# Autosome-wide Copy Number Variation Association Analysis for Rheumatoid Arthritis Using the WTCCC High-density SNP Genotype Data

MOHAMMED UDDIN, MITCH STURGE, PROTON RAHMAN, and MICHAEL O. WOODS

***ABSTRACT.*** ***Objective.*** Rheumatoid arthritis (RA) is a complex autoimmune rheumatic disease that is strongly influenced by genetic factors. Numerous genes are convincingly associated with RA, including genes in tumor necrosis factor signaling (TNF) and the nuclear factor-κB pathway. To date, except for genes within the HLA region, no data exist regarding potential copy number variations (CNV) involving RA-associated genes. We set out to identify genes affected by CNV that are associated with RA at a genome-wide level.

***Methods.*** Data from the Wellcome Trust Case Control Consortium (WTCCC) were used in our analyses. The initial WTCCC cohort genotyped 3004 controls and 1999 RA cases using the GeneChip 500k Mapping Array Set. We performed a comparative intensity analysis using the PennCNV algorithm, which uses a hidden Markov model to detect CNV. A total of 2271 controls and 1572 RA samples passed quality control criteria and were included for association analysis. Association analysis was performed in 2 phases: (1) to identify CNV that are < 1 Mb with a population frequency < 5%; and (2) to identify large CNV that are > 1 Mb. Fishers' exact test was performed to quantify significance of the CNV.

***Results.*** We observed that the genome-wide CNV burden is 2-fold higher in patients with RA compared with controls. We identified 11 rare copy number variable regions with < 5% frequency that had an association with RA that reached a $p < 1 \times 10^{-4}$. These include *TNFAIP3* and *TNIP1*, which has been implicated in association studies for RA, systemic lupus erythematosus, and psoriasis. We identified CNV involving *IRF1*, which functions as a transcription activator of genes induced by interferons; *ALOX5AP* and *LCP2*, involved in inflammatory mediation; *B2M*, an MHC-class I associated gene; and *PRKCH*, a gene involved in T cell signaling pathways. A 57 kb deletion with 1% frequency in RA cases at 7p21.3 was also observed. Six of these loci overlap with CNV catalogued in the Database of Genomic Variants.

***Conclusion.*** This is the first study to identify non-HLA RA-associated CNV using genome-wide analyses. Validation and functional significance of these deletions/duplications in RA and other autoimmune diseases need to be further investigated. (J Rheumatol First Release March 1 2011; doi:10.3899/jrheum.100758)

*Key Indexing Terms:*
COPY NUMBER VARIATION ASSOCIATION ANALYSIS RHEUMATOID ARTHRITIS

Rheumatoid arthritis (RA) is the prototypic seropositive inflammatory arthritis. The heritability of RA is between 50% and 60%, with HLA-DRB1 accounting for 30% of the genetic risk of developing RA[1]. The RA-associated genes can be broadly characterized as being involved in T cell activation (*PTPN22*, *STAT4*, and *CTLA4*) and the nuclear factor-κB (NF-κB) signaling pathway (*CD40*, *TRAF1*, *TNFSF14*, and *TNFAIP3*)[2].

Recently, structural variation of DNA, namely copy number variation (CNV), has been recognized as important for both normal genomic variability and in disease susceptibility. A CNV is a common type of genomic variation ranging in size from 1 kilobase (kb) to several megabases (Mb) and they cover at least 25% of the human genome[3]. Among different individuals, a single CNV can have different forms (i.e., duplication, deletion). The variable region that contains these different CNV is known as the copy number variation region (CNVR). CNV have the potential to disrupt genes and are associated with many complex diseases, including psoriasis, systemic lupus erythematosus (SLE), Crohn's disease, autism, and osteoporosis[4,5,6,7,8,9]. Thus, it is important to assess the influence of CNV to obtain a thorough understanding of genetic susceptibility in complex disease.

*From the Discipline of Genetics, Faculty of Medicine, and the Discipline of Rheumatology, Department of Medicine, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.*

*M. Uddin, MSc, Discipline of Genetics, Faculty of Medicine; M. Sturge, BSc; P. Rahman, MD, Professor of Medicine, Discipline of Rheumatology, Department of Medicine; M.O. Woods, PhD, Discipline of Genetics, Faculty of Medicine.*

*P. Rahman and M.O. Woods are co-senior authors and contributed equally to this report.*

*Address correspondence to Dr. P. Rahman, St. Clare's Mercy Hospital, 1 South, 154 LeMarchant Road, St. John's, Newfoundland A1C 5B8. E-mail:prahman@mun.ca*

*Accepted for publication December 21, 2010.*

Previous CNV studies in RA have noted association with chemokine ligand 3-like 1 (*CCL3L1*) and Fc fragment of IgG, low affinity IIIb, receptor (*FCGR3B*). *CCL3L1* binds to several proinflammatory cytokine receptors, including chemokine receptor 5 (*CCR5*). In a recent New Zealand study, a copy number higher than 2 at the *CCL3L1* locus was a risk factor for RA, but this was not replicated in a smaller UK cohort[10]. Two independent Dutch studies have reported that the 1q23 region, containing *FCGR3B*, has been associated with CNV that influence susceptibility to RA[11,12]. However, no such association was noted in the UK cohort[1], nor in the recent WTCCC CNV analysis[13].

We performed a comparative intensity analysis to detect CNV using the WTCCC single-nucleotide polymorphism (SNP) genotype array data. In addition, we performed a CNV genome-wide association analysis to identify disease-susceptible loci in the WTCCC RA cohort.

## MATERIALS AND METHODS

*Cohort*. The initial cohort data consist of 3004 shared controls and 1999 individuals with RA obtained from the WTCCC data center. The controls were taken from 2 groups: 1504 samples from the 1958 British Birth Cohort (58C) and 1500 additional controls recruited from the UK Blood Services (UKBS). A set of 500,568 SNP were genotyped using Affymetrix GeneChip 500K Mapping Array Set (Affymetrix Inc., Santa Clara, CA, USA)[14].

*Quality control*. We used PennCNV software to detect high-resolution copy number variation. PennCNV uses a hidden Markov model to detect kilobase resolution detection of CNV with low false positives[15]. Prior to the analysis, the sample inclusion/exclusion criteria were set based on WTCCC SNP analysis. All subjects that passed quality control criteria of the WTCCC original SNP association analysis were included for CNV analysis. Samples were excluded after a check for contamination, false identity, relatedness, and non-Caucasian ancestry. Affymetrix Power Tool, a BRLMM (Bayesian robust linear model with Mahalanobis distance classifier) algorithm, was used to make genotype calls for NSP and STY arrays separately[16]. The first array uses the NSP I restriction enzyme to genotype 250K SNP while the second array uses STY I restriction enzyme to genotype 250K SNP. To reduce signal-to-noise ratio we used the WTCCC data to model parameters for canonical genotype clustering, instead of using default canonical genotype clustering information provided by PennCNV, which is used to calculate log R ratio (LRR) and B allele frequency (BAF) values. After separate calculations of LRR and BAF, the 2 array sets, NSP and STY, were combined to make CNV detection calls in the autosomes. Many of the genomic samples can have below-optimum genomic wave quality control values; hence, in our sample inclusion criteria we used the default parameters of PennCNV to restrict case or control samples — LRR standard deviation < 0.25, BAF drift > 0.01, and wave factor > 0.05 or < –0.05. A total of 2271 controls (1123 58C and 1148 UKBS) and 1572 RA samples passed all the quality control criteria and were included for association analysis.

After exclusion of samples, we performed quality control on CNV calls. PennCNV tends to show false-positive CNV calls on centromeric and telomeric regions. Due to the complex nature of hemizygosity in sex chromosomes, we excluded all CNV detected from sex chromosomes. In addition, human immunoglobulin coding regions showed false-positive CNV calls using PennCNV, therefore we excluded CNV that overlapped at 50% or more of its length with the following immunoglobulin regions: chr2: 88.9–89.4 Mb, chr14: 21.1–22.0 Mb, chr15: 17.0–21.0 Mb, chr16: 31.8–36.8 Mb, and chr22: 20.7–21.5 Mb. Also, large CNV that overlap with centromeres and telomeres were excluded from the following regions: chr1: 12.0–14.1 Mb, chr10: 46.3–47.1 Mb, chr14: 19.2–19.4 Mb, chr15: 18.4–20.0 Mb, and chr19: 24.2–32.7 Mb (build 35). The boundaries for immunoglobulin, centromere, and telomere regions were obtained from a previous study that used PennCNV[17].

*Association analysis*. Association analysis was performed on the detected CNV that were < 1 Mb, and Fisher's exact test was performed to quantify the significance of the CNV with exact breakpoint match or within a CNVR. Association analysis was considered to be suggestively significant if (1) $p < 1.0 \times 10^{-4}$; (2) there were at least 15 SNP within the CNV; and (3) the population frequency was < 5%. To avoid possible plate/batch effects, we manually checked the significant CNV samples and ignored any association if a CNV was identified in > 50% of samples from the same plate/batch. We also excluded CNV with strong p values if the LRR standard deviation for that CNV call was between 0.20 and 0.25 to avoid borderline CNV calls. A second association analysis was performed on large CNV that were > 1 Mb. Large CNV are not frequent, but any CNV in fewer than 3 samples were ignored. Fisher's exact test was performed to quantify the significance of the CNV. The cutoff value for significance at $p < 1 \times 10^{-4}$ was arbitrary, since we are not aware of the total number of CNV in the human genome. However, this level of significance was also used in the recent CNV analysis by WTCCC that introduced the population frequency threshold for rare CNV of < 5%[13].

The Database of Genomic Variants (DGV) includes CNV that have been annotated. The detected CNV in our case and control samples were compared with the DGV (Build 35).

## RESULTS

*Genome-wide CNV distribution*. High-resolution SNP obtained from the WTCCC data center had an average call rate of 99.63%, demonstrating high quality intensity data. The log R spread for the controls and cases showed that ~90% of the samples were within mean standard deviation of 0.15 to 0.22 in controls and 0.13 to 0.20 for the RA cases. After quality control, we identified 5927 CNV in RA cases and 4333 in WTCCC controls that were at least 1 kb in length. There was almost a 2-fold increase in the total CNV burden of 1 kb or more in size among RA patients compared to controls. The excessive burden of CNV compared to controls has also been noted in patients with schizophrenia and autism[18,19]. However, without experimental validation of CNV it is difficult to exclude the possibility that this may be due to an experimental artefact. Most of these CNV were found in fewer than 3 samples and thus were omitted from the association analysis. Fifty-four percent of the CNV detected in this study overlapped with CNV in the DGV.

*Association analysis*. Our association analysis revealed significance for the HLA region chr6 30583394–30889981 ($p < 5 \times 10^{-7}$), which overlaps with DGV ID 3600. There were 11 CNV that achieved significance ($p < 1 \times 10^{-4}$), shown in Table 1. The association analysis showed mostly one-sided CNV occurrences (in cases more than in controls). The disease-associated CNV involved numerous genes, including the interferon regulatory factor 1 (*IRF1*) gene; the tumor necrosis factor (TNF)-induced gene *TNFAIP3* and its interacting protein 1 gene *TNIP1*; 2 autoimmune related genes, lymphocyte cytosolic protein 2 (*LCP2*) and beta-2-microglobulin (*B2M*); 2 genes, protein kinase C-eta (*PRKCH*) and phosphatidylinositol-3,4,5-trisphosphate-dependent rac

*Table 1*. The association results of the most significant rate copy number variation regions (CNVR). The genes are contained within, or overlap, with CNV breakpoints. All CNV positions are given in NCBI Build 35 coordinates.

| Gene | Chromosome | CNVR Start-End | Deletion | | Duplication | | Fisher Exact p | DGV ID* |
|---|---|---|---|---|---|---|---|---|
| | | | Case | Control | Case | Control | | |
| ALOX5AP | 13q12.3 | 30136207-30264780 | 0 | 0 | 35 | 2 | $5.06 \times 10^{-12}$ | |
| SRGN | 10q22.1 | 70423581-70542187 | 6 | 1 | 21 | 1 | $4.35 \times 10^{-9}$ | 48614 |
| Gene Desert (nearby gene THSD7A) | 7p21.3 | 11696007-11753538 | 20 | 0 | 0 | 0 | $1.60 \times 10^{-8}$ | 3665 |
| LCP2 | 5q35.1 | 169605980-169668498 | 0 | 0 | 24 | 2 | $5.26 \times 10^{-8}$ | |
| PREX1 | 20q13.13 | 46676571-46882578 | 0 | 0 | 18 | 0 | $9.70 \times 10^{-8}$ | 12195 |
| B2M | 15q21.1 | 42606873-42824209 | 0 | 0 | 26 | 4 | $2.71 \times 10^{-7}$ | 3961 |
| LITAF | 16p13.13 | 11506474-11695215 | 1 | 0 | 15 | 0 | $5.87 \times 10^{-7}$ | |
| PRKCH | 14q23.1 | 60986674-61130641 | 0 | 0 | 20 | 2 | $1.39 \times 10^{-6}$ | 49385 |
| IRF1 | 5q31.1 | 131735192-131863294 | 2 | 0 | 18 | 2 | $1.39 \times 10^{-6}$ | |
| TNFAIP3 | 6q23.3 | 138064233-138239517 | 0 | 0 | 11 | 0 | $5.25 \times 10^{-5}$ | |
| TNIP1 | 5q23.3 | 150379418-150470998 | 2 | 0 | 18 | 2 | $7.59 \times 10^{-5}$ | 6455 |

* From the Database of Genome Variants.

exchange factor 1 (*PREX1*), involved in cell-signaling pathways; 2 known inflammatory mediator genes, arachidonate 5-lipoxygenase-activating protein (*ALOX5AP*) and lipopolysaccharide-induced TNF factor (*LITAF*); and the cell proliferation gene serglycin (*SRGN*). One CNVR contained deletions located in an intergenic region on chromosome 7p21.3, with the nearest gene, thrombospondin type-1 domain-containing protein 7A (*THSD7A*), located approximately 50 kb away from the deletion.

The association analysis of large CNV with RA (Table 2) revealed a duplication on chromosome 16p11.2 that was 1.46 Mb ($p < 0.001$). Duplications and deletions on 16p11.2 have previously been reported and validated in 2 independent studies[17,20]. However, our data confirmation revealed there was very low probe intensity in this region and therefore this result should be interpreted with caution.

## DISCUSSION

Copy number variation may play a significant role in genetic susceptibility and expression of complex autoimmune disease. An example of this is the role of CNV in beta-defensins in Crohn's disease and psoriasis. Beta-defensins are small antimicrobial peptides that play an important role in the innate immune system. A cluster of beta-defensin genes contain an extensive array of CNV in humans. It has been demonstrated that individuals with Crohn's disease have significantly lower beta-defensin copy number than that in controls. It is hypothesized that this results in the reduction of the antimicrobial barrier in the gut leading to Crohn's disease[21]. Meanwhile, individuals with psoriasis have significantly higher copy numbers than controls. The association analysis suggests 6 or more copies of beta-defensin is correlated with susceptibility to psoriasis. Due to the cytokine-like properties of the beta-defensin this could lead to an inappropriate inflammatory response generating psoriatic lesions after minor injury, infection, or other environmental triggers[4].

In our study assessing CNV in RA by analyzing a genome-wide association study, we first report the greater genome-wide burden of CNV in RA patients compared with controls. We found an approximately 2-fold increase in the number of CNV carried by an individual with RA. We also identified 11 rare CNVR, with < 5% frequency, that showed evidence of association with RA. Some of the loci identified encoded genes within the CNVR previously implicated in autoimmune diseases. The 2 most interesting candidates are *TNFAIP3* and *TNIP1*. The products of the *TNFAIP3* and *TNIP1* genes are A20 and ABIN1/Nafla, respectively. These proteins physically interact with each other to influence the ubiquitin-mediated destruction of IKKg (IkB kinase-g)/NEMO, which is an essential nexus of NF-κB signaling[22]. A20 also regulates the degradation of several other components of the TNF signaling pathway. Polymorphisms of *TNFAIP3* have been associated with RA, SLE, and psoriasis, although the specific variants differ among these autoimmune diseases[23,24,25,26]. *TNFAIP3* knockout mice develop multiorgan inflammation and arthritis[23]. As well, gene-disruptive CNV in classical Hodgkin's lymphoma are

*Table 2*. Association analysis of rare large copy number variations (CNV) that are greater than 1 Mb.

| CNV Region | CNVR Start-End | Case CNV | Control CNV | Fisher Exact p |
|---|---|---|---|---|
| 16p11.2 | 29053928 - 30514652 | 7 | 0 | 0.001 |
| 22q11.21 | 17275227 - 18602641 | 0 | 7 | 0.02 |

found in *TNFAIP3*[27]. This may be of potential interest as there is increased incidence of lymphoma among patients with moderate to severe RA.

Another candidate of potential interest is interferon regulatory factor 1. Type I interferons are a family of cytokines typically produced during viral infection but their multiple immunomodulatory effects are increasingly being recognized, including upregulation of innate immune receptors, polarization of T cells towards a TH1 phenotype, and activation of B cells. Recent studies provide strong evidence for an association between *IRF-5* gene variants and RA, particularly for patients with RA who are negative for anti-cyclic citrullinated peptide[28]. An insertion-deletion polymorphism in the *IRF-5* gene has also been shown to confer risk to inflammatory bowel disease[29].

The most significant CNV detected in our study overlaps with *ALOX5AP*, also known as *FLAP*. Expression of *ALOX5AP* is regulated by the binding of TNF-α to the promoter of *ALOX5AP*, thus modulating its gene expression[30]. In a functional study using collagen-induced arthritis in the DBA/1 mouse, it was shown that the severity of arthritis in FLAP-deficient mice was substantially reduced[31]. This implies the potential importance of *ALOX5AP* in arthritis. As well, *LITAF* showed upregulation in mice with glucose-6-phosphate isomerase (GPI)-induced arthritis[32].

CNV involving 2 other genes reported here, *PRKCH* and *PREX1*, have also been observed in acute myeloid leukemia and gastric cancer cell line, respectively[33,34]. *PRKCH* was reported to be associated with RA in a Japanese population[35]. *LCP2* and *B2M* are 2 additional autoimmune-related genes (Table 1). *LCP2* is important for normal T cell development and *B2M* is an MHC class I associated gene that is associated with spondylitis[36]. The CNVR at locus 7p21.3 is of particular interest. First, it contains deletions rather than duplications, and, in general, deletions are known to have higher phenotypic consequences (or penetrance)[5]. Second, this deletion occurs in at least 1% of RA cases and it is a validated deletion annotated in DGV (ID 3665). Finally, in a recent study, it was shown that CNV can have long-range effects (up to 1 Mb) on gene expression[37]; therefore, validation of the detected CNVR in the intergenic region near *THSD7A* on 7p21.3 is warranted.

Variation in copy number of *FCGR3B* was shown to influence susceptibility to RA in 2 independent studies in a Dutch population[35,36]. However, in our analysis, no evidence of association was observed due to poor coverage of the regions [*FCGR1A* (chr1: 1146567361–146577146) and *FCGR-A/2A/2B/2C/3A/3B* (chr1: 158,288,275–158,382,415)]. Limitations of our study include the issue of multiple testing and possibly under-estimating the number of common CNV due to inaccurate measurement of CNV breakpoints using SNP microarray data. Technological developments in sequencing will allow efficient identification of these breakpoints. Although we do not know the number of CNV in the human genome, we have done an exploratory analysis detecting disease-associated CNV that are $p < 1.0 \times 10^{-4}$. The false-positive error in the detection of duplications is significantly higher in Affymetrix arrays than in Illumina arrays[38]. Hence, we used a more restricted LRR standard deviation in our analysis than that used in the previously reported Illumina array CNV analysis using PennCNV[17]. Default parameters were used for the BAF drift and wave factor adjustment. As in a previous study, for the SNP intensity parameter, we have used the criteria where intensities of at least 15 SNP are considered for CNV < 1 Mb and intensities of 50 SNP are considered for CNV > 1 Mb.

In a recent CNV analysis carried out by the WTCCC, there were no associations of 3432 common CNV with common diseases, including RA. That analysis used the same cohort as in our study and a custom-designed Agilent Comparative Genomic Hybridization (CGH) chip for CNV detection, which is much more accurate than using SNP arrays for identifying CNV. However, even with CGH they identified a 15% false-positive rate when detecting duplications. Among the true-positive CNV, 50% had a frequency ≥ 5%, all were > 500 bp in length, and most were tagged by nearby SNP[13]. Similarly, our analyses also showed no strong associations of common CNV with RA other than those within the HLA region. One possible reason for this observation is that the probes for Affymetrix 500k are not uniformly distributed across the genome (median 2.5 kb space between SNP), hence the poor coverage affects the detection capacity and the frequency of CNV[38].

In summary, we are at the early stages of identification of CNV involved in inflammatory rheumatic diseases. Our study identified 11 rare CNVR associated with RA, and these now need to be verified by additional molecular studies. The functional significance of the validated CNV would then need to be elucidated. We suggest that CNV should be routinely analyzed during a genome-wide association study, which may reveal additional alleles with low to modest disease risk associated with RA.

## ACKNOWLEDGMENT

We thank Dr. Kai Wang for guidance with the analysis of copy number variation and the Wellcome Trust Case Control Consortium for providing the data.

## REFERENCES

1. Wordsworth BP, Bell JI. The immunogenetics of rheumatoid arthritis. Springer Semin Immunopathol 1992;14:59-78.
2. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. Nature Genetics 2009;41:1313-8.
3. Kidd JM, Cooper GM, Donahuet WF, Hayden HS, Sampas N, Graves T, et al. Mapping and sequencing of structural variation from eight human genomes. Nature 2008;453:56-64.
4. Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Rodijk-Olthuis D, van de Kerkhof PC, et al. Psoriasis is associated with increased ß-defensin genomic copy number. Nat Genet 2008;40:23-5.

5. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, et al. Origins and functional impact of copy number variation in the human genome. Nature 2010;464:704-12.
6. Fanciulli M, Norsworthy PJ, Petretto E, Dong R, Harper L, Kamesh L, et al. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. Nat Genet 2007;39:721-3.
7. McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, et al. Deletion polymorphism upstream of IRiGM associated with altered IRGM expression and Crohn's disease. Nat Genet 2008;40:1107-12.
8. Fernandez BA, Roberts W, Chung B, Weksberg R, Meyn S, Szatmari P, et al. Phenotypic spectrum associated with de novo and inherited deletions and duplications at 16p11.2 in individuals ascertained for diagnosis of autism spectrum disorder. J Med Genet 2009;47:195-203.
9. Yang T-L, Chen XD, Guo Y, Lei S-F, Wang J-T, Zhou Q, et al. Genomewide copy number variation study identified a susceptibility gene, UGT2B17, for osteoporosis. Am J Hum Genet 2008; 83:663-74.
10. McKinney C, Merriman ME, Chapman PT, Gow PJ, Harrison AA, Highton J, et al. Evidence for an influence of chemokine ligand 3 like 1 (CCL3L1) gene copy number on susceptibility to rheumatoid arthritis. Ann Rheum Dis 2008;67:409-13.
11. McKinney C, Fanciulli M, Merriman ME, Phipps-Green A, Alizadeh BZ, Koeleman BPC, et al. Association of variation in Fcγ receptor 3B gene copy number with rheumatoid arthritis in Caucasian samples. Ann Rheum Dis 2010;69:1711-6.
12. Thabet MM, Huizinga TWJ, Marques RB, Stoeken-Rijsbergen G, Bakker AM, Kurreeman FA, et al. Contribution of Fcγ receptor IIIA gene 158V/F polymorphism and copy number variation to the risk of ACPA positive rheumatoid arthritis. Ann Rheum Dis 2009;68:1775-80.
13. The Wellcome Trust Case Control Consortium. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 2010;464:713-20.
14. The 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.
15. Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SFA, et al. Penn CNV: An integrated hidden Markov model designed for high resolution copy number variation detection in whole genome SNP genotyping data. Genome Res 2007;17:1665-74.
16. Rabbee N, Speed T. A genotype calling algorithm for affymetrix SNP arrays. Bioinformatics 2006;22:7-12.
17. Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL, et al. A genome-wide investigation of SNPs and CNVs in schizophrenia. PLoS Genetics 2009;5:e1000373.
18. Dalila P, Pagnamenta AT, Klei L, Anney R, Merico D, Regan R, et al. Functional impact of global rare copy number variation in autism spectrum disorders. Nature 2010;466:368-72.
19. The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 2008;455:237-41.
20. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305:525-8.
21. Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL, et al. A chromosome 8 gene cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. Am J Hum Genet 2006;79:439-48.
22. Elder JT. Genomewide association scan yields new insights into the immunopathogenesis of psoriasis. Genes Immun 2009;10:201-9.
23. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PIW, Maller J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477-82.
24. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, et al. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythmatosus. Nat Genet 2008;40:1059-61.
25. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genomewide scan reveals association of psoriasis with IL23 and NFkB pathways. Nat Genet 2009;41:199-204.
26. Turer EE, Tavares RM, Mortier E, Hitotsumatsu O, Advincula R, Lee B, et al. Homeostatic MyD88 dependent signals cause lethal inflammation in the absence of A20. J Exp Med 2008;205:451-64.
27. Schmitz R, Hansmann ML, Bohle V, Martin-Subero JI, Hartmann S, Mechtersheimer G, et al. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. J Exp Med 2009;206:981-9.
28. Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Alfredsson L, Toes R, et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 2007;56:2202-10.
29. Dideberg V, Kristjansdottir G, Milani L, Libioulle C, Sigurdsson S, Louis E, et al. An insertion-deletion polymorphism in the interferon regulatory factor 5 (IRF5) gene confers risk of inflammatory bowel diseases. Hum Mol Genet 2007;16:3008-16.
30. Reddy KV, Serio KJ, Hodulik CR, Bigby TD. 5 Lipoxygenase-activating protein gene expression: key role of CCAAT/enhancer binding proteins (C/EBP) in constitutive and tumor necrosis factor (TNF) alpha induced expression in THP1 cells. J Biol Chem 2003;278:13810-8.
31. Griffiths RJ, Smith M, Roach ML, Stock JL, Stam EJ, Milici AJ, et al. Collagen induced arthritis is reduced in 5 lipoxygenase activating protein deficient mice. J Exp Med 1997;185:1123-30.
32. Inoue A, Matsumoto I, Tanaka Y, Iwanami K, Kanamori A, Ochiai N, et al. Tumor necrosis factor α induced adipose related protein expression in experimental arthritis and in rheumatoid arthritis. Arthritis Res Ther 2009;11:r118.
33. Bullinger L, Krönke J, Schön C, Radtke I, Urlbauer K, Botzenhardt U, et al. Identification of acquired copy number alterations and uniparental disomies in cytogenetically normal acute myeloid leukemia using high resolution single nucleotide polymorphism analysis. Leukemia 2010;24:438-49.
34. Takada H, Imoto I, Tsuda H, Sonoda I, Ichikura T, Mochizuki H, et al. Screening of DNA copy number aberrations in gastric cancer cell lines by array-based comparative genomic hybridization. Cancer Sci 2005;96:100-10.
35. Takata Y, Hamada D, Miyatake K, Nakano S, Shinomiya F, Scafe CR, et al. Genetic association between the PRKCH gene encoding protein kinase C-eta isozyme and rheumatoid arthritis in the Japanese population. Arthritis Rheum 2007;56:30-42.
36. Clements JL, Yang B, Ross-Barta SE, Eliason SL, Hrstka RF, Williamson RA, et al. Requirement for the leukocyte-specific adapter protein SLP76 for normal T cell development. Science 1998;281:416-9.
37. Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, et al. Segmental copy number variation shapes tissue transcriptomes. Nat Genet 2009;41:424-9.
38. Carter PN. Methods and strategies for analyzing copy number variation using DNA microarrays. Nat Genet 2007;39:S16-S21.