Genome-Wide Meta-Analysis of Sciatica in Finnish Population

Susanna Lemmelä1, Svetlana Solovieva1, Rahman Shiri1, Christian Benner2,3, Markku Heliovaa4, Johannes Kettunen5,6,7, Verneri Anttila8,9, Samuli Ripatti2,3,10, Markus Perola2,11,12, Ilkka Seppälä13, Markus Juonala14,16, Mika Kähönen16, Veikko Salomaa17, Jorma Viikari14,15, Olli T. Raitakari16,19, Terho Lehtimäki13, Aarno Palotie2,8,9,20, Eira Viikari-Juntura21, Kirsti Husgafvel-Pursiainen1*

1 Health and Work Ability, Finnish Institute of Occupational Health, 00250 Helsinki, Finland, 2 Institute for Molecular Medicine Finland (FIMM), 00014 University of Helsinki, Helsinki, Finland, 3 Department of Public Health, 00014 University of Helsinki, Helsinki, Finland, 4 Population Health Unit, National Institute for Health and Welfare, 00251 Helsinki, Finland, 5 Faculty of Medicine, Institute of Health Sciences, University of Oulu, 90200 Oulu, Finland, 6 NMR Metabolomics Laboratory, University of Eastern Finland, Kuopio, Finland, 7 National Institute for Health and Welfare, Helsinki, Finland, 8 Analytic and Translational Genomics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, United States of America, 9 Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States of America, 10 Welcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, CB10 1SA, United Kingdom, 11 Public Health Genomics Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, 00271 Helsinki, Finland, 12 The Estonian Genome Center, University of Tartu, 51010 Tartu, Estonia, 13 Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere School of Medicine, 33520 Tampere, Finland, 14 Division of Medicine, Turku University Hospital, 20521 Turku, Finland, 15 Department of Medicine, University of Turku, 20521 Turku, Finland, 16 Department of Clinical Physiology, Tampere University Hospital, 33521 Tampere, Finland, 17 Department of Health, National Institute for Health and Welfare, 00251 Helsinki, Finland, 18 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, 20520 Turku, Finland, 19 Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, 20521 Turku, Finland, 20 Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts 02114, United States of America, 21 Disability Prevention Centre, Finnish Institute of Occupational Health, 00250 Helsinki, Finland

* kirsti.husgafvel-pursiainen@ttl.fi

Abstract

Sciatica or the sciatic syndrome is a common and often disabling low back disorder in the working-age population. It has a relatively high heritability but poorly understood molecular mechanisms. The Finnish population is a genetic isolate where small founder population and bottleneck events have led to enrichment of certain rare and low frequency variants. We performed here the first genome-wide association (GWAS) and meta-analysis of sciatica. The meta-analysis was conducted across two GWAS covering 291 Finnish sciatica cases and 3671 controls genotyped and imputed at 7.7 million autosomal variants. The most promising loci (p < 1x10^-8) were replicated in 776 Finnish sciatica patients and 18,489 controls. We identified five intragenic variants, with relatively low frequencies, at two novel loci associated with sciatica at genome-wide significance. These included chr9:14344410:I (rs71321981) at 9p22.3 (NFIB gene; p = 1.30x10^-8, MAF = 0.08) and four variants at 15q21.2: rs145901849, rs80035109, rs190200374 and rs117458827 (MYO5A; p = 1.34x10^-8, MAF = 0.06; p = 2.32x10^-8, MAF = 0.07; p = 3.85x10^-8, MAF = 0.06; p = 4.78x10^-8, MAF = 0.07, respectively). The most significant association in the meta-
analysis, a single base insertion rs71321981 within the regulatory region of the transcription factor NFIB, replicated in an independent Finnish population sample (p = 0.04). Despite identifying 15q21.2 as a promising locus, we were not able to replicate it. It was differentiated; the lead variants within 15q21.2 were more frequent in Finland (6–7%) than in other European populations (1–2%). Imputation accuracies of the three significantly associated variants (chr9:14344410:I, rs190200374, and rs80035109) were validated by genotyping. In summary, our results suggest a novel locus, 9p22.3 (NFIB), which may be involved in susceptibility to sciatica. In addition, another locus, 15q21.2, emerged as a promising one, but failed to replicate.

Introduction

Low back pain is a global health problem affecting all age groups [1, 2]. Sciatica—usually a clinical manifestation of lumbar disc herniation (OMIM 603932)—is a common low back disorder with a population prevalence of about 5%; it is often disabling in the working age population [3, 4]. Sciatica is a complex disorder, with both genetic and environmental factors involved [4, 5]. Sciatic pain or lumbar radicular pain—the typical symptom of sciatica— is defined as pain radiating from the back down to the leg, usually caused by compression or irritation of one of the lumbosacral nerve roots [6–9].

Twin and family studies have revealed a substantial genetic component in low back disorders with heritability estimates of approximately 20–40% for sciatica and 35–75% for lumbar disc degeneration (OMIM 603932) [10–12]. Candidate gene studies of low back disorders have typically focused on functional genes that associate with cartilage structure and stability, pain signaling, obesity, or inflammation [13–15].

The Finnish population is one of the most thoroughly characterized genetic isolates. It has its ancestry in a small founder population, followed by several bottle neck events and genetic drift that has led to the enrichment of certain rare and low frequency variants that are almost absent in many other European populations [16]. Presently, there are several representative Finnish population cohorts with genome-wide data available, and these have contributed to an array of successful large GWAS consortia [17–25].

The current study was conducted in two of the Finnish population cohorts, the Young Finns Study (YFS) and Health 2000 Study (H2000), both with data on physician-diagnosed sciatica available. We rationalized that the unique Finnish population features together with the representative population-based studies allow us to investigate possible influence of not only common variants but also discover rare or low frequency variants involved in the development of sciatica.

In the present study, we performed the first genome-wide association studies (GWAS) and GWAS meta-analysis of sciatica (291 sciatica cases and 3671 controls), utilizing the special benefits of the Finnish population isolate and the power of 1000 Genomes imputation. We validated imputation accuracies of three significantly associated variants by genotyping them in the discovery cohorts. The most promising loci (p < 1x10^-6) were replicated in an independent population-based sample of 776 Finnish sciatica patients and 18,489 controls.
Results

Genome-wide meta-analysis

We conducted a meta-analysis of sciatica across two Finnish GWAS covering 291 sciatica cases and 3,671 controls genotyped or imputed at 7.7 million autosomal variants (Fig 1; S1 Table). Both GWAS (YFS and H2000) (see URLs; S1 Text) are based on Finnish population-based cohorts with dense genome-wide genotyping and 1000 Genomes imputation data (Fig 1; Table 1; S1 Table).

In the meta-analysis of GWAS data, a total of five novel variants within two loci achieved genome-wide significance ($p < 5 \times 10^{-8}$). These were insertion chr9:14344410:I (also known as rs71321981) at 9p22.3 ($p = 1.30 \times 10^{-8}$; MAF = 0.08) and rs145901849, rs80035109, rs190200374 and rs117458827 at 15q21.2 ($p = 1.34 \times 10^{-8}$, MAF = 0.06; 2.32 $\times 10^{-8}$, MAF = 0.07; 3.85 $\times 10^{-8}$, MAF = 0.06; 4.78 $\times 10^{-8}$, MAF = 0.07, respectively). The between study heterogeneity ($I^2$) ranged between 0–0.63 and the imputation quality was high (0.77–0.99) for the associated variants (Table 1; S2 and S3 Tables). From these, the insertion chr9:14344410:I (rs71321981) is a regulation region variant in the first intron of the NFIB gene. The 15q21.2 variants (an intronic regulation region variant rs145901849; intronic SNPs rs80035109 and rs190200374, as well as a 3' UTR SNP rs117458827) are located within ~200kb region in the MYO5A gene and are in a strong linkage disequilibrium (LD) ($r^2 \geq 0.81$) (Table 1; S4 Table; Figs 2 and 3; 1000 Genomes Project; see URLs).

All five significantly associated variants have relatively low frequencies (MAF≤0.08). The insertion chr9:14344410:I at 9p22.3 (NFIB) has similar frequencies in the Finnish and other European populations (7% and 8%, respectively), whereas all four lead variants at the locus 15q21.2 (MYO5A) were more frequent in Finnish population (6–7%) than in other Europeans (0–2%, respectively) (Table 2; S5 Table; 1000 Genomes Project; see URLs).

In addition, 176 variants at 30 loci showed suggestive associations with sciatica ($p < 1.0 \times 10^{-5}$) (Fig 2; S1 Fig; S2 and S3 Tables). Of these, 45 variants were within a 1.4 Mb region at 6p21.32–33 in the HLA gene region, with six lead SNPs having $p < 1 \times 10^{-6}$. Four of these (s115949512, rs3094014, rs114615271, rs115688765) were in perfect LD ($r^2 \geq 0.97$) within or nearby the HCP5.
gene at 6p21.33, and two (rs190606317 and rs115488695) were in moderate LD ($r^2 \geq 0.31$) within or nearby the HLA-DRB5 gene at 6p21.32 (S6 Table). The locus 6p21.32 has previously been associated with lumbar disc degeneration and osteoarthritis (OMIM 165720) [26, 27] (Fig 3).

The genome-wide inflation factor in the meta-analysis was low ($\lambda_{GC} = 0.99$). Manhattan and Quantile-Quantile (QQ) -plots for meta-analysis of adjusted genome-wide association results (adjusted for age, sex and the seven principal components of the genetic data) are shown in Fig 2 and S1 Fig. Regional plots for the associated loci are shown in the Fig 3. Manhattan and QQ -plots of adjusted individual GWAS are shown in S2, S3, S4 and S5 Figs.

**Replication**

From the meta-analysis, we selected 30 most promising SNPs ($p < 1 \times 10^{-6}$) representing eight loci for replication in an independent Finnish sample of 776 sciatica cases and 18,489 controls from the FINRISK population survey (FINRISK; see URLs). The most significantly associated variant in the meta-analysis (insertion chr9:14344410:I; rs71321981 at 9p22.3, $p = 1.30 \times 10^{-8}$).

### Table 1. Sample demographics.

| Study   | Status | N   | Age   | Female % | BMI | Smoking % (Smokers/Non-smokers) | PA % (Very low or no/Active) |
|---------|--------|-----|-------|----------|-----|---------------------------------|-------------------------------|
| YFS     | All    | 2020| 37.7  | 5.0      | 55  | 26.0                            | 23 (450/1468)                | 23 (455/1561)                |
|         | Case   | 180 | 39.2  | 4.8      | 64  | 26.6                            | 28 (48/123)                  | 21 (38/141)                  |
|         | Control| 1840| 37.6  | 5.0      | 54  | 25.9                            | 23 (402/1345)               | 23 (417/1420)               |
| H2000   | All    | 1942| 50.4  | 10.9     | 51  | 27.2                            | 29 (571/1365)               | 25 (470/1443)               |
|         | Case   | 111 | 54.4  | 10.6     | 55  | 28.0                            | 29 (32/79)                  | 27 (30/80)                  |
|         | Control| 1831| 50.1  | 10.9     | 50  | 27.2                            | 30 (539/1286)               | 24 (440/1363)               |
| FINRISK | All    | 19 265| 48.1 | 13.3     | 55  | 26.8                            | 27 (5144/13961)             | NA                          |
| FINRISK | Case   | 776 | 50.7  | 12.3     | 48  | 27.7                            | 28 (212/558)                | NA                          |
| FINRISK | Control| 18 489| 48.0 | 13.3     | 55  | 26.8                            | 27 (4932/13403)             | NA                          |

N, Number of individuals; BMI, Body Mass Index; S.D., Standard Deviation; Smoking, Percentage of smokers (Numbers of smokers vs non-smokers given for each group); PA, Percentage of subjects with very low or no physical activity (Numbers of those with no physical activity or up to 3 times a month vs once a week or more frequently); NA, Not available. YFS, The Cardiovascular Risk in Young Finns Study; H2000, The Health 2000 Study; FINRISK, a subsample (years 1992, 1997, 2002, 2007) of the FINRISK Study. Values given represent those at the time of the questionnaire.

doi:10.1371/journal.pone.0163877.t001
showed association with sciatica in the replication sample (p = 0.04) (S7 Table). No other replications were identified (S7 Table). The rs190606317 at 6p21.32 as well as rs62100562 at 18q22.3 showing suggestive associations in the meta-analysis (p < 1x10^-6) (S7 Table) had a significant p-value in replication sample (p = 0.006, p = 0.03, respectively), but the effect direction was different and thus was not considered as replicated (S7 Table).

**Genotype validation.** For validating the imputation accuracy, we sequenced the insertion variant chr9:14344410:I (rs71321981) at 9p22.3 (NFIB) in the two discovery cohorts, with 184
Table 2. SNPs exhibiting genome-wide significant association with sciatica in the GWAS meta-analysis.

| SNP          | Type            | Chr  | Position* (bp) | Gene | EA/OA | Analysis (GWAS/Replication) | EAF          | Imput. quality# | OR* (95% CI)    | beta   | SE    | P value* | Phet | I²   |
|--------------|-----------------|------|----------------|------|-------|-----------------------------|--------------|-----------------|-----------------|--------|-------|----------|------|------|
| chr9:14344410:t; rs71321981 | regulatory region | 9p22.3 | 14344410       | NFB  | AG/A  | YFS                         | 0.08         | 0.13            | 0.78            | 3.17   | (1.94–5.18) | 1.15     | 0.25 | 4.09x10⁻⁶ |
| H2000        |                 |      |                |      |       | 2.89 (1.62–5.19)            | 1.06         | 0.30            | 0.54x10⁻⁴      |        |       |          |      |      |
| Meta-analysis|                 |      |                |      |       | 3.05 (2.08–4.49)            | 1.12         | 0.20            | 1.30x10⁻⁸      | 0.82   | 0     |          |      |      |
| Replication@ |                 |      |                |      |       | 1.17 (0.97–1.40)            | 0.16         | 0.09            | 0.04           |        |       |          |      |      |
| rs145901849  | regulatory region | 15q21.2 | 52640539       | MYO5A| T/C   | YFS                         | 0.06         | 0.12            | 0.91            | 3.84   | (2.38–6.20) | 1.35     | 0.24 | 3.83x10⁻⁹ |
| H2000        |                 |      |                |      |       | 2.01 (1.08–3.75)            | 0.70         | 0.32            | 2.8x10⁻²       |        |       |          |      |      |
| Meta-analysis|                 |      |                |      |       | 3.04 (2.07–4.45)            | 1.11         | 0.19            | 1.34x10⁻⁸      | 0.11   | 0.61  |          |      |      |
| Replication@ |                 |      |                |      |       | 1.00 (0.66–1.54)            | 0.45         | 0.39            | 0.22           |        |       |          |      |      |
| rs80035109   | intronic        | 15q21.2 | 52665890       | MYO5A| C/T   | YFS                         | 0.07         | 0.14            | 0.97            | 3.01   | (1.97–4.59) | 1.10     | 0.22 | 3.73x10⁻⁷ |
| H2000        |                 |      |                |      |       | 2.10 (1.19–3.68)            | 0.74         | 0.29            | 9.6x10⁻³       |        |       |          |      |      |
| Meta-analysis|                 |      |                |      |       | 2.65 (1.88–3.72)            | 0.97         | 0.17            | 2.3x10⁻²       | 0.32   | 0     |          |      |      |
| Replication@ |                 |      |                |      |       | 1.11 (0.91–1.35)            | 0.10         | 0.10            | 0.25           |        |       |          |      |      |
| rs190200374  | intronic        | 15q21.2 | 52811959       | MYO5A| T/G   | YFS                         | 0.06         | 0.13            | 0.84            | 3.64   | (2.28–5.83) | 1.29     | 0.24 | 6.72x10⁻⁸ |
| H2000        |                 |      |                |      |       | 1.88 (1.01–3.51)            | 1.88         | 0.32            | 4.7x10⁻²       |        |       |          |      |      |
| Meta-analysis|                 |      |                |      |       | 2.89 (1.98–4.21)            | 1.06         | 0.19            | 3.85x10⁻⁸      | 0.10   | 0.63  |          |      |      |
| Replication@ |                 |      |                |      |       | 1.10 (0.90–1.34)            | 0.10         | 0.10            | 0.33           |        |       |          |      |      |
| rs117458827  | 3’UTR           | 15q21.2 | 52600066       | MYO5A| A/G   | YFS                         | 0.07         | 0.14            | 0.99            | 2.85   | (1.88–4.32) | 1.05     | 0.21 | 8.19x10⁻⁷ |
| H2000        |                 |      |                |      |       | 2.05 (1.19–3.54)            | 0.72         | 0.28            | 9.9x10⁻³       |        |       |          |      |      |
| Meta-analysis|                 |      |                |      |       | 2.53 (1.81–3.53)            | 0.93         | 0.17            | 4.78x10⁻²      | 0.35   | 0     |          |      |      |
| Replication@ |                 |      |                |      |       | 1.06 (0.88–1.29)            | 0.06         | 0.10            | 0.51           |        |       |          |      |      |

The respective data from the two discovery GWAS (YFS, H2000), meta-analysis and replication cohort are shown.

*Chromosomal positions are based on NCBI build 37;

#Imputation quality score from IMPUTE;

*Additive model, adjusted for seven principal components, age and gender;

@The FINRISK Study.

Abbreviations: SNP, single nucleotide polymorphism; Type, type of variant; Chr, chromosomal locus; EA, effect allele; OA, other allele; EAF, effect allele frequency; —, not applicable; OR (95% CI), odds ratio (95% confidence interval); beta, effect size; SE, standard error of beta; P value, Cochran’s heterogeneity statistic’s p-value; I², heterogeneity index.

doi:10.1371/journal.pone.0163877.t002
individuals (92 cases and 92 controls) in YFS, and 184 individuals (89 cases and 95 controls) in H2000 (Fig 4). The concordance between the sequenced and the imputed genotypes were 88.2% in YFS (imputation quality 0.78; MAF 0.08) and 87.3% in H2000 (imputation quality 0.77; MAF 0.07). We also genotyped rs190200374 and rs80035109 at 15q21.2 (MYO5A) in a total of 1686 (152 cases and 1534 controls) and 1642 (154 cases and 1488 controls) individuals in YFS as well as 1405 (82 cases and 1323 controls) and 1392 (99 cases and 1293 controls) individuals in H2000, respectively. For rs190200374, the concordance between the genotyped and imputed genotypes was 96.7% in YFS (imputation quality 0.84; MAF 0.06) and 96.2% in H2000 (imputation quality 0.87; MAF 0.06). For rs80035109, these were 99.6% in YFS (imputation quality 0.97; MAF 0.07) and 98.3% in H2000 (imputation quality 0.97; MAF 0.07). All concordances as well as results from the association analysis using the real genotypes are given in S8 Table. Both rs190200374 and rs80035109 are in strong LD with two other significantly associated variants (rs145901849 and rs117458827) in the MYO5A gene region (15q21.2) (r² ≥ 0.81) (S4 Table). In summary, our additional genotyping assessments were able to validate the accuracy of the imputed genotypes for chr9:14344410:I (9p22.3, NFIB), rs190200374 and rs80035109 (15q21.2, MYO5A), with high concordances obtained for all.

Discussion

The present study is the first GWAS and meta-analysis of sciatica. For lumbar disc degeneration, two genome-wide association analyses have been reported [26, 30]. The strongest association signal (p = 1.30x10⁻⁸) for sciatica in our present study was obtained for a single base insertion -/-G (chr9:14344410:I; rs71321981) within NFIB gene at locus 9p22.3. To our knowledge, this locus has not been associated with sciatica earlier. We were able to replicate this association in an independent Finnish population cohort (FINRISK, rs71321981, p = 0.04). Frequency of rs71321981 is about 8% in Finns as well as in other European populations.

The insertion rs71321981 resides in the regulatory region within the first of eight introns of the NFIB gene causing a single base extension to the sequence with regard to the reference sequence. It overlaps also a novel antisense gene, RP11-120J1.1, being intronic or upstream gene variant for that depending on the transcript (Ensemble Variant effect predictor; see URLs).

NFIB is a member of the nuclear factor I (NFII) family of evolutionarily conserved genes (NFIА, NFIB, NFIC and NFIX) encoding sequence-specific DNA-binding proteins, transcription factors, which regulate transcription throughout the development in mammals and in
adenoviral replication [31]. NFI transcription factors bind to DNA as homo- and heterodimers [32–34] and control, through promoter and cell-type specific transactivation or repression, a diverse set of tissue-specific and developmentally regulated genes (e.g. in the central nervous system, muscle, and the lungs, as well as in various cell types such as fibroblasts, chondrocytes, osteoblasts, adipocytes) [31, 35–38]. The insertion discovered here is a regulatory region variant located in the C-terminal transcriptional activation and/or repression domain of NFIB [35] (1000 Genomes Project; see URLs).

In vitro work has suggested that NFIB plays a role in chondrocyte differentiation [39]. In murine mesenchymal ATDC5 cells, a cell line used as a model for chondrocyte differentiation, the creation of a NFIB truncation mutation resulted in lack of the C-terminal transactivation/repression domain, led to an impaired nodule formation, less accumulation of cartilaginous matrices, and the reduced expression of a set of marker genes for proliferating chondrocytes, namely Col2a1, Matn-1, PTHrP, and to some extent also of SOX9 [39]. Other studies have indicated that NFI/NFIB proteins bind to a promoter silencer region of Matn1 and modulate Sox9 transactivation in in vitro chondrogenesis [40, 41]. NFI has also been demonstrated to function as a positive regulator of Runx2-dependent skeletal development and osteoblast function [37].

At 15q21.2, four SNPs; rs145901849 (intronic regulatory region variant), rs117458827 (3’UTR SNP), rs80035109 and rs190200374 (both intronic), were significantly associated with sciatica (p < 5x10^-8). These variants are in strong LD (r^2 > 0.81) within a 200kb region in the MYO5A gene. However, we were not able to replicate the 15q21.2 variants in the Finnish replication sample and thus we consider 15q21.2 as a promising locus but needing further replication efforts. In addition, thirteen surrounding SNPs at 15q21.2–15q21.3 harboring the MYO5A, FAM214A, ARP19, LYSMD2, DMXL2, and MYO5C genes as well as an intergenic region showed suggestive associations (p < 1x10^-5).

Class V myosins (MYO5A, MYO5B, MYO5C; ~50–60% protein sequence identity shared) are an evolutionarily ancient group of molecular motors that mediate actin-dependent organelle trafficking [42, 43]. MYO5A is involved in the intracellular transport of organelles in melanocytes and neuronal cells [44–46], and mutations within it have been associated with rare human syndromes with neurological defects [47–49]. Interestingly, a suggestive association was reported for rs4802666 of MYH14 (19q13.33) in a GWAS meta-analysis of lumbar disc degeneration [26]. MYH14 is a member of the same myosin superfamily as MYO5A and MYO5C; all are expressed in human cell lines derived from bone (Human Protein Atlas; see URLs) and normal skeletal muscle tissue (Gene Cards; see URLs). It has been suggested that myosins may play a role in lumbar disc degeneration through mechanisms that affect multiple tissues, rather than cartilage alone [26]. To sum up, 15q21.2 (MYO5A) emerged as an interesting locus but failed to replicate in our replication sample. Further replication efforts in other populations are needed. In this context, it is of note that the MYO5A variants discovered appear enriched in the Finnish population (somewhat over 5%) and are rather rare in other European populations (up to 2%) (1000 Genomes Project; see URLs) as discussed in more detail below.

At locus 6p21.32–33, 45 SNPs located within a 1.4 Mb region in the HLA gene region (i.e., human major histocompatibility complex) showed suggestive associations with sciatica (p < 1x10^-5). Six of them had a p-value < 1x10^-6, of which four (rs115949512, rs3094014, rs114615271 and rs115688765) were in strong LD (r^2 > 0.97) within or nearby the HCP5 gene at 6p21.33 and two (rs190606317, rs115488695) were in moderate LD (r^2 > 0.31) within/nearby the HLA-DRB5 gene at 6p21.32 (S6 Table). We were not able to replicate association between these six lead SNPs and sciatica in our independent Finnish replication sample. The rs190606317 at 6p21.32 had a significant p-value (p = 0.006) in the sample, but the effect direction was different and thus was not considered as replicated (S7 Table).
The HLA gene region (6p21.32) has previously been associated with musculoskeletal disorders. Two SNPs (rs2187689 and rs7767277) were associated with lumbar disc degeneration (LDD) in a GWAS meta-analysis in Northern European populations [26], and two variants (rs10947262 and rs7775228) were associated with osteoarthritis (OA) in the GWAS of a Japanese population [27], although this association was not replicated in European or Han Chinese study populations [50, 51]. The OA-associated SNPs (rs7775228, rs10947262) from the Japanese study were not in LD with our six lead variants in the HLA gene region, and the LDD-associated variants (rs2187689 and rs7767277) were not included in our GWAS (not included in the 1000 Genomes 1KG pilot data) (S6 Table). It may be noted that neither sciatica, osteoarthritis, nor LDD is considered as an auto-immune disease. However, it has been suggested that there may be pro-inflammatory cytokine activation in herniated lumbar discs [5, 52], and anti-TNF has been used successfully to treat disc herniation-induced sciatica [5, 53]. Accordingly, inflammatory mediator genes are considered to be candidate genes for sciatica, lumbar disc degeneration and osteoarthritis [5, 14, 28, 54, 55].

It is also of note that COL11A2, previously associated with magnetic resonance-determined disc signal intensity (with rs2076311 as the lead SNP) and with degenerative lumbar spinal stenosis with radicular pain (the lead SNP rs1799907) [28, 29], lies 622 kb upstream of our rs190606317 at 6p2.32. However, our six lead variants were not in LD with these COL11A2 variants (rs2076311, rs1799907) (S6 Table), so it seems unlikely that COL11A2 would directly account for the observed association in the meta-analysis, although we cannot rule out some possible influence of long range effectors such as enhancers.

The primary phenotype in the two discovery GWAS was a physician-diagnosed sciatica. Sciatica is a syndrome involving nerve root impingement or inflammation that has progressed sufficiently to cause neurological symptoms in the areas that are supplied by the affected nerve roots [8]. While there is a range of definitions of sciatica, its specific clinical features—such as low back pain radiating below the knee, presence of numbness or pins and needles in a dermatomal distribution, positive results on a straight leg raise test, and weakness or reflex changes, or both, in a myotomal distribution—are used for diagnoses of sciatica [7, 9]. Further, sciatica is commonly associated with disc disorders such as herniated disk or spinal stenosis [9], but still in many cases with clinical symptoms of sciatica, no lumbar disc herniation is present on images [9, 56, 57].

In YFS, the physician-diagnosed sciatica cases were self-reported in an on-site examination and represented manifestations of sciatica in a relatively young population. In H2000, sciatica was diagnosed by a field physician, if the subject had a history of low back pain radiating down to the leg, and either positive clinical findings or a history of lumbar disc herniation that had previously been confirmed by imagining or required surgery. Finally, in the replication population (FINRISK, the largest of the study populations), the sciatica phenotype was based on ICD diagnoses available at the Finnish Hospital Discharge Register, thus likely to represent the severe cases requiring hospitalization. Collectively, while the phenotypes across the three study cohorts were not identical, they were all representative of sciatica. Therefore, we see that the study managed to capture representation from the whole sciatica spectrum. Furthermore, in all study populations, the diagnoses were made by Finnish physicians according to general medical practice in Finland, a country with high-level and uniform medical education and well organized health services. Obviously, some residual differences between the diagnoses and study populations may partly explain why only one of the loci (9p22.3) was replicated in the FINRISK population.

This study is based on three representative Finnish population samples representing a genetic isolate, where certain rare and low frequency variants are enriched due to population history (small founder population, several bottle neck events and genetic drift). Each of the five
lead SNPs in the two loci (9p22.3, 15q21.2) has a relatively low frequency (<8%) in the general population (1000 Genomes Project; see URLs). The insertion chr9:14344410:I at 9p22.3 (NFIB) has similar frequencies in the Finnish (7%) and other European populations (8%) as well as in American (6%) and South Asian populations (8%) (1000 Genomes Project; see URLs). In African population, the frequency of rs71321981 is even higher (24%), but in East Asian population this variant is absent (0%). Most of these minor allele frequencies would allow replication; we therefore conclude that further replication efforts are warranted. For 15q21.2, however, all four MYO5A variants were 3–6 times more frequent in our representative Finnish samples (6–7%) than in other European populations (0–2%) and in African, American, East and South Asian populations they were absent (Table 2; S5 Table; 1000 Genomes Project; see URLs). The differences in allele frequencies likely indicates a past bottleneck events and genetic drift in the Finnish population. In all, our data are in line with previous studies illustrating the high utility of population isolate and the dense genotype imputation based on representative data from multiple populations, in search for low frequency variants associated with complex human diseases [21, 58–60]. The five relatively low frequency variants associated with sciatica in this study, especially the differentiated variants in 15q21.2, would likely not have been identified without inclusion of Finnish individuals in the 1000 Genomes reference panel. Furthermore, the identification and replication of the differentiated variants would require much larger sample sizes in more mixed population.

In summary, we conducted the first GWAS meta-analysis of sciatica and identified a single base insertion at locus 9p22.3 (NFIB) associated with sciatica at genome-wide significance and replicated in an independent Finnish population sample. The insertion is within a regulation region of the transcription factor NFIB, which has been shown to be involved in chondrocyte differentiation and osteoblast function, thus making this gene and the insertion functionally interesting for sciatica. In addition, we identified four variants associated with sciatica at the locus 15q21.2 (MYO5A), which was promising but not replicated. Both loci merit further investigation and replication studies. As the first GWAS of sciatica, this study may serve as a starting point for further studies and shed light to the genetic susceptibility factors of sciatica.

Materials and Methods

Study populations and phenotypes

This study was carried out in accordance with the recommendations of the Declaration of Helsinki. All participants of studies have given written informed consent. Studies were approved by the local research ethic committees: Ethics Committee of the National Public Health Institute for the Health 2000 Study, Ethics Committee of the Hospital District of Southwest Finland for the Young Finns Study, and Ethics Committee of Helsinki and Uusimaa Hospital for the FINRISK Study.

Genome-wide association studies of sciatica were carried out in two Finnish population-based cohorts, the Young Finns Study (YFS; 180 sciatica cases and 1,840 controls) and the Health 2000 Study (H2000; 111 sciatica cases and 1,831 controls) (Fig 1; S1 Table; S1 Text). The primary phenotype analyzed was a physician-diagnosed sciatica, with diagnosis based on specific symptoms and clinical findings according to general medical practice in Finland. In the YFS, information on physician-diagnosed sciatica was inquired during on-site examinations using a self-administered questionnaire (“Do you currently have or have you had a long-term disease diagnosed by a physician, such as sciatica?”). In the H2000, participants attended a comprehensive health examination, with a physical examination of the musculoskeletal system performed by a field physician. The diagnosis of sciatica was based on the presence of chronic (>3 months) low back pain radiating down to the leg and either clinical findings of
lumbar nerve root compression or a history of lumbar disc herniation that had been previously verified by imaging or required surgery (S1 Text). Demographics of the population samples are given in Table 1 (see also S1 Text).

Replication analyses were carried out in an independent Finnish population sample consisting of four independent cross-sectional population surveys (carried out in years 1992, 1997, 2002, and 2007) of the FINRISK Study [61] (Table 1; S1 Text; see URLs). The FINRISK Study populations have been linked to the Finnish Hospital Discharge Register (currently the Finnish Care Register for Health Care) (see URLs), which provides personal identification code-based individual diagnoses (WHO ICD codes) at discharge. For the replication study, those diagnosed with one of the ICD-codes selected a priori by two experts on musculoskeletal diseases (EVJ and MH) as relevant for sciatica or sciatic syndrome (ICD8 353, 728.8; ICD9 724.3, 722.1, 722.10, 722.5, 722.52, 355.0; ICD10 M54.3, M51.1, M54.1, M54.4) were included as cases (amounting to 776 sciatica cases and 18,489 controls) (see URLs; S1 Text).

**Genome-wide scans and imputation**

Genotypes for both YFS and H2000 study populations were pre-existing, determined at the Wellcome Trust Sanger Institute (UK) using custom-generated Illumina Human Map 670K array for YFS and 610K BeadChip for H2000 study. Prior to genome-wide association analysis, quality control was performed independently in both two cohorts. Poor quality markers (those with genotyping failure >5% of samples) and poor quality DNA samples (those with genotyping failure >5% of markers) were removed. Moreover, markers with low minor allele frequency (MAF <0.01 in YFS and MAF <0.02 in H2000), Mendelian errors, or those violating the Hardy-Weinberg equilibrium (HWE \( \leq 1 \times 10^{-6} \) in YFS and in H2000) were excluded. Likewise, samples with gender inconsistency or cryptic relatedness (PI_HAT >0.2) and samples with excessive genome-wide heterozygosity (indicative of sample contamination) were removed. IBD sharing was also computed for the combined dataset, and duplicates and close relatives (PI_hat >0.4) were removed from the analyses. The genotype imputation data used were generated from cleaned data in both cohorts using IMPUTE2–program [62] and was based on the 1000 Genomes imputation reference in NCBI build 37, where the 1000 Genomes files were from March 2012 release for YFS and April 2012 release for H2000 (see URLs). The 1000 Genomes imputation reference includes Finnish imputation reference (FIN; see URLs). Quality control for imputed markers was conducted separately in both studies; markers with MAF <0.02 or imputation quality <0.7 were excluded. Genotyping, imputation and quality control measures are summarized in S1 Table.

**Genome-wide association analyses and meta-analysis**

The genome-wide scan data was analyzed for associations between genetic variants and sciatica separately in YFS and H2000 studies (Fig 1). Multidimensional Scaling was done for genetic data of both studies using PLINK v. 1.07 [63]. Genome-wide association analyses were performed for directly genotyped and imputed variants. Both studies included as covariates age, sex and the first seven principal components from the genetic data to correct possible population stratification. Frequentist/case-control-test, assuming an additive genetic model, was performed using SNPTEST v2 for both genotyped and imputed markers [64]. To combine the effect estimates from both studies, a fixed-effects meta-analysis was conducted for sciatica using GWAMA [65]. The presence of heterogeneity across studies was investigated with Cochran’s Q (weighted sum of squares test) and \( I^2 \) statistic (percentage of true heterogeneity to total observed variation) [66]. Only good quality autosomal markers passing the following criteria: imputation informativeness >0.7, and no heterogeneity in the effect sizes for the SNP...
between cohorts (Cochran’s Q statistic P-value < 1x10^-5), were included in further evaluations. The genome-wide inflation factor was measured in the individual GWAS and the GWAS meta-analysis by genomic control statistic [67]. There was no evidence for population stratification at the study level (genomic inflation factor; YFS λ_{GC} = 1.001 and H2000 λ_{GC} = 1.016) or at the meta-analysis level (λ_{GC} = 0.993). Test statistics of both GWAS were corrected by using a genomic inflation factors. The Quantile-Quantile and Manhattan plots were created using R-2.11 (see URLs) to visualize genome-wide association results. Regional plots of association were generated using LocusZoom [68] (see URLs). The genomic positions throughout this study are based on NCBI human genome build 37 (see URLs). GWAS, GWAS meta-analysis, and quality control measures are summarized in S1 Table.

**Replication analysis**

Variants with p-value < 1 x 10^{-6} in the genome-wide meta-analysis were selected for replication. Variant was considered replicated if it reached significance of p<0.05 and was consistent in terms of risk allele.

A large Finnish replication cohort, the FINRISK Study, was used for replication. Genotypes for FINRISK were pre-existing and, due to the large number of participants (close to 21 000), genotyping was performed in multiple batches/subpopulations using several standard genotyping arrays including: Sanger CoreExome batch1, Illumina HumanCoreExome Sanger CoreExome batch2, Illumina HumanCoreExome Broad CoreExome batch1, Illumina HumanCoreExome PredictCVD, Illumina OmniX Corogene, Illumina 610K SUMMIT, Illumina OmniX MIGEN—and Affy 6. Genotyping quality was examined by a detailed QC procedure consisting of success rate checks, duplicated water controls and Hardy Weinberg Equilibrium (HWE) testing. The genotype imputation data used was generated from cleaned data using IMPUTE2—program and was based on the 1000 Genomes imputation reference panel.

Frequentist association test assuming an additive genetic model was performed using SNPTEST v2 for selected variants [64].

**Genotype validation**

To validate imputation accuracy, the imputed variant chr9:14344410:I (rs71321981) at 9p22.3 was genotyped by direct sequencing using standard methods as described elsewhere [69]. The chromatograms were analyzed manually, and the corresponding nucleotide sequences were compared to the reference sequence at 1000 Genomes browser (1000 Genomes Project; see URLs). The primer sequences are available from the authors on request.

In addition, imputed variants rs190200374 and rs80035109 at 15q21.2 were genotyped using a TaqMan® chemistry-based PCR platform (Open Array™ system) and custom-made TaqMan® SNP Genotyping assays (Applied Biosystems). The allelic calling analysis was performed using TaqMan Genotyper v1.3 software and OpenArray™ SNP Genotyping Analysing software. For quality control, two independent readers interpreted the results. Random selection of all samples (about 5% in H2000 and 12% in YFS) was re-genotyped. No discrepancies were discovered in the replicate tests for the variants.

Concordances between the genotyped and imputed SNPs were calculated using threshold 0.7 for converting probabilistic genotypes of imputed SNPs to hard calls.

**Web Resources**

The URLs for data presented herein are as follows:

GWAMA, http://www.well.ox.ac.uk/gwama/

IMPUTE2, http://mathgen.stats.ox.ac.uk/impute/impute_v2.html
SNPTEST, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/old/snptest_v2.1.1.html
LocusZoom, http://csg.sph.umich.edu/locuszoo/m
R, http://www.r-project.org/
1000 Genomes Project, http://www.1000genomes.org/
Ensemble Variant Effect Predictor: http://www.ensembl.org/info/docs/tools/vep/index.html
Human Protein Atlas, http://www.proteinatlas.org/
Gene Cards, http://www.genecards.org/
Cardiovascular Risk in Young Finns Study, http://youngfinnsstudy.utu.fi/
Health 2000 Study, http://www.nationalbiobanks.fi/index.php/studies2/8-health2000
FINRISK, http://www.nationalbiobanks.fi/index.php/studies2/7-finrisk
Finnish Care Register for Health Care, https://www.thl.fi/fi/web/thlfi-en/statistics/information-on-statistics/register-descriptions/care-register-for-health-care

Supporting Information
S1 Appendix. Raw result data of the meta-analysis (with variants with $p < 0.05$).
(TXT)
S2 Appendix. Raw result data of the meta-analysis (with variants with $p < 0.001$).
(TXT)
S1 Fig. Quantile-quantile plot of observed against expected p-values for the genome-wide meta-analysis of sciatica.
(TIF)
S2 Fig. Manhattan plot for the Young Finns Study GWAS of sciatica. Results are adjusted for the first seven principal components of genetic data and for sex and age. Variants in red have p-value below genome-wide significance level ($p < 5 \times 10^{-8}$).
(TIFF)
S3 Fig. Quantile-quantile plot of observed against expected p-values for the Young Finns Study GWAS of sciatica. Results are adjusted for the first seven principal components of genetic data and for sex and age.
(TIF)
S4 Fig. Manhattan plot for the Health 2000 GWAS of sciatica. Results are adjusted for the first seven principal components of genetic data, and for sex and age.
(TIF)
S5 Fig. Quantile-quantile plot of observed against expected p-values for the Health 2000 GWAS of sciatica. Results are adjusted for the first seven principal components of genetic data, and for sex and age.
(TIF)
S1 Table. Study analysis methods. Methods used in genotyping, imputation, genome-wide association study and meta-analysis.
(DOCX)
S2 Table. Results of the genome-wide meta-analysis of sciatica showing variants with $p < 1 \times 10^{-5}$.
(DOCX)
S3 Table. Variants associated with sciatica in the meta-analysis (p < 1x10^{-6}) with effect allele frequencies, imputation quality and p-values in two discovery GWAS.

(DOCX)

S4 Table. LD estimates from the YFS (upper diagonal in green) and H2000 (lower diagonal in yellow) for SNPs with p<5x10^{-8} within the locus 15q21.2.

(DOCX)

S5 Table. Minor allele frequencies for the five most promising variants in the Finnish and other European populations.

(DOCX)

S6 Table. LD estimates from the YFS (upper diagonal in green) and H2000 (lower diagonal in yellow) for the SNPs in the HLA region within the locus 6p21.32.

(DOCX)

S7 Table. The most promising variants in the GWAS meta-analysis of sciatica (p<1x10^{-6}) as tested for replication in a Finnish population-based cohort (FINRISK).

(DOCX)

S8 Table. Genotype validations with three variants showing genome-wide significant association in the GWAS meta-analysis subsequently genotyped in individuals of the Young Finns Study and Health 2000 discovery cohorts.

(DOCX)

S1 Text. Study populations and phenotypes.

(DOCX)

S2 Text. Supplementary References.

(DOCX)

Acknowledgments

We thank all the Finnish volunteers who participated in the two cohort studies. We wish to thank Tuula Suitsiala, Specialized Laboratory Technician, for technical assistance, and Satu Hämäläinen, MSc, PhD student, for technical advice in the genotyping assays; both are from Finnish Institute of Occupational Health, Helsinki. Antti-Pekka Sarin, PhD student, Institute for Molecular Medicine Finland (FIMM), is thanked for part of the imputations. We are grateful to Dr. Mark Daly, the Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA, for critical reading of the manuscript, and to Dr. Josep M. Mercader, Barcelona Supercomputing Center, Barcelona, Spain, for advice in the analyses and comments on the revised manuscript. The personnel of the DNA Laboratory of the National Institute for Health and Welfare, Helsinki, Finland are thanked for skillful biobanking and handling of the samples used in this study. We thank Dr. Ewen MacDonald, University of Eastern Finland, Kuopio, Finland, for language revision. The Technology Center of the Institute for Molecular Medicine Finland kindly provided the computational facilities required in this study.

Author Contributions

Conceptualization: KHP SL SS EVJ.

Data curation: SL CB SS.

Formal analysis: SL CB.
Funding acquisition: KHP EVJ SS JK VA SR MK VS JV OTR TL AP.
Investigation: SL CB IS SS RS EVJ KHP.
Methodology: SL JK MP VA.
Project administration: KHP SL.
Resources: SS RS IS TL OTR MJ MK JV MH VS SR AP EVJ KHP.
Software: SL.
Supervision: KHP EVJ SS AP.
Validation: SL CB AP SR MH VS EVJ KHP.
Visualization: SL.
Writing – original draft: SL KHP SS EVJ RS AP.
Writing – review & editing: SL KHP SS EVJ RS CB MH JK VA SR MP IS MJ MK VS JV OTR TL AP.

References
1. Balague F, Mannion AF, Pellise F, Cedraschi C. Non-specific low back pain. Lancet. 2012; 379(9814):482–91. doi: 10.1016/S0140-6736(11)60610-7 PMID: 21982256
2. GBD Collaborators. Murray CJ, Barber RM, Foreman KJ, Abbasoglu Ozgoren A, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. Lancet. 2015; 386(10009):2145–91. doi: 10.1016/S0140-6736(15)61340-X PMID: 26321261
3. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. Spine. 2008; 33(22):2464–72. doi: 10.1097/BRS.0b013e318183a4a2 PMID: 18923325
4. Stafford MA, Peng P, Hill DA. Sciatica: a review of history, epidemiology, pathogenesis, and the role of epidural steroid injection in management. Br J Anaesth. 2007; 99(4):461–73. doi: 10.1093/bja/aem238 PMID: 17704089
5. Karppinen J. New Perspectives on Sciatica. In: De Leo JA, Sorkin LS, Watkins LR (eds). Immune and Glial Regulation of Pain. Seattle: IASP Press; 2007. p. 385–403.
6. Deyo RA, Loeser JD, Bigos SJ. Herniated lumbar intervertebral disk. Annals of internal medicine. 1990; 112(8):598–603. PMID: 2139310
7. Koes BW, van Tulder MW, Peul WC. Diagnosis and treatment of sciatica. Bmj. 2007; 334(7607):1313–7. doi: 10.1136/bmj.39223.428495.BE PMID: 17581610
8. Tarulli AW, Raynor EM. Lumbosacral radiculopathy. Neurologic clinics. 2007; 25(2):387–405. doi: 10.1016/j.ncl.2007.01.008 PMID: 17445735
9. Ropper AH, Zafonte RD. Sciatica. N Engl J Med. 2015; 372(13):1240–8. doi: 10.1056/NEJMra1410151 PMID: 25806916
10. Heikkila JK, Koskenvuo M, Heliovaara M, Kurppa K, Riihimaki H, Heikkila K, et al. Genetic and environmental factors in sciatica. Evidence from a nationwide panel of 9365 adult twin pairs. Ann Med. 1989; 21(5):393–8. PMID: 2605032
11. Battie MC, Videman T, Levalahli E, Gill K, Kaprio J. Heritability of low back pain and the role of disc degeneration. Pain. 2007; 131(3):272–80. doi: 10.1016/j.pain.2007.01.010 PMID: 17335977
12. Sambrook PN, Macgregor AJ, Spector TD. Genetic influences on cervical and lumbar disc degeneration: a magnetic resonance imaging study in twins. Arthritis Rheum. 1999; 42(2):366–72. doi: 10.1002/1529-0131(199902)42:2<366::AID-ANR20>3.0.CO;2-6 PMID: 10025932
13. Kalb S, Martirosyan NL, Kalani MY, Broc GG, Theodore N. Genetics of the degenerated intervertebral disc. World Neurosurg. 2012; 77(3–4):491–501. doi: 10.1016/j.wneu.2011.07.014 PMID: 22120330
14. Eskola PJ, Lemmela S, Kjaer P, Solovieva S, Mannikko M, Tommerup N, et al. Genetic association studies in lumbar disc degeneration: a systematic review. PLoS One. 2012; 7(11):e49995. doi: 10.1371/journal.pone.0049995 PMID: 23185509
15. Zhang Y, Sun Z, Liu J, Guo X. Advances in susceptibility genetics of intervertebral degenerative disc disease. Int J Biol Sci. 2008; 4(5):283–90. PMID: 18781226

16. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. An integrated map of genetic variation from 1,092 human genomes. Nature. 2012; 491(7422):56–65. doi: 10.1038/nature11632 PMID: 23128226

17. Surakka I, Isaacs A, Karssen LC, Laurila PP, Middelberg RP, Tikkannen E, et al. A genome-wide screen for interactions reveals a new locus on 4p15 modifying the effect of waist-to-hip ratio on total cholesterol. PLoS Genet. 2011; 7(10):e1002333. doi: 10.1371/journal.pgen.1002333 PMID: 22028671

18. Anttila V, Winsvold BS, Gormley P, Kurth T, Bettella F, McMahon G, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. Nat Genet. 2013; 45(8):912–7. doi: 10.1038/ng.2676 PMID: 23793025

19. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014; 46(11):1173–86. doi: 10.1038/ng.3097 PMID: 25282103

20. Kurki MI, Gaal EI, Kettunen J, Lappalainen T, Menelaou A, Anttila V, et al. High risk population isolate reveals low frequency variants predisposing to intracranial aneurysms. PLoS Genet. 2014; 10(1):e1004134. doi: 10.1371/journal.pgen.1004134 PMID: 24497844

21. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, et al. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet. 2015; 11(10):e1005378. doi: 10.1371/journal.pgen.1005378 PMID: 26426971

22. Kettunen J, Demirkan A, Wurtz P, Draiisma HH, Hailer T, Rawal R, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commucnic. 2016; 7:11122.

23. Kilpelainen TO, Carli JF, Skowronsni AA, Sun Q, Kriebel J, Feitosa MF, et al. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. Nat Commucnic. 2016; 7:10494.

24. Lu Y, Day FR, Gustafsson S, Buchkovick ML, Na J, Bataille V, et al. New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat Commun. 2016; 7:10495.

25. Singh T, Kurki MI, Curtis D, Purcell SM, Crooks L, McRae J, et al. Rare loss-of-function variants in SETD1A are associated with schizophrenia and developmental disorders. Nat Neurosci. 2016; 19(4):571–7. doi: 10.1038/nn.4267 PMID: 26974950

26. Williams FM, Bansal AT, van Meurs JB, Bell JT, Meulenbelt I, Suri P, et al. Novel genetic variants associated with lumbar disc degeneration in northern Europeans: a meta-analysis of 4600 subjects. Ann Rheum Dis. 2013; 72(7):1141–8. doi: 10.1136/annrheumdis-2012-201551 PMID: 22993228

27. Nakajima M, Takahashi A, Kour I, Rodriguez-Fontena C, Gomez-Reino JJ, Fujuchi T, et al. New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study. PLoS One. 2010; 5(3):e9723. doi: 10.1371/journal.pone.0009723 PMID: 20305777

28. Videman T, Saarela J, Kaprio J, Nakki A, Levalahti E, Gill K, et al. Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing. Arthritis Rheum. 2009; 60(2):470–81. doi: 10.1002/art.24268 PMID: 19180518

29. Noronen-Hietala N, Kylionen E, Mannikko M, Ilkko E, Karppinen J, Ott J, et al. Sequence variations in the collagen IX and XI genes are associated with degenerative lumbar spinal stenosis. Ann Rheum Dis. 2003; 62(12):1206–14. doi: 10.1136/ard.2003.008334 PMID: 14648461

30. Song YQ, Karasugi T, Cheung KM, Chiba K, Ho DW, Miyake A, et al. Lumbar disc degeneration is linked to a carbohydrate sulfotransferase 3 variant. The Journal of clinical investigation. 2013; 123(11):4909–17. doi: 10.1172/JCI69277 PMID: 24216480

31. Gronostajski RM. Roles of the NFI/CTF gene family in transcription and development. Gene. 2000; 249(1–2):31–45. PMID: 10831836

32. Pjanic M, Pjanic P, Schmid C, Ambrosini G, Gaussin A, Piasari G, et al. Nuclear factor I revealed as family of promoter binding transcription activators. BMC Genomics. 2011; 12:181. doi: 10.1186/1471-2164-12-181 PMID: 21473784

33. Qian F, Kruse U, Lichter P, Sippel AE. Chromosomal localization of the four genes (NFIA, B, C, and X) for the human transcription factor nuclear factor I by FISH. Genomics. 1995; 28(1):66–73. doi: 10.1006/geno.1995.1107 PMID: 7590749

34. Gronostajski RM, Guaneri J, Lee DH, Gallo SM. The NFI-Regulome Database: A tool for annotation and analysis of control regions of genes regulated by Nuclear Factor I transcription factors. J Clin Bioinforma. 2011; 1(1):4. doi: 10.1186/2043-9113-1-4 PMID: 21884625

35. Mason S, Piper M, Gronostajski RM, Richards LJ. Nuclear factor one transcription factors in CNS development. Mol Neurobiol. 2009; 39(1):10–23. doi: 10.1007/s12035-008-8048-6 PMID: 19058033
36. Steele-Perkins G, Plachez C, Butz KG, Yang G, Bachurski CJ, Kinsman SL, et al. The transcription factor gene Nfib is essential for both lung maturation and brain development. Mol Cell Biol. 2005; 25(2):685–98. doi: 10.1128/MCB.25.2.685-698.2005 PMID: 15632069

37. Xiao Z, Zhang S, Magenheimer BS, Luo J, Quarles LD. Polycystin-1 regulates skeletogenesis through stimulation of the osteoblast-specific transcription factor RUNX2-II. J Biol Chem. 2008; 283(18):12624–34. doi: 10.1074/jbc.M710407200 PMID: 18321855

38. Chang CY, Pasolli HA, Giannopoulou EG, Guasch G, Gronostajski RM, Elemento O, et al. NFIB is a governor of epithelial-melanocyte stem cell behaviour in a shared niche. Nature. 2013; 495(7439):98–102. doi: 10.1038/nature11847 PMID: 23389444

39. Uchihashi T, Kimata M, Tachikawa K, Koshimizu T, Okada T, Ihara-Watanabe M, et al. Involvement of nuclear factor I transcription/replication factor in the early stage of chondrocytic differentiation. Bone. 2007; 41(6):1025–35. doi: 10.1016/j.bone.2007.08.028 PMID: 17904922

40. Szabo P, Moitra J, Rencendorj A, Rakhely G, Rauch T, Kiss I. Identification of a nuclear factor-I family protein-binding site in the silencer region of the cartilage matrix protein gene. J Biol Chem. 1995; 270(17):10212–21. PMID: 7730325

41. Nagy A, Kenezi E, Rentsendorj O, Molnar A, Szenasi T, Sinko I, et al. Evolutionarily conserved, growth plate zone-specific regulation of the matrilin-1 promoter: L-Sox5/Sox6 and Nfi factors bound near TATA finely tune activation by Sox9. Mol Cell Biol. 2011; 31(4):686–99. doi: 10.1128/MCB.00019-10 PMID: 21173167

42. Trybus KM. Myosin V from head to tail. Cell Mol Life Sci. 2008; 65(9):1378–89. doi: 10.1007/s00018-008-7507-6 PMID: 18239852

43. Rodriguez OC, Cheney RE. Human myosin-Vc is a novel class V myosin expressed in epithelial cells. J Cell Sci. 2002; 115(Pt 5):991–1004. PMID: 11870218

44. Wu X, Bowers B, Rao K, Wei O, Hammer JA 3rd. Visualization of melanosome dynamics within wild-type and dilute melanocytes suggests a paradigm for myosin V function in vivo. J Cell Biol. 1998; 143(7):1899–918. PMID: 9864363

45. Rudolf R, Kogel T, Kuznetsov SA, Salm T, Schlicker O, Hellwig A, et al. Myosin Va facilitates the distribution of secretory granules in the F-actin rich cortex of PC12 cells. J Cell Sci. 2003; 116(Pt 7):1339–48. PMID: 12615975

46. Desnos C, Huet S, Darchen F. 'Should I stay or should I go?': myosin V function in organelle trafficking. Biocell. 2007; 99(8):411–23. doi: 10.1042/BC20070021 PMID: 17635110

47. Pastural E, Barrat FJ, Dufourcq-Lagelouse R, Certain S, Sanal O, Jabado N, et al. Griscelli disease maps to chromosome 15q21 and is associated with mutations in the myosin-Va gene. Nat Genet. 1997; 16(3):289–92. doi: 10.1038/ng0797-289 PMID: 9207796

48. Miyata M, Kishimoto Y, Tanaka M, Hashimoto K, Hirashima N, Murata Y, et al. A role for myosin Va in cerebellar plasticity and motor learning: a possible mechanism underlying neurological disorder in myosin Va disease. J Neurosci. 2011; 31(16):6067–78. doi: 10.1523/JNEUROSCI.5651-10.2011 PMID: 21508232

49. Van Gele M, Dynoodt P, Lambert J. Griscelli syndrome: a model system to study vesicular trafficking. Pigment Cell Melanoma Res. 2009; 22(3):268–82. doi: 10.1111/j.1755-148X.2009.00558.x PMID: 19243575

50. Shi D, Zheng Q, Chen D, Zhu L, Qin A, Fan J, et al. Association of single-nucleotide polymorphisms in HLA class II/III region with knee osteoarthritis. Osteoarthritis Cartilage. 2010; 18(11):1454–7. doi: 10.1016/j.joca.2010.07.009 PMID: 20691797

51. Valdes AM, Styrkarsdottir U, Doherty M, Morris DL, Mangino M, Tamm A, et al. Large scale replication study of the association between HLA class II/BTNL2 variants and osteoarthritis of the knee in European-descent populations. PLoS One. 2011; 6(8):e23371. doi: 10.1371/journal.pone.0023371 PMID: 21853121

52. Kang JD, Georgescu HI, McIntyre-Larkin L, Stefanovic-Racic M, Donaldson WF 3rd, Evans CH. Herniated lumbar intervertebral discs spontaneously produce matrix metalloproteinases, nitric oxide, interleukin-6, and prostaglandin E2. Spine. 1996; 21(3):271–7. PMID: 8742201

53. Korhonen T, Karppinen J, Paimela L, Malmivuo A, Lindgren KA, Bowman C, et al. The treatment of disc-herniation-induced sciatica with infliximab: one-year follow-up results of FIRST II, a randomized controlled trial. Spine. 2006; 31(24):2759–66. doi: 10.1097/01.brs.0000245873.23876.1e PMID: 17108825

54. Kim DH, Lee SH, Kim KT, Yu SD. Association of interleukin-1 receptor antagonist gene polymorphism with response to conservative treatment of lumbar herniated nucleus pulposus. Spine. 2010; 35(16):1527–31. doi: 10.1097/BRS.0b013e3181e4ef66 PMID: 20581747
55. Valdes AM, Hasset G, Hart DJ, Spector TD. Radiographic progression of lumbar spine disc degeneration is influenced by variation at inflammatory genes: a candidate SNP association study in the Chingford cohort. Spine (Phila Pa 1976). 2005; 30(21):2445–51.

56. Modic MT, Obuchowski NA, Ross JS, Brant-Zawadzki MN, Grooff PN, Mazanec DJ, et al. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. Radiology. 2005; 237(2):597–604. doi: 10.1148/radiol.2372041509 PMID: 16244269

57. Wassenaar M, van Rijn RM, van Tulder MW, Verhagen AP, van der Windt DA, Koes BW, et al. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic systematic review. Eur Spine J. 2012; 21(2):220–7. doi: 10.1007/s00586-011-2019-8 PMID: 21922287

58. Surakka I, Kristiansson K, Anttila V, Inouye M, Barnes C, Moutsianas L, et al. Founder population-specific HapMap panel increases power in GWA studies through improved imputation accuracy and CNV tagging. Genome Res. 2010; 20(10):1344–51. doi: 10.1101/gr.106534.110 PMID: 20810666

59. Consortium UK, Walter K, Min JL, Huang J, Crooks L, Memari Y, et al. The UK10K project identifies rare variants in health and disease. Nature. 2015; 526(7571):82–90. doi: 10.1038/nature14962 PMID: 26367797

60. Zeggini E. Using genetically isolated populations to understand the genomic basis of disease. Genome medicine. 2014; 6(10):83. doi: 10.1186/s13073-014-0083-5 PMID: 25473423

61. Varilainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol. 2010; 39(2):504–18. doi: 10.1093/ije/dyp330 PMID: 1999603

62. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009; 5(6):e1000529. doi: 10.1371/journal.pgen.1000529 PMID: 19543373

63. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007; 81(3):559–75. doi: 10.1086/519795 PMID: 17701901

64. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39(7):906–13. doi: 10.1038/ng2088 PMID: 17572673

65. Magi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics. 2010; 11:288. doi: 10.1186/1471-2105-11-288 PMID: 20509871

66. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003; 327(7414):557–60. doi: 10.1136/bmj.327.7414.557 PMID: 12958120

67. Devlin B, Roeder K. Genomic control for association studies. Biometrics. 1999; 55(4):997–1004. PMID: 11315092

68. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics. 2010; 26(18):2336–7. doi: 10.1093/bioinformatics/btp419 PMID: 20634204

69. Lemmela S, Forsman E, Onkamo P, Nurmi H, Laivuo H, Kivelä T, et al. Association of LOXL1 gene with Finnish exfoliation syndrome patients. Journal of human genetics. 2009; 54(5):289–97. doi: 10.1038/jhg.2009.28 PMID: 19343041