Warner, Sophie C. and van Meurs, Joyce B.J. and Schiphof, Dieuwke and Bierma-Zeinstra, Sita M. and Hofman, Albert and Uitterlinden, Andre G. and Richardson, Helen and Jenkins, Wendy and Doherty, Michael and Valdes, Ana M. (2017) Genome-wide association scan of neuropathic pain symptoms post total joint replacement highlights a variant in the protein-kinase C gene. European Journal of Human Genetics . ISSN 1476-5438

Genome-wide association scan of neuropathic pain symptoms post-total joint replacement highlights a variant in the protein-kinase C (*PRKCA*) gene

The genetics of neuropathic pain symptoms post-TJR

Sophie C Warner[1], Joyce BJ van Meurs[2], Dieuwke Schiphof[3], Sita M Bierma-Zeinstra[3,4], Albert Hofman[5], Andre G Uitterlinden[2,5], Helen Richardson[1], Wendy Jenkins[1], Michael Doherty[1,6], Ana M Valdes[1,6]

1 Academic Rheumatology, Clinical Sciences Building, Nottingham City Hospital, Nottingham, UK
2 Department of Internal Medicine, Erasmus MC, Rotterdam, Netherlands
3 Department of General Practice, Erasmus MC, Rotterdam, Netherlands
4 Department of Orthopaedics, Erasmus MC, Rotterdam, Netherlands
5 Department of Epidemiology, Erasmus MC, Rotterdam, Netherlands
6 Arthritis Research UK Pain Centre, Nottingham, UK

Corresponding author: Sophie C. Warner, Academic Rheumatology, Clinical Sciences Building, Nottingham City Hospital, Hucknall Road, Nottingham, NG5 1PB, UK. Telephone: +44 (0)115 823 1759. Fax: +44 (0)115 823 1757. Email: sophiecwarner@gmail.com

The authors declare no conflict of interest

Abstract

Neuropathic pain-like joint symptoms (NP) are seen in a proportion of individuals diagnosed with osteoarthritis (OA) and post-total joint replacement (TJR). In this study we performed a genome-wide association study (GWAS) using NP as defined by the painDETECT questionnaire (score >12 indicating possible NP) in 613 post-TJR participants recruited from Nottinghamshire (UK). The prevalence of possible NP was 17.8%. The top four hits from the GWAS and one other biologically relevant SNP were replicated in individuals with OA and post-TJR from an independent study in the same area (N=908) and in individuals from the Rotterdam Study (N=212). Three of these SNPs showed effect sizes in the same direction as in the GWAS results in both replication cohorts. The strongest association upon meta-analysis of a recessive model was for the variant allele in rs887797 mapping to the protein kinase C alpha (*PRKCA*) gene OR_{possNP}=2.41 (95%CI 1.74-3.34, p= $1.29x10^{-7}$). This SNP has been found to be associated with multiple sclerosis and encodes a functional variant affecting splicing and expression of the *PRKCA* gene. The *PRKCA* gene has been associated with long-term potentiation, synaptic plasticity, chronic pain and memory in the literature, making this a biologically relevant finding.

Keywords: Genome-wide association scan; neuropathic pain-like symptoms; neuropathic pain; osteoarthritis; total joint replacement; pain

Introduction

Neuropathic pain-like joint symptoms (NP) have been reported in people with osteoarthritis (OA) of the knee or hip and in some people who have undergone total-joint replacement (TJR) for OA (ref. 1, 2). Estimates for NP post-TJR range from 1% to 63% in the literature depending on the methodology (ref. 2, 3, 4).

Neuropathic pain is defined as "pain arising as a direct consequence of a lesion or disease affecting the somatosensory system", adapted from the International Association for the Study of Pain (IASP) definition (ref. 5). Symptoms can include burning, hypersensitivity, prickling and numbness in both the affected areas and areas of the body distant from the site of damage (ref. 6). Treatments for NP have been reported to be of limited effectiveness for many individuals and the condition can have a large impact on quality of life (ref. 7, 8). There are numerous risk factors for NP identified in the literature such as nerve damage from surgery, chronic nociceptive input (as seen in chronic pain), complications from herpes zoster infection and diabetes (ref. 9, 10). There are some common risk factors for OA pain and NP such as age, past joint surgery and psychological factors (ref. 1, 7, 11).

Heritability of NP has been estimated at 37% in the single published twin study on NP in humans (ref. 12). This is within the range of heritability estimates for other painful conditions such as back pain, migraine and sciatica which range from 21% to 58% (ref. 13, 14, 15, 16, 17).

There have been numerous candidate gene studies on pain, including chronic pain post-surgery (ref. 18). Genes reported in the literature on NP from candidate gene studies

include the *COMT* gene, *TRPV1* gene, *P2X* receptor genes and the *CACNG2* gene (ref. 19, 20, 21, 22). The genetics of NP are still not fully understood (ref. 23). NP is thought to have distinct genetic mechanisms, and different types of hypersensitivity (e.g. to heat or mechanical stimuli) and, according to mouse studies, different molecular mechanisms may be involved depending on the method for inducing NP (ref. 23).

A genome-wide association scan (GWAS) can be used to study the genetic basis of complex traits so is an appropriate design to study NP which can have a complex aetiology. GWAS identifies the genetic locations (single nucleotide polymorphisms; SNPs) which differ significantly between cases and controls for a specific phenotype. The genes in which these loci are located offer clues about the mechanisms behind the phenotype.

To date only one GWAS has been published on NP, in individuals with diabetic neuropathy. Results from this GWAS identified SNPs in the *GFRA2* and *ZSCAN20* genes (ref. 24, 25). Zinc finger proteins are potentially relevant in the treatment of NP (ref. 26). Previous GWAS for migraine and chronic widespread pain (CWP) have identified susceptibility loci relating to genes involved in synaptic plasticity and some types of neuropathy, respectively (ref. 27, 28). A GWAS has also been published on acute post-surgical pain (ref. 29).

The aim of this study was to identify genes associated with the risk of NP in individuals post-TJR using a genome-wide approach. The replication analysis aimed to reproduce these findings in other groups containing individuals with knee and hip OA and knee pain.

Methods

Participants

Nottingham discovery cohort: Participants were recruited post-total hip or knee replacement (n=613) from secondary care in the Nottinghamshire area (see **Figure S1, Supplements**).

Nottingham replication cohort: Participants from an independent Nottingham-based study (n=908) including individuals with knee OA, hip OA, or both and individuals post-total hip or knee replacement were used as a replication cohort (see **Figure S1, Supplements**).

The North Nottinghamshire Research Ethics Committee gave approval for the ethics of both studies. All participants gave written, informed consent.

To improve statistical power, in each of the above two Nottingham groups, total hip replacement participants and total knee replacement participants were combined into one post-TJR group, as seen in previous GWAs analyses

The Rotterdam Study: The selected individuals were part of Rotterdam Study III (RS-III) which was started in 2006 and comprised of in total 3 932 participants. A total of 212 women that reported knee pain had data on painDETECT and genetic data (see **Figure S1, Supplements**). This population-based cohort study has been previously described and is studied in the context of chronic disabling diseases in older adults (ref. 30). The Erasmus University Medical School medical ethics committee gave approval for this study. All participants gave written, informed consent.

Stage 1: GWAS

Blood samples from the participants in this study were processed to obtain genotype data. Genotype data was analysed using the Illumina 610k array (https://www.ebi.ac.uk/ega/studies/EGAS00001001017). Only directly typed SNPs were used. Genotyping and QC were carried out as previously described (ref. 31), gPLINK software (version 1.07) was used to analyse GWAS data from this array (ref. 32). The results of this association are a list of genetic variants (SNPs) and information about their location in the genome, as well as an odds ratio (OR), chi square value, and p value to indicate the level of association of the variants with the specified phenotype. The statistics program R (version 3.0.2) was used to create Manhattan and QQ plots using the "ggplot2" library and "qqplot" script.

Post-genomic analysis was undertaken using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (ref. 33). This is an online tool to which a list of genes can be submitted and subsequently results are generated regarding the genes' involvement in biological processes (ref. 33). The gene list was comprised of genes corresponding to all SNPs with a p value of $p<0.0001$ in the GWAS analysis. The BioCarta and Kegg pathways maps were used for functional annotation.

Stage 2: replication cohorts

Five SNPs with a nominal p value of $p<10^{-4}$ after the stage 1 GWAS analysis and one additional lower-ranking but potentially relevant SNP were selected for replication (rs1133076; see **Discussion**). Genotype information for these SNPs from in silico and de

novo genotype data were used for further analysis. In total six SNPs were selected for replication analysis.

Stage 3: meta-analysis

The "meta" library in the statistics program R (version 3.0.2) was used to run the meta-analysis using the four cohorts described above. Meta-analysis takes the effect size, standard error and sample size into account to give an overall effect from the different groups studied. If heterogeneity was significant between the cohorts in the meta-analysis, a Han Eskin random effects model was used as an alternative meta-analysis method as, compared to traditional models, it allows for more heterogeneity in the data (ref. 34).

Phenotype

Individuals were assigned a phenotype by classifying them according to their scores on the painDETECT questionnaire. This is a seven-item questionnaire scored from 0-39 which uses a Likert scale for participants to describe the nature of their pain, in order to distinguish hit from nociceptive pain. Questions are included on qualities such as burning pain, tingling, sudden pain and sensitivity to heat and cold. In all cohorts, scores of >12 were classified as "possible neuropathic pain" as described by Freynhagen et al. (ref. 35).

Results

Stage 1: GWAS

The results of the unadjusted GWAS on NP can be seen in **Table 1** and **Figure 2**. A total of 548 382 SNPs were tested for association with NP. The genomic control inflation factor

for the p values was low (λ=0.99) and the quantile-quantile (QQ) plot indicated no substantial population stratification due to cryptic relatedness, population substructure or other biases (Figure 2).

The results of the GWAS are summarised in Manhattan plots of the p values (Figure 3). **Table 1** shows the odds ratios (OR) and significance of the results from the Illumina array NP GWAS for five of the top-scoring SNPs and a SNP of biological relevance. The highest-scoring SNP was rs887797 in the protein kinase C (*PRKCA*) gene: OR=2.04 (1.51-2.77), $p=3.76x10^{-6}$.

Pathway analysis: Pathway analysis was carried out on the GWAS results using a list of genes corresponding to SNPs with p values less than $p<0.0001$ in the GWAS (n=62; see **Table S1, Supplements**). If the SNP mapped to an area within a gene, this gene was used. For intergenic SNPs the two closest flanking genes on each side were used. The results of this analysis (see **Table S2, Supplements**) report no significant findings after adjusting for multiple testing with a Bonferroni correction (see **Table S2, Supplements**).

Stage 2: Replication cohorts

We sought to replicate the 6 selected SNPs for their association with NP in two independent replication cohorts. The results are shown in **Table 1**. As shown in **Table 1**, two of the SNPs selected from the GWAS in stage 1 for replication analysis show significant p values and effects in the same direction in one of the replication cohorts.

Stage 3: meta-analysis

We then combined discovery and replication results in a joint meta-analysis. The results can be seen in **Table 1**. Heterogeneity of the loci was tested using the Cochran Q test.

Due to the significant heterogeneity introduced to the model by the replication data in the rs887797, rs4866176, rs7734804, rs298235 and rs12596162 meta-analyses, a Han Eskin random effects model was used to account for this (see **Table 1**). The additive model for the rs887797 SNP after this analysis gave a result of: OR=1.47 (95% CI 1.24-1.76), p=1.33 $x10^{-5}$. A recessive model for the rs887797 SNP was also used in a meta-analysis. A recessive model was used to test the nature of the effect of the risk allele, that is, to test if two copies of the risk allele were needed to increase the risk of possible NP. After Han Eskin analysis, the recessive model for rs887797 gave a result of OR=2.41 (95% CI 1.74-3.34, $p=1.29x10^{-7}$) (see **Figure 4**).

After adjusting for age, sex and BMI, Han Eskin analysis of the rs887797 SNP gave values of: OR_{possNP}=1.44 (95% CI 1.21-1.73, $p=7.13x10^{-5}$) and OR_{possNP}=2.33 (95% CI 1.67-3.27, $p=8.67x10^{-7}$) for the additive and recessive models, respectively. Upon combining the data from the two replication cohorts used, it was found that overall this SNP was significant. The additive model for the rs887797 SNP in the Nottingham replication cohort and Rotterdam Study cohort gave OR=1.25 (95% CI 1.01-1.55), p=0.040 and the recessive model gave OR=1.75 (95% CI 1.15-2.64), p=0.008.

Finally, we attempted to replicate two of the top hits from the only published GWAS on NP. These SNPs were reported to be suggestively associated with diabetic neuropathy: rs17428041 (*GFRA2*, OR=0.67, $p=1.77x10^{-7}$) (ref. 24) and rs71647933 (*ZSCAN20*,

OR=2.31, $p=4.88x10^{-7}$) (ref. 25). The effect of rs17428041 was not replicated in the results of our GWAS: OR=1.47, p=0.016. Similarly, after using a proxy for rs71647933 (rs12565140, $r^2=0.947$) we found no association with NP in the results of our GWAS: OR=0.71 (95% CI 0.46-1.09), p=0.12).

Discussion

We report that a variant in the *PRKCA* gene is associated with NP in people with knee pain, knee or hip OA and post-TJR. Despite not reaching genome-wide significance (GWS) the replication of effect sizes for four of these SNPs in one or both of the replication cohorts, and the improvement of the p value for one of these SNPs after meta-analysis suggest that these are true associations. The findings are also biologically plausible and supported by previously published work in the literature. We were unable to confirm the recently published association between SNPs in the *GFRA2* and *ZSCAN20* genes and diabetic neuropathy (ref. 24, 25). However, it should be noted that diabetic neuropathy is not necessarily the same phenotype as neuropathic pain-like joint symptoms. The definition of NP used in these studies was partly based on use of prescription analgesic medication and partly on the results of sensory testing. However, this type of medication is commonly used even by people with no NP, including people post-TJR with no NP. In our study, a validated screening questionnaire (painDETECT) was used, the location of pain is exclusively that of the OA-affected joint and further clinical history and demographics have been collected for all participants.

The top hit from our GWAS and replication analysis maps to the *PRKCA* gene. This gene codes for protein kinase C alpha, a protein which has been linked with the nervous system and may contribute to central sensitisation in dorsal horn neurons (ref. 36). The *PRKCA* gene has also been found in the literature to be involved in long-term potentiation (LTP), a process involved in both memory and chronic pain (ref. 37). As well as this, the *PRKCA* gene has been implicated in related processes such as memory capacity and post-traumatic stress disorder (PTSD) (ref. 38) and genetic variation in this gene has been linked to the neural basis of episodic memory (ref. 39). Although we do not reach the $p<5x10^{-8}$ threshold for GWS, we show plausible, reproducible genetic effects on NP post-TJR and after replication analysis. The National Human Genome Research Institute (NHGRI) keeps a record of all SNP-trait associations $p<10^{-5}$ (ref. 40) which supports the relevance of the findings in this study and their suggestive role in NP, despite not achieving GWS. More importantly if we combine the data from the two replication cohorts used we still achieve a significant p-value. A role for the *PRKCA* gene in pain has been previously reported (ref. 41). The rs887797 variant identified in this paper is a variant already associated with multiple sclerosis (ref. 42). Therefore, although this association may not reach GWS it remains highly biologically plausible.

In the present GWAS the intergenic rs12596162 SNP near the *FOXL1* gene was associated with NP: OR=1.96 (95% CI 1.45-2.64), $p=1.09x10^{-5}$. This gene codes for a forkhead/winged helix-box transcription factor (ref. 43). This gene and the rest of the *FOX* gene family are involved in many cellular processes (ref. 43). *FOXL1* in particular was

found in one study to be involved in the Wnt/β-catenin pathway (ref. 44) which is important in the nervous system and has been implicated in NP and hip OA (ref. 45, 46).

Thyroglobulin, encoded by the *TG* gene, is a protein necessary for normal thyroid function which has previously been related to NP and central sensitisation in the literature (ref. 47). The rs1133076 SNP mapping to this gene was suggested in this analysis to be associated with possible NP at the discovery stage with $p=8.09 \times 10^{-4}$. However this variant did not replicate in the additional cohorts and the evidence for association with NP for this gene is very weak.

The effect sizes we report here are larger than those reported in previous GWAS on pain traits such as migraine and CWP (OR=1.18 and OR=1.23, respectively) (ref. 27, 28) despite our study having a smaller sample size. The effect size for the GWAS on NP in diabetes was 2.31 for the SNP with the lowest p value, which is consistent with our finding for the rs887797 SNP in the GWAS analysis (OR=2.05, see **Table 1**).

There are a number of limitations to this study. None of the variants identified by this study reaches GWS. This is not surprising given the small discovery and replication sample sizes available for this kind of study. A major issue with the use of GWAS is the potential for inflated associations (ref. 48). The statistical power for the rs887797 recessive model with the observed OR=2.41 was 56% for GWS. For the observed p value the statistical power was 66% given the observed minor allele frequency and the rare homozygote frequency (which is in HWE). Although the study was underpowered for GWS, the effect size is relatively large. To achieve 80% power with this effect size and the same proportion of

cases to controls we would have needed 417 cases and 1 767 controls, a 25% larger sample size, assuming that in the additional sample the effect was the same (ref. 48). Only the most extreme p values and effect sizes are selected for further study after a GWAS (ref. 49). This is called the "winner's curse" (ref. 49) and means that the effect size reported here is likely to be an overestimate given the small sample size used for the discovery phase, and sample sizes of at least twice those that were used are likely to be needed. Furthermore, heterogeneity between the groups used in the meta-analysis can limit the effects seen in the results though we attempted to address this by the use of a Han Eskin Random Effects analysis (ref. 34).

The absence of a clinical NP diagnosis in these participants is another limitation of this study. However the results of this questionnaire have been shown to correlate with brain activity in areas associated with NP in people with NP and OA (ref. 50).

In summary, this study has found biologically plausible and reproducible genetic effects when analysing possible NP in individuals with knee pain, OA and post-TJR. Replication in further cohorts could improve sample size and p values and it is hoped that this GWAS of neuropathic pain-like symptoms of the joint may encourage the collection of DNA and of painDETECT and similar instruments in other cohorts.

Acknowledgements

SCW is funded by a PhD studentship awarded by the University of Nottingham. This work was supported by a EULAR project grant to AMV (grant 108239) and by the Arthritis Research UK Pain Centre (grant 18769). The authors gratefully acknowledge the contributions of Sally Doherty and Maggie Wheeler to patient assessments at baseline, data

collection and entry. JM was funded by The Netherlands Society for Scientific Research (NWO) VIDI Grant 917103521. The generation and management of GWAS genotype data for the Rotterdam Study (RS I, RS II, RS III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, M.Sc., and Carolina Medina-Gomez, M.Sc., for their help in creating the GWAS database, and Karol Estrada, Ph.D., Yurii Aulchenko, Ph.D., and Carolina Medina-Gomez, M.Sc., for the creation and analysis of imputed data. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

References

1. Hochman, J.R., Gagliese, L., Davis, A. M., Hawker, G. A., Neuropathic pain symptoms in a community knee OA cohort. *Osteoarthritis Cartilage*, 2011. **19**(6): p. 647-54.

2. Wylde, V., S. Hewlett, I.D. Learmonth, and P. Dieppe, Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. *Pain*, 2011. **152**(3): p. 566-72.

3. Buvanendran, A., J.S. Kroin, C.J. Della Valle, et al., Perioperative oral pregabalin reduces chronic pain after total knee arthroplasty: a prospective, randomized, controlled trial. *Anesthesia & Analgesia*, 2010. **110**(1): p. 199-207.

4. Haroutiunian, S., L. Nikolajsen, N.B. Finnerup, and T.S. Jensen, The neuropathic component in persistent postsurgical pain: a systematic literature review. *Pain*, 2013. **154**(1): p. 95-102.

5. Treede, R.-D., T.S. Jensen, J. Campbell, et al., Neuropathic pain redefinition and a grading system for clinical and research purposes. *Neurology*, 2008. **70**(18): p. 1630-1635.

6. Kehlet, H., Jensen, T.S., Woolf, C.J., Persistent postsurgical pain: risk factors and prevention. *Lancet*, 2006. **367**: p. 1618-1625.

7. Valdes A.M., S., A.K., Doherty, S.A., Jenkins, W., Doherty, M., History of knee surgery is associated with higher prevalence of neuropathic pain-like symptoms in patients with severe osteoarthritis of the knee. *Seminars in Arthritis and Rheumatism*, 2013.

8. van Hecke, O., Austin, S. K., Khan, R. A., Smith, B. H., Torrance, N., Neuropathic pain in the general population: a systematic review of epidemiological studies. *Pain*, 2014. **155**(4): p. 654-62.

9. Dualé, C., L. Ouchchane, P. Schoeffler, et al., Neuropathic Aspects of Persistent Postsurgical Pain: A French Multicenter Survey With a 6-Month Prospective Follow-Up. *The Journal of Pain*, 2014. **15**(1): p. 24.e1-24.e20.

10. Graven-Nielsen, T., T. Wodehouse, R.M. Langford, L. Arendt-Nielsen, and B.L. Kidd, Normalization of widespread hyperesthesia and facilitated spatial summation of deep-tissue pain in knee osteoarthritis patients after knee replacement. *Arthritis & Rheumatism*, 2012. **64**(9): p. 2907-2916.

11. Novak, J.C., Lovell, J. A., Stuesse, S. L., Cruce, W. L. R., McBurney, D. L., Crisp, T., Aging and neuropathic pain. *Brain Research*, 1999. **833**(2): p. 308-310.

12. Momi, S.K., S.M. Fabiane, G. Lachance, G. Livshits, and F.M. Williams, Neuropathic pain as part of chronic widespread pain: environmental and genetic influences. *Pain*, 2015. **156**(10): p. 2100-2106.

13. Honkasalo, M.L., J. Kaprio, T. Winter, et al., Migraine and concomitant symptoms among 8167 adult twin pairs. *Headache: The Journal of Head and Face Pain*, 1995. **35**(2): p. 70-78.

14. Ziegler, D.K., Y.M. Hur, T.J. Bouchard, R.S. Hassanein, and R. Barter, Migraine in twins raised together and apart. *Headache: The Journal of Head and Face Pain*, 1998. **38**(6): p. 417-422.

15. Heikkila, J.K., M. Koskenvuo, M. Heliovaara, et al., Genetic and environmental factors in sciatica. Evidence from a nationwide panel of 9365 adult twin pairs. *Ann Med*, 1989. **21**(5): p. 393-8.

16. Larsson, B., B. Bille, and N.L. Pedersen, Genetic influence in headaches: a Swedish twin study. *Headache: The Journal of Head and Face Pain*, 1995. **35**(9): p. 513-519.

17. Bengtsson, B. and J. Thorson, Back pain: a study of twins. *Acta geneticae medicae et gemellologiae: twin research*, 1991. **40**(01): p. 83-90.

18. Clarke, H., J. Katz, H. Flor, et al., Genetics of chronic post-surgical pain: a crucial step toward personal pain medicine. *Canadian Journal of Anesthesia/Journal canadien d'anesthésie*, 2015. **62**(3): p. 294-303.

19. Belfer, I., H. Shnol, and P. Finelli, Molecular Genetics of Variability in Human Pain, in *eLS*. 2013, John Wiley & Sons, Ltd.

20. Nissenbaum, J., M. Devor, Z.e. Seltzer, et al., Susceptibility to chronic pain following nerve injury is genetically affected by CACNG2. *Genome Research*, 2010. **20**(9): p. 1180-1190.

21. Tsuda, M., Kuboyama, K, Inoue, T, Nagata, K, Tozaki-Saitoh, H, Inoue, K, Behavioral phenotypes of mice lacking purinergic P2X4 receptors in acute and chronic pain assays. *Molecular Pain*, 2009. **5**(28).

22. Valdes, A.M., G. De Wilde, S.A. Doherty, et al., The Ile585Val TRPV1 variant is involved in risk of painful knee osteoarthritis. *Annals of the Rheumatic Diseases*, 2011. **70**(9): p. 1556-1561.

23. Young, E.E., M. Costigan, T.A. Herbert, and W.R. Lariviere, Heritability of Nociception IV: Neuropathic pain assays are genetically distinct across methods of peripheral nerve injury. *Pain*, 2014. **155**(5): p. 868-880.

24. Meng, W., H.A. Deshmukh, N.R. van Zuydam, et al., A genome-wide association study suggests an association of Chr8p21.3 (GFRA2) with diabetic neuropathic pain. *Eur J Pain*, 2015. **19**(3): p. 392-9.

355 25. Meng, W., H.A. Deshmukh, L.A. Donnelly, et al., A Genome-wide Association Study
356 Provides Evidence of Sex-specific Involvement of Chr1p35.1 (ZSCAN20-TLR12P) and
357 Chr8p23.1 (HMGB1P46) With Diabetic Neuropathic Pain. *EBioMedicine*, 2015. **2**(10): p.
358 1386-1393.
359 26. Krishna, S.S., I. Majumdar, and N.V. Grishin, Structural classification of zinc fingers:
360 SURVEY AND SUMMARY. *Nucleic Acids Research*, 2003. **31**(2): p. 532-550.
361 27. Anttila, V., Stefansson, H., Kallela, K., Unda Todt, U., Terwindt, G.M., Calafato, M.S.,
362 et al., Genome-wide association study of migraine implicates a common susceptibility
363 variant on 8q22.1. *Nat Genet*, 2010. **42**(10): p. 869-873.
364 28. Peters, M.J., L. Broer, H.L.D.M. Willemen, et al., Genome-wide association study
365 meta-analysis of chronic widespread pain: evidence for involvement of the 5p15.2 region.
366 *Annals of the Rheumatic Diseases*, 2013. **72**(3): p. 427-436.
367 29. Kim, H., E. Ramsay, H. Lee, S. Wahl, and R.A. Dionne, Genome-wide association
368 study of acute post-surgical pain in humans. *Pharmacogenomics*, 2009. **10**(2): p. 171-179.
369 30. Hofman, A., M.B. Breteler, C. van Duijn, et al., The Rotterdam Study: objectives and
370 design update. *European Journal of Epidemiology*, 2007. **22**(11): p. 819-829.
371 31. Zeggini, E., Panoutsopoulou, Kalliope,; Southam, Lorraine; Rayner, Nigel W; Day-
372 Williams, Aaron G; Lopes, Margarida C, et al., Identification of new susceptibility loci for
373 osteoarthritis (arcOGEN): a genome-wide association study. *Lancet*, 2012.
374 32. Purcell S., N.B., Todd-Brown K., Thomas L., Ferreira M.A.R., Bender D., Maller J.,
375 Sklar P., de Bakker P.I.W., Daly M.J., Sham P.C., PLINK: a toolset for whole-genome
376 association and population-based linkage analysis. 2007, American Journal of Human
377 Genetics: Harvard University, Cambridge, MA, USA.

33. Huang da, W., B.T. Sherman, and R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 2009. **4**(1): p. 44-57.

34. Han, B. and E. Eskin, Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *The American Journal of Human Genetics*, 2011. **88**(5): p. 586-598.

35. Freynhagen, R., Baron, R., Gockel, U., Tolle, T. R., painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Curr Med Res Opin*, 2006. **22**(10): p. 1911-20.

36. Kawasaki, Y., Kohno, T., Zhuang, Z.Y., Brenner, G.J., Wang, H., Van Der Meer, C., Befort, K., Woolf, C.J, Ji, R.R., Ionotropic and Metabotropic Receptors, Protein Kinase A, Protein Kinase C, and Src Contribute to C-Fiber-Induced ERK Activation and cAMP Response Element-Binding Protein Phosphorylation in Dorsal Horn Neurons, Leading to Central Sensitization. *J Neurosci*, 2004. **24**(38): p. 8310-8321.

37. Price, T.J. and K.E. Inyang, Commonalities between pain and memory mechanisms and their meaning for understanding chronic pain. *Prog Mol Biol Transl Sci*, 2015. **131**: p. 409-34.

38. de Quervain, D.J.F., I.-T. Kolassa, S. Ackermann, et al., PKCα is genetically linked to memory capacity in healthy subjects and to risk for posttraumatic stress disorder in genocide survivors. *Proceedings of the National Academy of Sciences of the United States of America*, 2012. **109**(22): p. 8746-8751.

39. MacLeod, C.A. and D.I. Donaldson, PRKCA Polymorphism Changes the Neural Basis of Episodic Remembering in Healthy Individuals. *PLoS ONE*, 2014. **9**(5): p. e98018.

40. Welter, D., J. MacArthur, J. Morales, et al., The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, 2014. **42**(Database issue): p. D1001-D1006.

41. Olah, Z., L. Karai, and M.J. Iadarola, Protein Kinase Cα Is Required for Vanilloid Receptor 1 Activation: EVIDENCE FOR MULTIPLE SIGNALING PATHWAYS. *Journal of Biological Chemistry*, 2002. **277**(38): p. 35752-35759.

42. Paraboschi, E.M., V. Rimoldi, G. Soldà, et al., Functional variations modulating PRKCA expression and alternative splicing predispose to multiple sclerosis. *Human Molecular Genetics*, 2014. **23**(25): p. 6746-6761.

43. NCBI. FOXL1 forkhead box L1 [Homo sapiens (human)]. *Gene* 2014; Available from: http://www.ncbi.nlm.nih.gov/gene/2300.

44. Jiang, D., Hwang, K.S., Bordelon, Y., Apostolova, L.G., Plenary Paper - Cortical Atrophy and Gene Expression in Parkinson's Disease with Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, 2013. **61**: p. S3.

45. Zhang, Y.-K., Z.-J. Huang, S. Liu, et al., WNT signaling underlies the pathogenesis of neuropathic pain in rodents. *The Journal of Clinical Investigation*, 2013. **123**(5): p. 2268-2286.

46. Castaño Betancourt, M.C., F. Cailotto, H.J. Kerkhof, et al., Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. *Proceedings of the National Academy of Sciences*, 2012. **109**(21): p. 8218-8223.

47. Penza, P., R. Lombardi, F. Camozzi, C. Ciano, and G. Lauria, Painful neuropathy in subclinical hypothyroidism: clinical and neuropathological recovery after hormone replacement therapy. *Neurol Sci*, 2009. **30**(2): p. 149-51.

48. Ioannidis, J.P., Why most discovered true associations are inflated. *Epidemiology*, 2008. **19**(5): p. 640-648.

49. Garner, C., Upward bias in odds ratio estimates from genome-wide association studies. *Genetic Epidemiology*, 2007. **31**(4): p. 288-295.

50. Gwilym, S.E., Keltner, J.R., Warnaby, C.E., Carr, A.J., Chizh, B., Chessell, I., Tracey, I., Psychophysical and functional imaging evidence supporting the presence of central sensitization in a cohort of osteoarthritis patients. *Arthritis Care Res*, 2009. **61**(9): p. 1226-1234.

Tables and Figures

							Stage 1 (GWAS)		Stage 2a (Rotterdam)	Stage 2b (Nottingham replication)	Stage 3 (Meta-analysis)		
SNP ID	Description	Chr	BP	MAF	Effect allele/No effect	Gene or closest gene	GWAS p value	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	I^2	p value†
rs887797	hg19 chr17:g.64579445G>A	17	64,579,445	0.335	A/G	*PRKCA*	$4.29x10^{-6}$	2.00 (1.48-2.70)	1.12 (0.64-1.95)	**1.28 (1.02-1.61)**	1.48 (1.23-1.75)	71.3%	$1.65x10^{-5}$
rs4866176	hg19 chr5:g.20245445C>T	5	20,245,554	0.064	A/G	*CDH18*	$1.19x10^{-5}$	2.86 (1.76-4.66)	1.12 (0.36-3.51)	0.88 (0.54-1.43)	1.52 (1.08-2.12)	81.8%	$1.39x10^{-3}$
rs1133076*	hg19 chr8:g.134125682G>A	8	134,125,682	0.469	A/G	*TG*	$8.09x10^{-4}$	1.66 (1.23-2.24)	1.52 (0.88-2.64)	1.18 (0.94-1.48)	1.35 (1.14-1.60)	42.7%	$5.45x10^{-4}$
rs7734804	hg19 chr5:g.164919530G>T	5	164,346,536	0.025	A/C	*MAT2B* (16.09 kbp) | *ODZ2* (690.20 kbp)	$5.25x10^{-6}$	4.64 (2.26-9.53)	**12.92 (1.13-147.20)**	1.50 (0.77-2.91)	2.61 (1.62-4.22)	68.4%	$7.80x10^{-5}$
rs298235	hg19 chr2:g.157306688A>G	2	157,306,688	0.016	A/G	*GPD2*	$3.41x10^{-5}$	6.72 (2.67-16.92)		1.12 (0.41-3.08)	2.97 (1.51-5.93)	85.2%	$5.32x10^{-4}$
rs12596162	hg19 chr16:g.87117889C>T	16	87,151,495	0.303	A/G	*FOXL1* (536.19 kbp) | *C16orf95* (184.91 kbp)	$3.53x10^{-5}$	2.05 (1.51-2.79)	1.68 (0.93-3.03)	0.87 (0.67-1.13)	1.26 (1.05-1.52)	88.5%	$2.80x10^{-4}$

Chr=chromosome, BP=nucleotide location, MAF=minor allele frequency

*indicates a SNP within a gene of biological relevance and interest which was hypothesised to be associated with NP

†if the heterogeneity (I^2) between the groups was significant, a Han Eskin random effects model was used to calculate this value

Table 1: The results of interest from the unadjusted Illumina array NP GWAS, followed by the results of replication analysis and meta-analysis

Figure 1: Study design

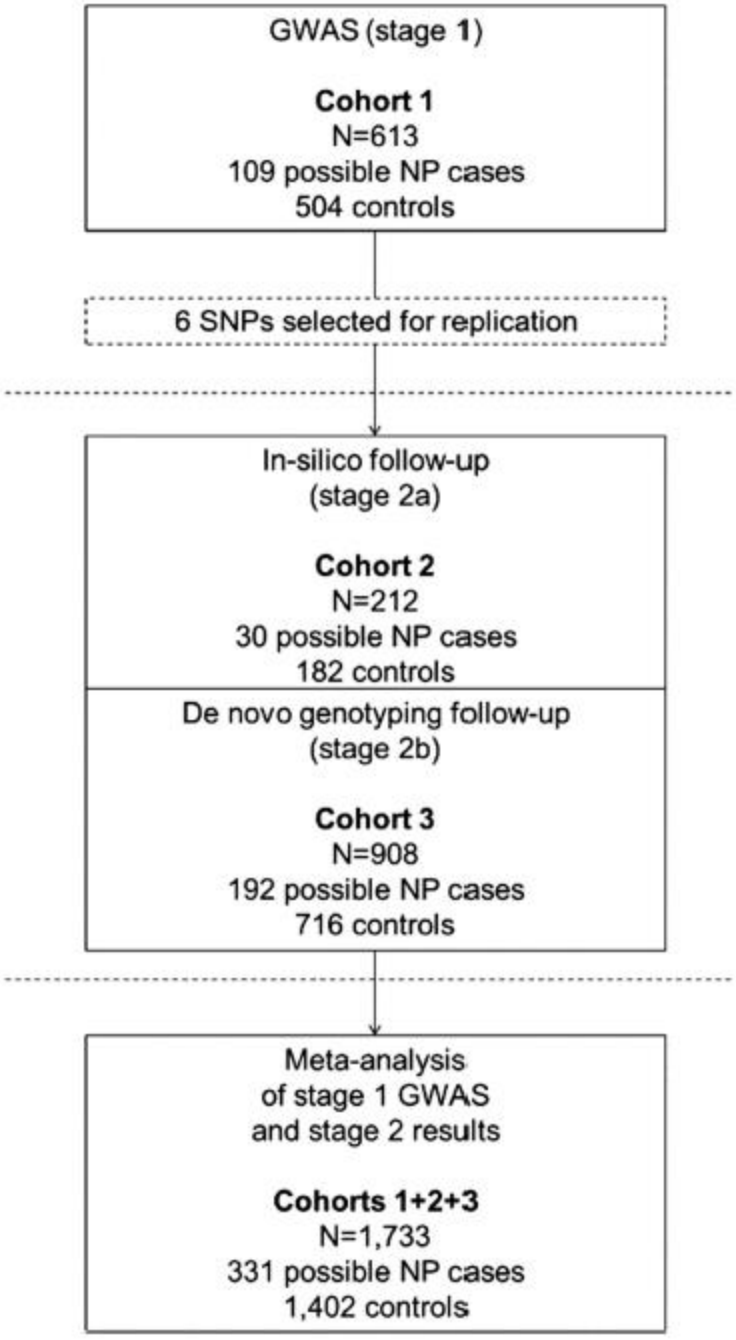

Figure 2: QQ plot for the results of the GWAS (λ=0.99)

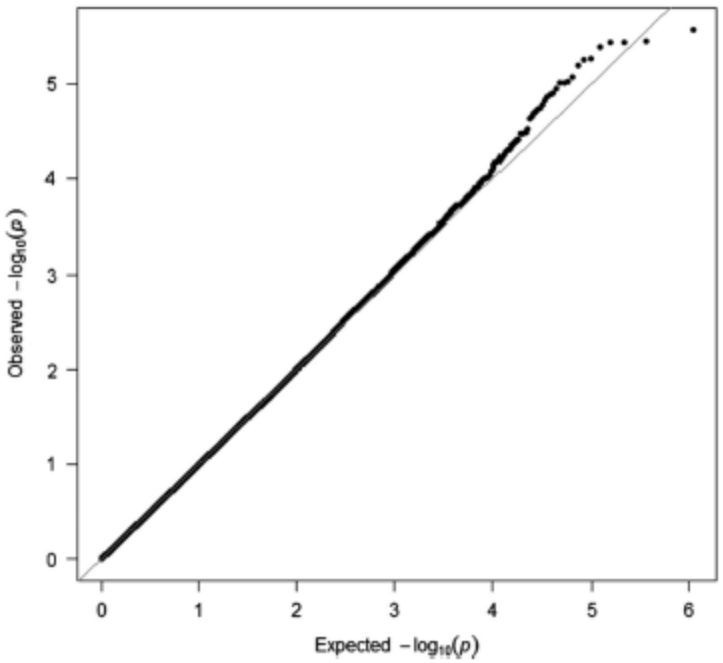

Figure 3: Manhattan plot showing the p value of association tests for SNPs with possible NP in the Illumina array GWAS. P values represent the association of the SNPs with possible NP

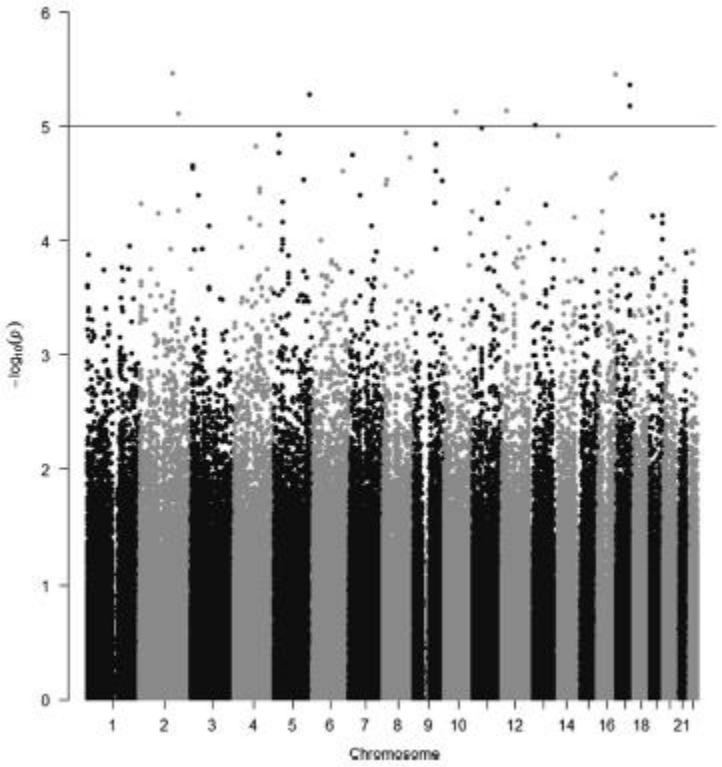

Figure 4: Forest plot showing the results of an unadjusted Han Eskin analysis on the rs887797 SNP using a recessive model