0.938) using the logistic regression model in the matched case-control study.

Conclusion Our study failed to replicate the association of the CMYA5 rs4704591 variant with RA and therefore, we cannot confirm the association between CMYA5 gene polymorphisms and RA in Caucasian population. However, further investigation might help to unravel the association of CMYA5 gene variants with RA.

ZUSAMMENFASSUNG

Einleitung Einzel nukleotid polymorphismen (SNPs) im Cardiomyopathy-Associated Protein 5 (CMYA5)-Gen wurden in genomweiten Assoziationsstudien mit rheumatoider Arthritis (RA) in Verbindung gebracht. In dieser Studie wollten wir die Assoziation zwischen CMYA5-Genpolymorphismen und RA in unabhängigen kaukasischen Fall-Kontroll-Kohorten replizieren und eine Metaanalyse durchführen, um die Assoziation von CMYA5-Genpolymorphismen mit RA in kaukasischen Populationen zu untersuchen.

Methoden Wir analysierten 2731 RA-Fälle und 1783 gesunde Kontrollen in vier unabhängigen kaukasischen Probensätzen. rs4704591 im CMYA5-Gen wurden unter Verwendung des Taq-

Man SNP-Genotypisierungsassays genotypisiert. Die Metaanalyse wurde über kaukasische Kohorten hinweg durchgeführt.

Ergebnisse Unsere Analyse ergab keine Hinweise auf eine Assoziation von rs4704591 mit RA in den Replikationsprobensätzen (P = 0,941, OR = 0,997). Die Metaanalyse zeigte eine schwache Assoziation zwischen dem kleinen Allel der CMYA5-Variante rs4704591 (C) und RA in den kombinierten RA-Kohorten

(P = 0,023, OR = 0,938) unter Verwendung des logistischen Regressionsmodells in der Matched-Case-Control-Studie.

Schlussfolgerung Unsere Studie war nicht erfolgreich darin, die Assoziation der CMYA5-Variante rs4704591 mit RA zu replizieren. Daher können wir die Assoziation zwischen CMYA5-Genpolymorphismen und RA in der kaukasischen Bevölkerung nicht bestätigen.

Introduction

Rheumatoid arthritis (RA) is a multigenic auto-immune disorder in which both genetic and environmental factors contribute to its pathophysiology [1]. More than 100 genetic loci have been associated with the RA as a result of genome-wide association studies. The genetics of RA is complex and each of associated risk loci contribute small but significant effect in the development of RA as a systemic autoimmune disease [1].

The Cardiomyopathy Associated Protein 5 (*CMYA5)* gene or Myospryn has been linked to cardiomyopathy and is highly expressed in skeletal muscle and heart whereas moderately expressed in brain [2, 3]. *CMYA5* was reported as a susceptibility gene for RA [4–6]. *CMYA5* interacts with Dystrobrevin Binding Protein 1 (*DTNBP1*) gene [7, 8]. Functional studies correlated *CMYA5* with protein kinase A in the regulation of the cyclic adenosine monophosphate (cAMP) signaling pathway and also with the *DTNBP1* gene in the biogenesis of lysosome-related organelles complex 1 (BLOC-1), processes [2, 9, 10].

In this study, we aimed to replicate the association between *CMYA5* and RA by testing SNPs within *CMYA5* locus for association with RA in independent Caucasian sample sets.

Methods

Study participants

In order to conduct a case-control association study, RA patients were recruited and compared with healthy normal individuals as controls in this study. Our study consisted of European Caucasian subjects derived from four separate RA case-control sample sets consisted of Australia and New Zealand (ANZ), Oxford, UK and Dutch cohorts in addition to the Wellcome Trust Case Control Consortium (WTCCC) data (publically available). RA patients were diagnosed by rheumatologists and were confirmed to met the ACR classification system criteria for the diagnosis of RA [11]. Control participants have had no history of autoimmune disease and being of European Caucasian ancestry (self-reported) in all of the sample sets. Written informed consent was obtained from all study subjects.

Ethics approval and consent to participate

Written informed consent was obtained from all study subjects. The study was approved by the multi-region Ethics Committee and the Lower South Ethics Committee of New Zealand, the Research

and Ethics Committee of the Repatriation General Hospital of Australia, the Lewisham Hospital and Guy's and St Thomas' Hospitals local research ethics committees and the European Collection of Authenticated Cell Culture research ethics committee, the Oxford Research Ethics Committee and, the ethical committee of the Radboud University Nijmegen Medical Centre.

DNA Extraction

Genomic DNA was extracted from participant blood samples by using the robust and standard method. The quality and quantity of the extracted DNA samples were obtained using the NanoDrop spectrophotometery. All samples were stored in 96-well 0.8ml deep well boxes. The storage plate was used to replenish the 6 ng/mL working stock plates to minimise cross-contamination of the stock DNA. Each sample set contained a series of plates for cases and controls.

SNPs Selection

The Wellcome Trust Case Control Consortium (WTCCC) study was utilized as the initial dataset for investigation of single nucleotide polymorphisms (SNPs) in this study. RA data was derived from the WTCCC dataset. *CMYA5* were selected based on the association in the Wellcome Trust Case Control Consortium (WTCCC) data for RA [4]. Genotype information was obtained using PLINK software for data analysis. The linkage disequilibrium (LD) for SNPs within the *CMYA5* gene was conducted to select the candidate SNP using SNiPa v3.3 [12] (▶ **Fig. 1**). There are some SNPs in the *CMYA5* region that have been associated with schizophrenia and bipolar disorder [13]. Therefore, only rs4704591 was selected for genotyping in Caucasian case-control sample sets for replication. The direction of association of the minor allele (protective or susceptible) was compared based on the OR in each dataset.

Genotyping

We genotyped the rs4704591 (assay ID C__27865739_10) using TaqMan SNP genotyping assay technology (Applied Biosystems, Foster City, USA) and a LightCycler 480 Real-Time Polymerase Chain Reaction System (Roche, Indianapolis, IN, USA).

Quality control of genotyping data was routinely performed to warrant the correct genotype calling for samples and minimize potential false findings. Known negative and positive controls were included in each run. To confirm the correct orientation of plates during genotyping, eight samples from each of the sample boxes were aliquoted into a separate 96-well box. This box was analysed
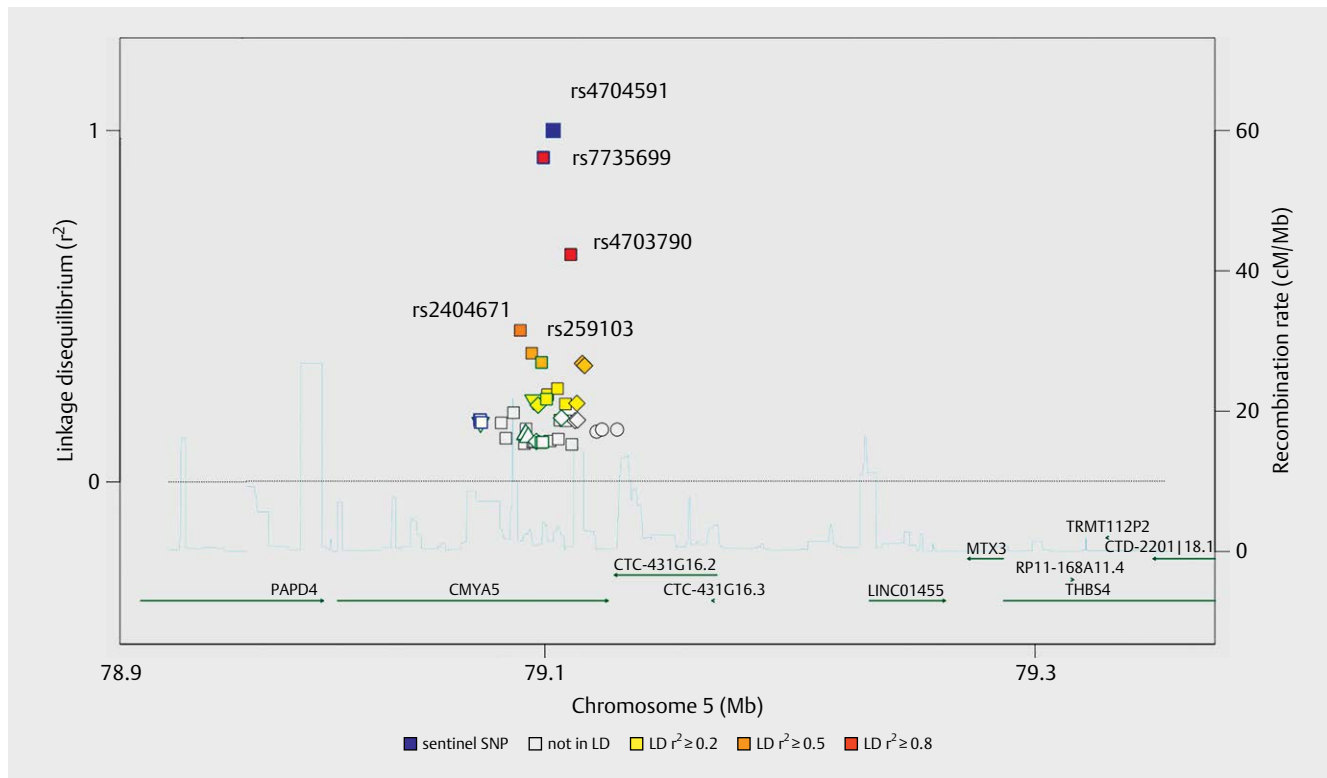
▶ **Fig. 1** Linkage disequilibrium plot shows the correlation between the rs4704591 variant and its surrounding variants generated by SNiPA using 1000 Genomes, Phase 3 v5 variant set for the European population (GRCh37.p13, Ensemble 87). The y-axis shows the correlation coefficient (r2), the x-axis shows chromosomal position of each SNP.

in the same way and the data were compared to the previous genotypes. If the results did not match, the whole plate was checked and the sample genotyping was repeated where necessary.

Statistical analysis

The significance of differences in the minor allele frequencies between cases and controls, odds ratios (OR) and departures from Hardy-Weinberg equilibrium, were calculated using the SHEsis and PLINK software packages [14, 15]. Genotype based ORs and 95 % confidence intervals (CI) were estimated using the logistic regression model in this matched case-control study. We used logistic regression model as it assumes odds for a disease risk increases exponentially as a number of minor allele increases and fits the multiplicative model [16]. Meta-analysis was performed to combine results from several studies or sample sets to increase the study power to detect small to moderate effects of SNPs on disease. Meta-analysis was undertaken using the STATA 8.0 metan software package (http://www.stata.com) to obtain the combined Mantel-Haenszel (M-H) ORs, to calculate the allelic and Breslow-Day (B-D, test for heterogeneity) P values and to generate the odds-ratio meta-analysis plot.

Results

CMYA5 rs4704591 was selected for association analysis using the ANZ, UK, Oxford and Dutch case-control sample sets. No association was found with RA in the ANZ sample set (P = 0.732,

OR = 1.029), UK sample set (P = 0.899, OR = 0.982), Dutch sample set (P = 0.518, OR = 1.049) or Oxford sample set (P = 0.254, OR = 0.957) for this variant (▶ **Table 1**). Meta-analysis was performed to combine the replication sample sets in order to increase the power of the association analysis. Meta-analysis showed no evidence for association with RA in the combined genotyped (replication) sample sets (P = 0.941, OR = 0.997) (▶ **Table 1**). By adding the WTCCC RA data (P = 0.002, OR = 0.874) to the combined dataset, the combined allelic P-value showed a weak association with RA in the combined RA cohorts (P = 0.023, OR = 0.938). The OR for combined overall RA cohorts was conveying a weak protective effect against RA (P = 0.023, OR = 0.938) (▶ **Fig. 2**). This protective effect is driven by WTCCC with no support from the replication sample sets. The Breslow-Day test revealed no evidence of genetic heterogeneity between the RA combined sample sets for rs4704591 (combined genotyped RA P = 0.741, combined overall RA P = 0.175).

Discussion

The *CMYA5* gene, that codes for cardiomyopathy-associated protein 5 (myospryn), is moderately expressed in the brain [2, 3]. *CMYA5* gene has been associated with RA susceptibility gene [4, 5]. We found no evidence for association of rs4704591 in Caucasian RA case-control sample sets including ANZ, UK, Dutch or Oxford sample sets. Meta-analysis showed no evidence for association with RA in the combined replication sample sets.

▶ **Table 1** Case-control association analysis for CMYA5 rs4704591 in RA sample sets.

Cohort	Genotype			MAF	OR (CI)	Allelic P	B-D P
	G/G	G/A	A/A				
ANZ RA							
Case	280 (0.380)	354 (0.481)	102 (0.139)	0.379	1.029 (0.874–1.211)	0.732	
Control	213 (0.403)	238 (0.450)	78 (0.147)	0.372			
UK RA							
Case	133 (0.422)	138 (0.438)	44 (0.140)	0.359	0.982 (0.735–1.310)	0.899	
Control	60 (0.411)	66 (0.452)	20 (0.137)	0.363			
Dutch RA							
Case	396 (0.365)	504 (0.465)	184 (0.170)	0.402	1.049 (0.908–1.211)	0.518	
Control	228 (0.380)	275 (0.458)	97 (0.162)	0.391			
OX RA							
Case	228 (0.383)	281 (0.471)	87 (0.146)	0.382	0.957 (0.790–1.160)	0.254	
Control	184 (0.362)	236 (0.465)	88 (0.173)	0.406			
Combined							
Case	1037 (0.380)	1277 (0.467)	417 (0.153)	0.386	0.997 (0.913–1.088)	0.941	0.741
Control	685 (0.384)	815 (0.457)	283 (0.159)	0.387			
WTCCC RA							
Case	745 (0.402)	848 (0.458)	259 (0.140)	0.369	0.874 (0.802–0.952)	0.0019	
Control	1064 (0.363)	1384 (0.472)	482 (0.164)	0.401			
Combined RA							
Case	1782 (0.389)	2125 (0.464)	676 (0.147)	0.379	0.938 (0.882–0.998)	0.023	0.175
Control	1749 (0.371)	2199 (0.466)	765 (0.162)	0.396			

MAF: minor allele frequency, OR: odds ratio, CI: 95 % confidence interval, B-D P: Breslow-Day P value for heterogeneity.
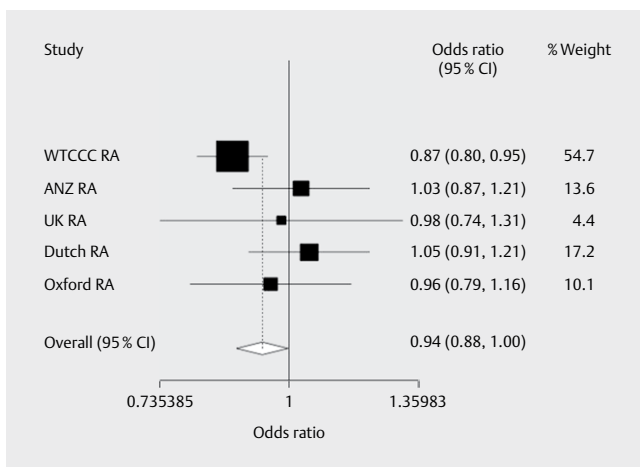
▶ **Fig. 2** Meta-analysis of CMYA5 rs4704591 in 5 independent RA Caucasian datasets with Mantel-Haenszel combined OR = 0.94, 95 %CI (0.88–1.00) with P = 0.042 using fixed-effect model.

Our study was not successful to replicate the association of the *CMYA5* rs4704591 variant with RA. This finding is consistent with the European RA GWAS dataset [17] that there was no association of rs4704591 with RA (*P* = 0.58, OR = 1.05). Although the number of recruited controls was lower than number of cases in our study, however, the Okada et al. (2014) study is the largest RA GWAS that analysed 103,638 samples that had sufficient statistical power to detect true associations. The lack of association in the current study as well as that from Okada et al. (2014) indicate that *CMYA5* rs4704591 variant is not genetically associated with RA.

In our study, the association of *CMYA5* rs4704591 with RA was only detected in the discovery cohort (WTCCC) and it was not associated in any other cohorts . The inconsistency in detection of associations in this study might be explained by population heterogeneity between the discovery and replication cohorts and also the differences in distribution of genetic and environmental risk factors in the studied cohorts. Manchia et al. (2013) [18] reported that the presence of phenotypic and genetic heterogeneity would reduce the statistical power and also influence estimation of the effect sizes of the observed association in GWAS. A lack of power to detect small effect sizes and population heterogeneity are 2 major issues in association study of RA as a genetically complex disorder. Population stratification due to systematic ancestry differences between sample sets (population structure) or unmatched sample sets can confound genetic association findings. Given the diverse geographic distribution of the sample sets gene-environ-

ment interactions might be responsible for the inconsistence in results between sample sets. Wei et al. (2017) [19] showed evidence for the presence of a geographically-dependent gene-environment interaction in distinct geographic locations by genotypic variability-based GWAS meta-analysis of multiple RA cohorts represented geographical locations.

One limitation of this study is that selection criteria for the candidate loci were based on the association in the publicly-available GWAS case-control studies. It is possible that some associated loci were not captured in those studies since only about 80 % of the common SNPs in the Caucasian genome were genotyped. Another limitation of this study is that SNPs with the association P-values less than 0.05 in WTCCC RA dataset as primary discovery dataset were of interest for replication as candidate SNPs. However, replication of association for a candidate SNP was not successful in this study that might be explained by the fact that selected candidate SNP was of weaker associations (P = 0.002, OR = 0.874). Although it is arguable whether the selected SNP reflect a true effect as they reached only nominal statistical significance (P-values less than 0.05). On the other hand, RA has 2 genetically distinct subsets (anti-citrullinated protein antibodies (ACPA) positive and negative) each with some specific set of associated SNPs while they also share some susceptibility loci [1]. These phenotypic or genetic subtypes challenge the assumption of being a single biologically valid category. In addition to the relatively small size of this study, one limitation of this project is that RA cases were not subdivided into ACPA negative and positive cohorts to be studied separately. Therefore it is desirable to investigate the association of candidate SNP(s) in more specifically subdivided study cohorts in future.

The CMYA5 protein has been reported to interact with Dystrobrevin Binding Protein 1 (DTNBP1) that has been confirmed to be linked to schizophrenia by several studies [7, 20–23]. *CMYA5* is involved in protein kinase A (PKA) signaling and vesicular trafficking [3]. Myospryn, the product of the *CMYA5* gene, interacts with the product of the *DTNBP1* gene or dysbindin [9]. DTNBP1 has been suggested to promote neuronal viability through PI3-kinase-Akt signaling [24]. There is well-established evidence for the role of the cAMP/PKA pathway in regulation of TNF-α and cytokine IL-10 in RA [25]. There is also evidence for interaction between the myospryn and the cAMP signaling pathway [9].

In conclusion, the findings of this study did not support the implication of the *CMYA5* gene in the development of RA. However, further investigations are required to investigate the variants within *CMYA5* gene with regulation of the *DTNBP1* and *PKA* genes. Functional studies would be helpful to unravel the molecular basis of the *CMYA5* gene's relationship in these complex pathways.

Data Accessibility Statement

Data is available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions. The WTCCC data was derived from https://www.wtccc.org.uk/info/access_to_data_samples.html.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Zamanpoor M. The genetic pathogenesis, diagnosis and therapeutic insight of rheumatoid arthritis. Clinical genetics 2019; 95: 547–557. doi:10.1111/cge.13498

[2] Chen X, Lee G, Maher B et al. GWA study data mining and independent replication identify cardiomyopathy-associated 5 (CMYA5) as a risk gene for schizophrenia. Molecular psychiatry 2011; 16: 1117–1129

[3] Sarparanta J. Biology of myospryn: what's known? Journal of muscle research and cell motility 2008; 29: 177–180. doi:10.1007/s10974-008-9165-6

[4] Burton PR, Clayton DG, Cardon LR et al. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661–678

[5] Julia A, Ballina J, Canete JD et al. Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. Arthritis and rheumatism 2008; 58: 2275–2286. doi:10.1002/art.23623

[6] Tsoupri E, Capetanaki Y. Muyospryn: a multifunctional desmin-associated protein. Histochemistry and cell biology 2013; 140: 55–63. doi:10.1007/s00418-013-1103-z

[7] Talbot K, Cho D-S, Ong W-Y et al. Dysbindin-1 is a synaptic and microtubular protein that binds brain snapin. Human molecular genetics 2006; 15: 3041–3054

[8] Benson MA, Tinsley CL, Blake DJ. Myospryn is a novel binding partner for dysbindin in muscle. J Biol Chem 2004; 279: 10450–10458. doi:10.1074/jbc.M312664200

[9] Zhang R, Zhang H, Li M et al. Genetic analysis of common variants in the CMYA5 (cardiomyopathy-associated 5) gene with schizophrenia. Prog Neuropsychopharmacol Biol Psychiatry 2013; 46: 64–69. doi:10.1016/j.pnpbp.2013.05.015

[10] Kouloumenta A, Mavroidis M, Capetanaki Y. Proper perinuclear localization of the TRIM-like protein myospryn requires its binding partner desmin. The Journal of biological chemistry 2007; 282: 35211–35221. doi:10.1074/jbc.M704733200

[11] Arnett F, Edworthy S, Bloch D et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis & Rheumatism 1988; 31: 315–324

[12] Arnold M, Raffler J, Pfeufer A et al. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015; 31: 1334–1336. doi:10.1093/bioinformatics/btu779

[13] Wang KS, Liu XF, Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophrenia research 2010; 124: 192–199. doi:10.1016/j.schres.2010.09.002

[14] Shi Y, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. Cell Research 2005; 15: 97–98

[15] Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics 2007; 81: 559–575. doi:10.1086/519795

[16] Horita N, Kaneko T. Genetic model selection for a case-control study and a meta-analysis. Meta gene 2015; 5: 1–8. doi:10.1016/j.mgene.2015.04.003

[17] Okada Y, Wu D, Trynka G et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature 2014; 506: 376–381. doi:10.1038/nature12873

[18] Manchia M, Cullis J, Turecki G et al. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. PLoS One 2013; 8: e76295. doi:10.1371/journal.pone.0076295

[19] Wei WH, Viatte S, Merriman TR et al. Genotypic variability based association identifies novel non-additive loci DHCR7 and IRF4 in sero-negative rheumatoid arthritis. Scientific reports 2017; 7: 5261. doi:10.1038/s41598-017-05447-1

[20] Williams NM, O'Donovan MC, Owen MJ. Is the dysbindin gene (DTNBP1) a susceptibility gene for schizophrenia? Schizophrenia bulletin 2005; 31: 800–805

[21] Owen MJ, Williams NM, O'Donovan MC. Dysbindin-1 and schizophrenia: from genetics to neuropathology. J Clin Invest 2004; 113: 1255–1257. doi:10.1172/JCI21470

[22] Straub RE, Jiang Y, MacLean CJ et al. Genetic variation in the 6p22.3 gene DTNBP1, the human ortholog of the mouse dysbindin gene, is associated with schizophrenia. Am J Hum Genet 2002; 71: 337–348. doi:10.1086/341750

[23] Zamanpoor M. Schizophrenia in a genomic era: a review from the pathogenesis, genetic and environmental etiology to diagnosis and treatment insights. Psychiatric genetics 2020; 30: 1–39. doi:10.1097/YPG.0000000000000245

[24] Numakawa T, Yagasaki Y, Ishimoto T et al. Evidence of novel neuronal functions of dysbindin, a susceptibility gene for schizophrenia. Human Molecular Genetics 2004; 13: 2699

[25] Foey AD, Field S, Ahmed S et al. Impact of VIP and cAMP on the regulation of TNF-alpha and IL-10 production: implications for rheumatoid arthritis. Arthritis Res Ther 2003; 5: R317–R328. doi:10.1186/ar99