EXECUTIVE SUMMARY

PAIN-OMICS is a multidisciplinary consortium comprising leading clinical, academic and SME researchers in pain and different omics technologies. The burden of chronic pain and chronic pain syndromes in the EU is very large, through it affecting adults of working age. While acute pain is an important part of human physiology, leading to protection from injury, chronic pain appears largely maladaptive. This consortium set out to explore the mechanisms leading from acute to chronic pain by examining a range of biomarkers. A small number of genetic loci have been associated with chronic pain, but the level of knowledge about underlying mechanisms of different pain syndromes as well as individual variation in the conditions remains inadequate. PainOMICS capitalised on its existing high quality clinical, genetic, biochemical and pharmacological data and biological samples collected on over 2500 patients with low-back pain (LBP) and controls available to our EU and US clinical partners. In addition, with the availability of large consortia having huge genetic samples with phenotype information on chronic pain, we were able to perform the largest to date genetic study of back pain. Novel technological approaches have been exploited by the consortium to understand the mechanism LBP underlying the transformation of acute LBP to chronic LBP. These comprise cutting edge genomic, glycomic, and proteomic approaches which reflect signal transduction and membrane dynamics.

These approaches helped identify pathways and biomarkers of chronic pain through which individual differences affects symptoms and response to therapy. A complex system biology approach has been used to integrate, interrogate and understand this multidimensional dataset in order to achieve the aims of identifying novel diagnostic and prognostic bio-markers as well as new targets for therapeutic intervention. Our results point towards a role for chronic inflammation in the development of chronic LBP – something that was not anticipated and may offer a role for novel therapies. In addition, the genetic studies have identified three novel genes associated with back pain which point to the role of structural spine genes. In keeping with other common complex traits, we believe that our findings from the largest ever back pain genome-wide association study will be put to best use if combined with existing strategies (such as questionnaires) which stratify patients with back pain into low and high risk groups for chronic pain. The next step will be the study of such a combination approach and if successful a clinical trial of its implementation.

PROJECT CONTEXT AND THE MAIN OBJECTIVES

Acute low back pain is one of the most common reasons for adults to consult a doctor or other healthcare practitioner and the majority of people will experience back pain at some point in their life. Although most patients recover quickly with minimal intervention, about 10-15% of these patients develop chronic symptoms (defined variously but often as pain persisting 3 months or more). In Europe, more than 40% of adults suffer from at least one episode of low back pain (LBP).

The major concerns about CLBP syndrome are:

- lack of exact knowledge of its complex pathophysiology;
- lack of biomarkers that predict risk of developing this syndrome;
- lack of biomarkers and imaging data that could help interpret clinical symptoms and pain intensity;
- lack of biomarkers and clinical data that could help to predict response to interventions and to intervention adverse effects.

Low-back pain is a diverse group of mixed pain syndromes (neuropathic and nociceptive) with different molecular pathologies at a structural level displaying similar clinical manifestations. Despite many decades of research to identify precise anatomical triggers (disc degeneration, facet joint, muscle) there has been little progress in diagnosis and prognosis in the vast majority of patients. The adoption of the term 'simple mechanical back pain' as a catch-all diagnosis has not reassured patients. In fact, the Global burden of disease (Lancet 2016) has placed back pain as the leading cause of disability worldwide, and a recent Lancet back pain series has implicated our current systems of healthcare in actually making the problem worse.

The aim of the current project was to move towards an understanding of the mechanisms of pain chronicity as well as a brief exploration of the anatomical triggers for acute LBP leading to chronic LBP. Currently, there are limited biomarkers (mostly imaging) or clinical findings that can be used objectively to help the physician in precise anatomic diagnosis. It is clearly understood, however, that the relationship between imaging of the spine and symptoms of LBP is poor. While there is good epidemiological evidence that the burden of intervertebral disc degeneration is predictive of episodes of severe and disabling LBP, translating these findings into improved patient care provides a massive challenge. Human variability and different common comorbidities complicate the picture, and make stratification of patients into correct subgroups difficult. Drugs act by targeting specific molecular pathways and so may be efficient only in subgroups of patients sharing common molecular pathology and common genetics. Both chronic LBP and disc degeneration are known to be heritable. A few genetic variants have been identified and confirmed as associated with disc degeneration, but little investigation has taken place for genetic variants in chronic LBP – which might be those

influencing personality such as tendency to pain catastrophising, peripheral or central pain transmission pathways, or local pathological processes such as inflammation.

LBP has been demonstrated in a number of studies to be heritable. That is, genetic factors are responsible in part for its manifestation. Acute (short lived) LBP has lower heritability than chronic LBP. Chronic LBP is variously defined but most often is considered to be of duration of 3 months in clinical practice, and a duration of 6 months is also commonly used particularly in research. Furthermore, chronic LBP is thought to contribute most to the social problem and disability that is associated with the condition. Thus LBP is now the number one cause of disability worldwide (Lancet Global Burden of Disability 2016). It follows that in LBP both genetic and omic factors can be identified as associated with it. We sought, in the PainOMICs study, not only to identify those factors associated with chronic LBP but to identify specifically those factors contributing to the development of chronic pain from acute, thereby shedding light on pathogenesis and the mechanism of transformation of one type of pain to another.

At the outset of the PainOMICS study, no genome-wide association studies (GWAS) had been performed for LBP and only one GWAS meta-analysis of intervertebral disc degeneration. The individual effects of the identified loci are generally small and explain only a small fraction of the trait or disease variation. As such, they do not substantially improve predictions over those based on known factors such as family history. But they may be useful in identifying new biological pathways which are greatly needed in many conditions leading to chronic LBP.

PAIN-OMICS aimed to identify genetic variants associated with chronic LBP as well as “omics biomarkers” by linking clinical data to a multiple omics analysis to identify:

- Patients who might develop CLBP;
- Diagnosis;
- Objective measure of pain intensity in order to correlate to its pathophysiology, and validate predictors of response to specific (drug) treatments.

The Pain-OMICS consortium was able to perform this type of a study because it combines leading clinical centres, laboratories performing novel omics analyses and leading biostatisticians in Europe, United States and Australia. Our clinical centres UNIPR and OSM (Italy), ZOL (Belgium), St-Cat (Croatia), ECU (Australia) and CPI (USA) are treating over 4,000 new patients with chronic LBP each year. In addition, KCL partner brought in a unique resource of over 13,000 monozygotic and dizygotic twins, excellent model for distinguishing environmental and genetic factors that contribute to

chronic LBP. Extremely well characterized large patient/control cohorts with stored biological samples and related data have been shared with our analytical partners, and, in addition to genetic analysis, glycomics and Activomics analyses have been performed by our partners, Genos and Photeomix, respectively.

In this context, PAIN-OMICS objectives were:

- To perform a large retrospective study and identify multiple “omics biomarkers” for stratification of patients with chronic LBP;
- To validate identified biomarkers for progression of acute to chronic LBP in a prospective study and possible molecular-pathophysiological pathways that could be involved in the genesis and maintaining of CLBP;
- To validate identified biomarkers and test their heritability in a large twin cohort;
- To identify pathways and relevant individual variations for generation, propagation and quenching of pain.

One of the main strengths of this project has been the large number of well-phenotyped samples of subjects that are either currently available to the partners, or will become accessible through prospective collections.

There were some challenges to the execution of this project. The role of co-ordinator of the PainOMICS project was formally taken over by King's College London (KCL) in June 2018. This change was made necessary because in 2017 the scientific coordinator at Università degli Studi di Parma (UNIPR), Dr Massimo Allegri, was accused of violating research ethics and impropriety with a police investigation still ongoing. To mitigate the impact on the PainOMICs project, the consortium voted at their annual meeting (Dubrovnik, Croatia, 17-18 June, 2017) to change the co-ordinating institution to KCL, represented by Prof Frances Williams, taking over the role of the scientific co-ordinator. UNIPR supported this decision as they were no longer in a position to lead the project. The change in leadership had a considerable impact on a number of WPs and caused a considerable delay in the submission of an amendment. The amendment described not only the change to the co-ordinating institution but also a variation to the scientific strategy such that resource could be re-allocated to the studies which were proving the most fruitful and novel.

MAIN S & T RESULTS/FOREGROUNDS

WP2

Clinical phenotyping of patients with low back pain and sample management

The PAIN-OMICS Consortium had to collect a great amount of clinical data and use the samples to ship to the various assaying centres around Europe, then collect the data back to allow for the omics analyses. After the definition of the minimal shared diagnostic dataset, the centres have worked to collect new samples for the retrospective study and organizing more than 2,000 cases already existing for in the TwinsUK study. The PAIN-OMICS consortium used theREDCap (Research Electronic Data Capture Database) to support data capture for research studies, providing 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources. REDCap electronic data capture tools were hosted at UNIPR. After the end of the recruitment phase, patients with LBP had been stratified: 2,811 cases with chronic pain and 712 controls were enrolled for the retrospective study.

The prospective study was the most powerful part of the project design. This is because we recruited acute LBP patients, without the knowledge of their final outcome, thereby reducing bias in participant selection. Participants were followed up and those that developed chronic LBP became the ‘cases’ while those in whom pain resolved became the ‘controls’. This study design is powerful because it reduces confounding between the two arms of the study, by minimising differences between cases and controls. For example selecting people early in the course of their symptoms would serve to reduce any differences in treatment between cases and controls. However, it was much harder than anticipated to recruit adequate numbers having acute LBP and in some centres, such as KCL in London, difficult to maintain follow-up. This is because in large cities there is a more transient population and the UK NHS does not encourage (perhaps correctly) the long term follow-up and medicalisation of chronic LBP as a condition. Thus at the KCL centre recruitment was switched from NHS-based centres to the local osteopathic services at the British School of Osteopathy, with whom we made a productive and beneficial collaboration. The approach in Croatia was to involve many hospitals, each contributing patients to the selection.

We elected, early in the course of the project when the difficulties of recruitment had been appreciated, to allow slightly differing protocols in the different centres and to extend the period for recruitment. This was permitted in order to optimise the number of patients recruited to the study

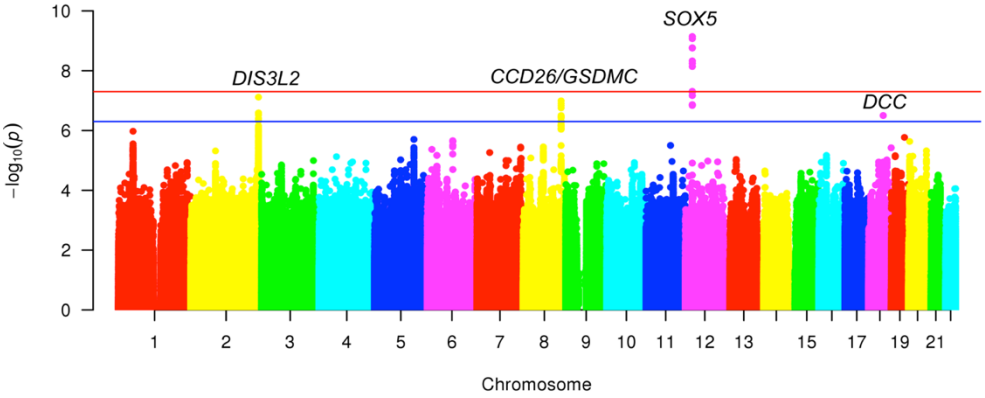
and the material available for analysis. In UNIPR, for example, a very difference approach to follow up is taken to the management of LBP and the patients were seen regularly over 12 months allowing for very intensive follow-up observations. In the prospective study, the following numbers were enrolled: 697 patients with acute low-back pain (baseline); 1,366 at the follow-up after 3 months;473 at the phone follow-up after 6 months and192 during the last follow-up (9-12 months after the acute episode). Additional 106 patients have been enrolled by UNIPR in the epiduroscopy study.

WP3

Genetics and epigenetics of low back pain

The genetic studies planned by the PainOMICS consortium were informed by an important preliminary study. This used existing material from TwinsUK and Generation Scotland where both samples had collected information on LBP and had genotype. A total of 10,215 individuals were examined for association including 2693 cases of LBP defined as “pain between costal margin and gluteal fold on most days in the past 3 months AND/OR disabling LBP for the last 12 months” for TwinsUK (*n* = 4136); and “back pain or discomfort for more than 3 months” for Generation Scotland (*n* = 6079). Genotyping, imputation (1000 Genomes, Phase 3) and association analysis were carried out for each study independently followed by QC, filtering and meta-analysis in a single centre (YuriiA). In TwinsUK, the analysis revealed a significant signal on chromosome 1 ($p < 5e-8$) and suggestive signals on chromosomes 5, 10, 16, and 17 ($p < 5e-7$). In Generation Scotland, suggestive signals were established for loci on chromosomes 5, 15, and 17. After applying filters for minor allele frequency and imputation quality, and adjustment for age and sex, meta-analysis established several regions of suggestive association, strongest for chromosomes 10, 12 and 17 (p = 9.7e-7, 6.3e-7, and 8.6e-7, respectively). However no signal was considered to be genome-wide significant. Thus it became apparent that despite a study of approximately 10,000 individuals which is a reasonable sample size, the study lacked sufficient statistical power to reveal significantly associated loci. This preliminary study also highlighted the marked polygenicity of the chronic LBP trait.

When it became clear that a very large sample would be needed, because of the complex polygenic nature of LBP, we decided to form collaborations and leverage power by use of existing datasets. We achieve this through two ways. First, we joined a CHARGE consortium effort led by Dr Pradeep Suri (University of Washington, US) to study genetic factors in chronic low back pain. Second, we applied for pain data from UK Biobank (n=450,000) so that we could examine not only back pain, but also put it in the context of other chronic pain syndromes as well as exploring known risk factors such as age, body mass index, smoking and social class.



The collaboration with the CHARGE consortium was very fruitful and is continuing beyond the end of the project. We identified four variants associated with chronic back pain on chromosomes 2, 8, 12 and 18 (see Manhattan plot above). The CHARGE study used 158,025 individuals of European ancestry from 16 cohorts in Europe and North America. In order to prove these were real findings and not false positives, we attempted to replicate our findings using the UK Biobank participants of European ancestry. Our study confirmed novel genome-wide significant associations with chronic back pain at three of the loci *SOX5* on chromosome 12, *CCDC26/GSDMC* on chromosome 8 and *DCC* on chromosome18. This study was deposited in Biorxiv initially (https://www.biorxiv.org/content/early/2018/01/20/244483) and then accepted for publication in the peer-reviewed Plos Genetics (https://doi.org/10.1371/journal.pgen.1007601).

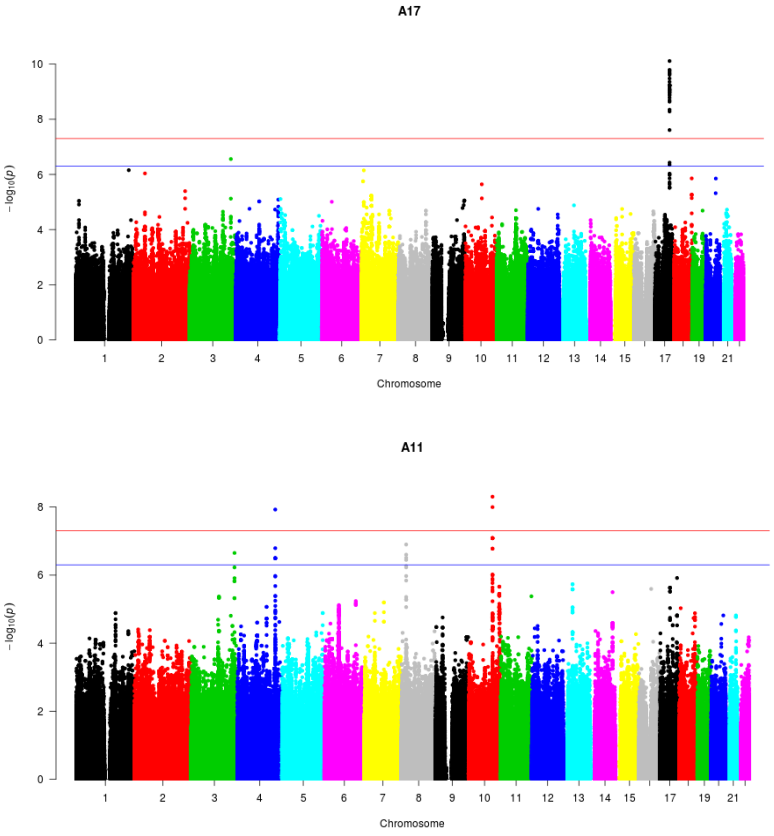
Furthermore, we explored the genetic relationship between chronic LBP and its known risk factors. Our analysis of UK Biobank dataset including about 450,000 individuals of European, African and Asian ancestry along with additional samples from the CHARGE Consortium provided evidence for three loci associated with back pain, including one novel region implicating *SPOCK2/CHST3* genes (under review in Pain – and accepted pending revision, October 2018). We provided evidence for pleiotropic effects of genetic factors underlying back pain, height, and intervertebral disc problems. We also identified independent genetic correlations between back pain and depression symptoms, neuroticism, sleep disturbance, overweight, and smoking. A significant enrichment for genes involved in central nervous system and skeletal tissue development was observed. The study of pleiotropy and genetic correlations, supported by pathway analysis, suggested at least two strong molecular axes of back pain genesis, one related to structural/anatomic factors such as intervertebral disk problems and anthropometrics; and another related to the psychological component of pain perception and pain processing. Overall, the results demonstrated that back pain has an extremely complex genetic architecture that overlaps with the genetic predisposition to its biopsychosocial risk factors.

Genetics of Activomics

A study of 412 cases of chronic LBP from UNIPR was performed. GWAS were carried out for four Activomics traits A2, A11, A17, and A25. We found a statistically significant association for a locus on chromosome 17 for A17 trait (lead SNP 17:61576429, p=6.4E-11) as well as for loci on chromosome 4 (lead SNP 4:161597020, p=5.4E-09) and chromosome 10 (lead SNP 10:101829514, p=4.7E-09) for A11. For the other two traits, suggestive associations were found (p<5e-7). The SNPs associated with A11 and A17 are located inside or around *CPN*1 and *ACE* genes, respectively. Both of these loci are highly plausible candidates for Activomics traits because they contain genes encoding enzymes of relevance. The results provide a step towards understanding the putative mechanisms of genetic regulation of Activomic biomarkers associated with LBP.

Out of the 3,504 PainOMICS samples genotyped, 1,200 samples with Activomics data were selected and genotyped using the Infinium Global Screening Array-24 v2.0 +MD BeadChip according to standard protocols on the Illumina (San Diego, CA, USA) platform at the Genome Analysis Center (GAC) of HMGU. GenomeStudio software v2.0 was used for visual inspection of control probe performance, for initial quality control of raw data and for generation of export files. Mismatches in reported and analysed genetic sex were investigated and samples with bad genotyping quality, duplicates and contaminated samples were excluded. We further investigated the relatedness of PainOMICS samples as well as population outliers and substructure. Furthermore, SNPs with insufficient genotyping quality and deviations from Hardy-Weinberg Equilibrium were excluded. The samples were imputed using the Haplotype Reference Consortium panel (HRC version r1.1) on the Michigan Imputation Server (imputationserver.sph.umich.edu/).

Results are shown for the most striking Activomics, A17 and A11 which demonstrated 1 and 2 genome-wide significant associations respectively (see below Manhattan plots). These peaks were on chromosomes 17, 4 and 10. There were also a number of suggestive associations in each of these 2 traits where peaks fell between the lines of significance shown in blue and red. A final piece of work, in which the additional genotyping will be utilized, will involve trying to replicate these initial findings.



The consortium has decided to go beyond the requirements of FP7 and follow the H2020 standards for making the data generated within the consortium available to other researchers following FAIR principles. For WP3 this means that the imputed genetic data will be uploaded to the European Genome-phenome Archive (EGA, http://ega-archive.org) as soon as proper QC has been done and metadata from the REDCap clinical has been properly linked to the genotype data. The EGA was chosen because it focuses on data requiring controlled access. As such, it will allow other researchers to get access to the PAIN-OMICS genotype and clinical data after sending an application to the PAIN-OMICS Data Access Committee, which will be chaired by partner KCL.

Epigenetic study of LBP

Further progress was made in the analysis of microRNAs and CpG methylation patterns associated with the development and maintenance of chronic LBP after an acute episode, opioid tolerance and responsiveness to therapy. A cohort of 100 cases was sampled during the acute episode of which half progressed to chronic pain during the follow-up visit at 3 months.

To analyse the association of DNA methylation, we used custom arrays (8x60K) on the Agilent platform as previously reported, by choosing 57,899 ChIP probes, 5,077 Ctrl probes and 4,456 unique probes of pain-related genes previously associated with the development and maintenance of

chronic pain. Several CpG sites were significantly associated with chronic pain. In particular we found differential methylation in *FSTK11* (regulation of cell polarity; tumor suppressor), *SS18L1* (Ca^{++} dependent dendritic growth and branching in cortical neurons), *KCNV2* (K^+ channels dependent), *piR-41927* (involvement in maternal epigenetics) *FOXC1* (regulation of embryonic and ocular development), *BAIAP2L1* (function of adapter protein in the formation of clusters of actin bundles, playing a role in the reorganization of the actin cytoskeleton in response to bacterial infection), *SFMBT2* (transcriptional repressor of HOXB13 gene), *RASIP1* (Ras Interacting Protein), *KCNQ2* (K^+ channel slowly activating and deactivating potassium channel that plays a critical role in the regulation of neuronal excitability),*TNRC18* (chromatin binding and transcription regulatory region sequence-specific DNA binding), *PPP2R3B* (Ser/Thr phosphatases implicated in the negative control of cell growth and division) and *ANKRD11* (chromatin regulator which modulates histone acetylation and gene expression in neuronal precursor cells).

Comparing the group of miRNAs from patients with chronic pain in respect to the same patients with acute pain, we identified 75 miRNAs with differential expression between the groups. In particular, 52 miRNAs were down-regulated and 23 miRNAs were up-regulated. By evaluating pathways regulated by modulation of the miRNAs during chronic pain conditions, we found that the difference in miRNAs regulation between resolver and not resolver are primarily related to natural killer cells, neutrophils, macrophages and T CD4+ve cells. Thus epigenetic mechanisms might play a key role in the development of chronic low back pain.

WP4

Glycomic and glycoproteomics phenotyping of samples

Within WP4 of PAIN-OMICS project N-glycosylation changes in LBP patients and healthy individuals were addressed on two levels - the level of total plasma glycoproteins (plasma N-glycome) and on the level of individual isolated glycoprotein, immunoglobulin G (IgG N-glycome). Additionally, N-glycosylation changes were explored both in retrospective case/control cohort and in a prospective setting (converters and nonconverters from acute LBP to chronic LBP) collected by clinical partners of the consortia.

Retrospective study of glycomics

Robust high-throughput ultra-performance liquid chromatography (UPLC) technology was used to analyse plasma N-glycome. Total plasma N-glycans were enzymatically cleaved off from plasma glycoproteins, labelled with fluorescent dye (2-aminobenzamide) and cleaned-up by solid phase

extraction from excess of reagents before the UPLC analysis. Chromatograms were processed with a recently developed semi-supervised Automatic Chromatogram Extraction (ACE) method for automated alignment and detection of glycan peaks in chromatograms. IgG N-glycome was analysed on the level of subclass-specific glycopeptides by mass detection after chromatographic separation (nano-liquid chromatography-mass spectrometry, nanoLC-MS). IgG was first fished out by affinity chromatography from plasma samples, treated with trypsin to obtain glycopeptides that were then enriched with C18-chromatography and analysed by the aforementioned technology.

Both plasma N-glycome (on the level of free N-glycans) and IgG N-glycome (on the level of N-glycopeptides) were analysed in the same subset of retrospective PainOMICS cohort (400 LBP cases and 400 healthy controls collected by Italian centres UNIPR and OSM). Statistical analysis of obtained data has shown that IgG N-glycome on the level of subclass-specific glycopeptides is not significantly different in LBP patients compared to healthy individuals for this sample size. On the other hand, plasma N-glycome profiling by the UPLC in the same set of samples showed far greater biomarker potential for LBP condition compared to profiling of IgG N-glycome by nanoLC-MS technology.

Therefore, plasma N-glycome was analyzed in a retrospective study of total 1888 chronic LBP patients and healthy controls (Trbojević-Akmačić *et al.*, 2018; doi: 10.1016/j.bbagen.2018.07.003) for which the samples were collected in several partnering clinical centres in Italy (UNIPR and OSM), Belgium (ZOL) and Croatia (St-Cat). Relative amounts of high-branched plasma N-glycans were found increased in chronic LBP compared to healthy individuals. These plasma protein N-glycosylation changes in chronic LBP patients are consistent with N-glycosylation changes usually seen in chronic inflammatory conditions such as rheumatoid arthritis. Relative amounts of specific groups of plasma N-glycans (disialylated and trisialylated) were found to be changed in chronic LBP, indicating changes on the level of specific glycoproteins (in their concentration or glycosylation pattern).

Additionally, the analysis of plasma N-glycome was carried out in TwinsUK retrospective samples for validation of findings in other cohorts. Overall, the results were in keeping with a trend towards a pro-inflammatory pattern of plasma glycans in patients with LBP.

Prospective study of glycomics

Observed specific total plasma N-glycome changes have a potential as a diagnostic and prognostic biomarker for chronic LBP and were evaluated in a prospective PAIN-OMICS study. Here, plasma N-glycome was analysed in 1821 samples collected in several partnering clinical centres in the UK (KCL), Italy (UNIPR and OSM), Belgium (ZOL) and Croatia (St-Cat) at two time points (t0 - baseline and t1 -

after 3 months follow-up). We used the same robust high-throughput ultra-performance liquid chromatography (UPLC) technology as for the analysis of plasma N-glycome in the retrospective cohort.

This study was used to address three questions:-

Q1: Can we predict who is going to have pain after three months based on the glycome profile at baseline?

Q2: Is there a difference between people with LBP and controls (validation of the results from the retrospective study)?

Q3: Whether changes in glycans between time points associate with LBP.

We validated the results from the retrospective study by showing the same patterns of changes in glycan levels in cases vs. controls. However, the obtained results suggested that glycomic data reflect the current state of the patients and cannot be used as a predictor for future risk of chronic LBP. That said, perhaps the most important findings of these studies is the nature of the glycosylation changes in chronic LBP. The suggestion of a pro-inflammatory response in this condition was a surprising one. We have investigated this further, to exclude the possibility that a tiny minority of patients with inflammatory spine disease (such as ankylosing spondylitis) might have been recruited to the study and provide results which are driving this signal. This does not appear to be the case: the distribution of the glycomes was almost normal, suggesting a shift in the overall pattern in patients who develop chronic LBP. This does raise interesting possibilities about the inflammatory nature of the chronic LBP, something that is also supported by the results of the epiduroscopy study.

WP5

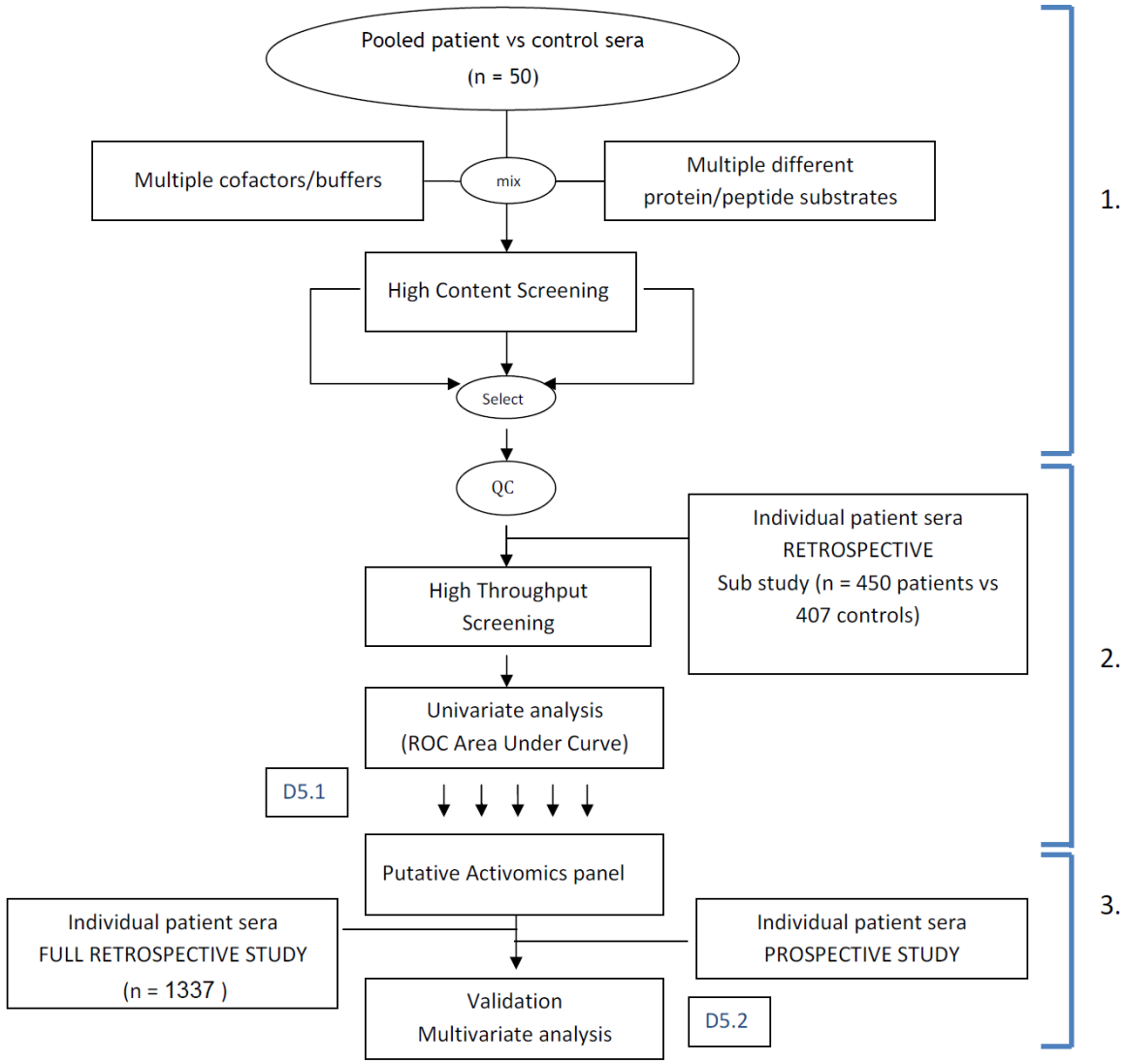
Activomics phenotyping of patients with low-back pain

Activomics® represents a novel approach to biomarker discovery. Permeation of blood with disease-specific patterns of protein modification enzymes is a promising and for the most part uncharted path to diagnose, stratify and monitor chronic disease progression and its treatment. The Activomics® platform employs a systematic approach to detect and quantify post-translational modification (PTM) enzymatic activities, offering novel means to identify and/or stratify diverse pathological mechanisms. High content (low throughput) assays, drawing on a wide array of proprietary protein and peptide substrates, inhibitors, cofactors and assay components, are first used to screen patient samples for the presence of PTM activities and to provide initial evidence of prospective differences in patient versus control samples. Enzymatic activities of interest are then

migrated onto a sensitive, highly accurate microfluidic platform for the high throughput screening (HTS) of large patient cohorts, such as those in the PAINOMICS study.

Methodology

Optimal target substrates are chosen on the basis of their ability to report quantitative differences in activities between biological samples. The principal Activomics® substrate panels screened to date include: proteases (metalloproteinases, serine, cysteine, aspartic proteases, caspases), kinases (ser/thr and tyr), phosphatases and deacetylases. Selected substrates showing putative differential activities with cLBP samples have been fine-tuned by reengineering the substrate recognition motif. For high throughput analyses, the development of high quality screening assays in 384-well format is formally assessed by calculating statistical parameters (e.g. Z'-factor) required as evidence of the quality and robustness of the screen before screening patient samples. The performance characteristics of each selected marker in the HTS have been assessed initially for the capacity to differentiate disease versus control samples in univariate analyses. Sensitivity and specificity for each have been evaluated in Receiver Operator Characteristic (ROC) analyses to define suitability in the design of the optimal predictive biomarker panel.



Overview of the Activomics screening process. Step 1, Activomics (protein modification) substrates are selected via high content assays designed to identify serum active markers. Initial assessment of differential activity is determined on pooled samples from patients versus healthy controls. This stage is low throughput and involves the analysis of multiple reaction conditions to optimize and/or differentiate activities. Step 2, markers selected from step1 are developed for HTS in 384-well format into high quality screening assays (Z' factor > 0.7). Step 3, Activomics markers are screened against large patient populations from retrospective studies. Biomarkers showing promising selectivity and sensitivity in ROC analyses are selected to create the 'Putative Activomics panel' for subsequent validation and multivariate analyses in the full retrospective cohort and in a prospective multicentre study. D5.1 and D5.2 represent the key deliverables of WP5.

Activomics panel in the retrospective cohort

Serum samples obtained for the retrospective study were assessed for the ability of the Activomics approach to stratify patients with chronic LBP quantitatively from those without chronic LBP. The panel of putative Activomics® biomarker activities was selected from an initial high content screen (see above) and subsequently assessed in high throughout analyses with patient samples in the retrospective study being differentiated solely on the clinical manifestation of chronic LBP.

The panel of Activomics markers selected from the retrospective study comprised four key enzymatic activities:

- **A2**: measurement of serum protein kinase A-like activity.
- **A11**: measurement of serum carboxypeptidase N (and B or U) activity.
- **A17**: measurement of serum Zn^{2+}-dependant metalloproteinase activity (unknown protein isoform).
- **A25:** measurement of serum serine protease activity (unknown protein isoform).

The subset, representing the putative Activomics cLBP biomarker panel from this heavily 'discovery' led retrospective analysis of cLBP patient sera, identified two putative Activomics biomarkers capable of differentiating patients with cLBP from controls with good sensitivity and specificity (ROC AUC 0.68 (Act2) and 0.9 (Act25)). Two additional normalisation markers (Act 11 and Act 17), furnished us with an initial putative Activomics panel for the subsequent prospective study.

Activomics panel in the prospective cohort

Clinical samples from the prospective study were screened against the selected Activomics chronic LBP biomarker panel. A total of 1337 samples were screened against each of the four Activomics markers, including 879 baseline samples and 459 samples collected at the follow-up study visit (3 months or later). The Activomics screening results for the prospective study results were obtained and collated blindly, without knowledge of patient clinical characteristics, so as to minimise bias.

Direct comparison showed reasonably consistent results between individual centres with notable exceptions.

- Low levels of kinase activity (A2) were detected in samples from St Cats.
- Low values of serine protease (A25) were detected in OSM patient samples.
- Low levels of Zn^{2+}-dependant metalloproteinase activity (A17) were associated with ZOL samples

In general, large differences in Activomics activities were not detected in comparison between patients with and without chronic LBP (t1 versus t1NC results). However, further comparisons suggest potential differences in the Activomics panel can be detected in chronic (t2) LBP vs non-chronic (t1NC) LBP patient samples.

- A17: patients with non-chronic LBP at follow-up *vs.* chronic LBP (both visits) showed diverse activity values of 41.5 ± 17.5% *vs*. 34.2 ± 17.5% associated with a moderate ROC AUC of 63%. When only the second follow-up visit was considered, the ROC AUC rose to 0.72, suggesting that the persistence of chronic pain might be a factor in differentiation by this marker
- A2: the second follow-up visit showed greater differences in this biomarker in non-chronic LBP than the first study visit. In this case, kinase activity in non-chronic patients was higher at the 3 month study visit than obtained with patients suffering chronic LBP at the second visit (46.9 ± 16.8% *vs*. 40.3 ± 19.1%; ROC AUC = 0.62).

Conclusions of the Activomics study

An important secondary endpoint complying with the published clinical study protocol was to assess the potential for Activomics biomarker measurements to stratify patients enrolled in the prospective study into those that develop chronic pain, defined as lasting more than 3 months, from those whose pain resolves naturally within the same period. None of the Activomics biomarkers selected from the results of the retrospective study demonstrate a clear association with this particular clinical manifestation.

Differences in serine protease (A25) and protein kinase (A2) activities reported previously for chronic LBP patients *vs*. controls in the retrospective study were not reproduced in the prospective study, when based on samples from patients declaring chronic LBP versus non-chronic LBP at the 3 month follow-up study visit.

Only A17, measuring serum metalloprotease levels, showed any significant correlation (P value = 0.02). The markers displayed only poor diagnostic predictive power when assessed via the receiver operator curve area under the curve (ROC AUC) criteria; 0.5 is considered as essentially random compared with 0.7 having good or > 0.9 showing excellent diagnostic power. In conclusion, the selected panel cannot be used to predict which of the patients suffering pain will go on to develop chronic pain status.

In order to better understand these differing results between the retrospective and prospective studies, the medical data uploaded to Redcap on patient demographics and clinical characteristics

was combed for associations with the measured Activomics data, most particularly pain status, current pain and other medications, and comorbidities.

This approach revealed a putative association between current pain or other medication with specific Activomics measurements. Two Activomics markers showed significant associations with current therapy for specific medications. The marker that appeared to be most clearly associated with therapy was the metalloproteinase measured by A17. Several therapies were associated with significant drop in serum metalloprotease activity levels. Whilst the reason for this is unclear, the therapies showing putative associations tend to be related to cardiovascular treatments. Due to the imbalance in the baseline demographics of patients on therapy versus those who are not, ROC AUCs tended to be less convincing than they might be with more balanced populations. Nonetheless, this is an important observation relating therapy to apparently unrelated serum enzymatic activities and should be explored further.

Potentially the most exciting result was the apparent association of serum protein kinase A levels and the use of non-steroid anti-inflammatory compounds (NSAIDs). Patients on NSAID therapy showed a statistically significant drop in serum PKA activities (35.4 ± 18.4 on therapy versus 46.1 ± 17.7 off therapy; p value <0.0001; ROC AUC 0.68). This association was maintained and extended at the subsequent 3 month visit (28.7 ± 18.6 on therapy versus 44.4 ± 17.0 off therapy; p value <0.0001; ROC AUC 0.76). No correlations between two markers, A11 and A25, were noted with pain therapy. Interestingly, A11 was recently associated via Activomics-GWAS with genetic polymorphisms in or around the carboxypeptidase N gene (see above). We have not found any other demographic or clinical feature associated with this Activomics marker. This might suggest that the natural fluctuations in the screening results between individuals might associate more with genetic factors than environmental ones. Further research is warranted to further investigate these observations for potential clinical exploitation.

In order to understand the association with clinical data better, we examined which type of NSAIDs are associated with lower serum PKA (A2) levels. All non-aspirin NSAIDs appeared to display the same correlation leading to lower serum PKA activities. By contrast, acetylsalicylic acid (aspirin), also a NSAID with a similar mode of action in preventing prostaglandin synthesis by inhibition of the cyclooxygenase (COX) enzyme, did not show association with lowered serum PKA.

The different compound structures for the different NSAIDS suggested that the effect is unlikely to be targeted against the serum PKA enzyme itself, i.e. by direct interference through catalytic site inhibition. Indeed, such inhibition would essentially require strong or irreversible inhibition of the

enzyme in order to observe its impact in our *in vitro* tests. It is also difficult to explain these effects simply through inhibition of COX-mediated prostaglandin synthesis, as both aspirin and non-aspirin NSAIDs operate with similar modes of action on the COX enzyme whilst aspirin does not appear to lower serum PKA activity. Nonetheless, the prevalence of these over the counter non-prescription drugs and their potential effects on PKA, an important enzyme in the front line of the innate immune response against microbial infection, warrants further investigation for two major reasons. Firstly, NSAID interference with serum PKA may expose users to risks of unintended comorbidities. Secondly, essentially repositioning or repurposing an approved drug therapy with the capacity to influence PKA activity in this way may offer an innovative route to modulate the kinase activity for alternative therapeutic means. Indeed, the apparent influence of therapeutic agents on non-intended targets described herein may offer potentially important medical vantage points, and raises several opportunities for further research with prospective medical exploitation.

WP6

Integrated models for the identification of biomarkers and potential new therapeutic targets

During the course of the project WP6 was active in development of new methods and models for multivariate and integrated biomarker and target identification; WP6 was in charge of integration of multiple kinds of big data generated in this consortium; for providing storage and computational analysis facilities for the integrated data generated in this consortium and other big relevant data generated elsewhere (e.g. UK Biobank, functional genomic data). WP6 coordinated integrated data analyses with primary goals of identification of back pain biomarkers and therapeutic targets.

Our case-control studies of N-glycosylation of IgG in the context of chronic lower back pain indicated that antibody-dependent cell-mediated cytotoxicity (ADCC) has a distinctive role in CLBP. We have also observed changes of total plasma N-glycome: higher abundance of highly branched sialylated glycan structures in people with chronic lower back pain (CLBP) compared to controls, indicating another mechanism involved in CLBP. Studies of Activomics allowed to establish several significant associations with CLBP.

In other fields of common complex traits it is becoming evident the the translation to benefit patients is rarely obtained from the top associated genetic variants, but may be afforded by combining the other associated variants in a polygenic risk score. This has recently been shown to explain much more of the variance in an individual trait and may therefore improve prediction of the dvelopment of a given trait – for example coronary disease (https://www.nature.com/articles/s41588-018-0183-z).

The identification of at-risk subjects is considered by epidemiologists now to provide the most tractable way of getting people to change their habits. Recent efforts, for example, to increase fruit and vegetable consumption by the whole population have been largely unsuccessful. Similarly, most people acknowledge the need to be more active, but few healthy individuals possess the degree of motivation to make longstanding changes in their lifestyle. Thus public health messages provided in an untargeted fashion to the entire population do little to change behaviour - while legislation may be much more effective (for example, smoking ban in public places). However a test which provides an individual with an accurate measure of their risk based on their own unique characteristics is a powerful tool for driving change. That is the rationale behind the personalised medicine efforts and the aim of the PainOMICS study. We wished to make a test which would help identify those at greatest risk of chronic LBP. We now know that such individuals are most likely to have occupational characteristics, be of low social class and have beliefs about the pain which are not accurate – catastrophizing and assuming that the pain was telling them to rest and avoid activity so as to lessen the chance of further injury. While our management of LBP many decades ago did promote the use of rest, we now know that outcomes are better if people remain active during an episode of acute LBP.

Our genetic studies have complemented and reinforced these findings. The largest ever—more than half a million samples—genome-wide association study of back pain suggested at least two strong molecular axes of BP genesis, one related to structural/anatomic factors such as intervertebral disk problems and anthropometrics; and another related to the psychological component of pain perception and pain processing. The studies of genetic control of the plasma glycome and Activomics were performed using a genome-wide association approach, led by investigators from the PAIN-OMICS consortium. Large external functional genomics data were collected. Integration of the results of the back pain genome-wide association with big functional genomic data provided a list of candidate causative genes which, when up- or down-regulated, may lead to back pain, and thus represent potential drug targets. We have also used genetic approach to infer other possible interventional targets. In particular, using this approach, we confirm obesity and anxiety/depression as potential therapeutic intervention targets for back pain.

The consortium has decided to go beyond the requirements of FP7 and follow the H2020 standards for making the data generated within the consortium available to other researchers following FAIR principles (see D6.4 “Posting results from retrospective study”), and WP6 coordinated this activity. We have selected European Genotyping Archive (EGA) for the data requiring controlled access. This

will allow other researchers to get access to the PAIN-OMICS genotype and clinical data after sending an application to the PAIN-OMICS Data Access Committee, which will be chaired by partner KCL. The upload of genotype and corresponding clinical data to EGA complements the glycomics and Activomics data uploaded to Zenodo (http://www.zenodo.org) as part of the PainOmics Zenodo community at https://zenodo.org/communities/painomics. WP6 has also shaped the development of the GWASarchive, a browsable interface to GWAS results which can be found at http://www.gwasarchive.org, where one can currently find summary statistics of three GWASes of back pain and the GWAS of plasma glycome, describing more than a billion of genotype-phenotype association (also available from the PainOmics Zenodo community).

WP7

Validation of biomarkers in a prospective cohort

In WP7 clinical investigators in the PAIN-OMICS consortium recruited patients experiencing an acute episode of low back pain. These patients did not have a prior history of chronic low back pain. Patients recruited in WP7 have been followed prospectively to determine those who go on to develop chronic low back pain versus those who do not. Chronic low back pain included patients who continued to experience low back pain 3 months following the initiation of their acute low back episode.

Historically, it is obvious that all patients who suffer from chronic pain experience an acute phase prior to later developing chronic pain. This is true with LBP as well as other acute and chronic pain diagnoses. There has been much interest in the past several years in the "chronification" of acute low back pain, much of which is non-specific in its diagnostic features. Clearly, many patients who experience an acute pain experience such as LBP improve and resolve their pain and do not progress to develop chronic low back pain. While there is a growing understanding of the features and patient characteristics of patients at risk of developing chronic pain after a particular acute pain injury, little if anything is known about the genetic and epigenetic features that may contribute to this process. A better understanding of this progression to chronic low back pain would be very important to improve clinical care. If we can better understand the factors of why a subset of patients are likely to develop chronic low back pain, then we may be able to change the natural history of an acute low back pain episode in the susceptible individual. It would be extremely helpful to identify a biomarker such that the clinician would know the patient most at risk of developing chronic low back pain. Currently, we know that for the vast majority for whom specific triggers cannot be identified, those at increased risk of chronic LBP are those with severe pain and psychological distress, those with pain at multiple other sites, having occupations involving in which the individual does not feel in control,

and those of lower social class. There is also good evidence in developed countries that compensation systems contribute to the societal burden of disability from LBP.

Specific diagnosis of anatomical pain generators in patients with non-specific low back pain can be difficult for all clinicians, both in primary care or pain specialists. This is particularly true when there is no available radiologic studies (plain films and/or CT scans and/or MRIs) and indeed it is recognised that the correlation between imaging and patient characteristics is poor. So patients who have imaging studies often present with symptoms or physical findings that do not correlate with radiologic examination. This often further complicates the diagnostic considerations and may, indeed, adversely influence the course of the pain.

A primary factor of the prospective study was to stratify patients with low back via diagnostic considerations. Categories were given to clinicians and included: zygapophyseal/facet joint, sacroiliac join, stenosis, widespread degeneration, radiculopathy, or discal (discogenic) pain. Clinicians could also check “unknown/cannot say” as one of the diagnoses or pain generators. While more than one pain generator may exist in the same patient, practitioners were asked to choose the pain generator most likely or most important to the complaint of low back pain.

Prior to WP7 the PAIN-OMICS consortium had performed a retrospective study to look at biomarkers that could be associated with chronic low back pain. The results were used to aid in determining appropriate biomarkers to examine in this prospective cohort of acute low back patients and examining differences between those whose acute low back episode resolved versus those patients whose acute low back pain progressed temporally to chronic low back pain.

Adult patients with acute low back pain who met inclusion criteria for the prospective cohort study were eligible for enrolment. Informed consent was obtained in all patients and blood samples drawn at the baseline visit. The pain generator for patients at baseline was determined in a majority of patients enrolled in this prospective cohort. It was recognized that determining pain generators in a group of patients with non-specific acute low back pain would undoubtedly be imprecise based on current understanding of the heterogeneity of low back pain.

Epiduroscopy study

We enrolled patients in the prospective cohort and subsequently performed epiduroscopy. Epiduroscopy allows for direct visualization of nerve roots and other structures in the epidural space. Epiduroscopy may provide better diagnostic criteria for the aetiology and pain generator in patients with acute low back pain. The consortium examined concentrations of IL-6, IL-10 and IFN-γ in the serum of patients in whom epiduroscopy was performed. Samples were collected and sent for

GWAS, genetic, Activomic, and glycomic analysis at baseline, 3 month, 6 month, 9 month, and 12 month from date of onset of acute low back pain.

Following the secondary objectives of the epiduroscopy study, the consortium also analysed other genetic variants related to inflammation by a gene candidate approach and measured the concentration of IL-6, IL-10 and IFN-γ in the serum of patients by ELISA, using specific monoclonal antibody (Endogen Tema). In analysing genetic variants such as IL-6 and demographic data such as obesity (example of variable linked to development of chronic low back pain in other results suggestion) the data from this cohort has been examined.

The analysis to date shows no correlation or association of those patients who went on to develop chronic low back with the specific inflammatory marker IL-6. This analysis was performed with and without BMI as a covariate. This may be due to the small sample size, that SNP is not important, or specific to this particular cohort of patients that were selected from one site and had agreed to epiduroscopy. Preliminary data showed that there is a statistically significant variation in the levels of IL6 during the course of the transition of patients from baseline to follow-up.

To reach the primary purpose of the study, the consortium genotyped *IL6* (rs1800795) SNP, in the promoter of the *IL6*, by using Taqman SNP Genotyping Assays (Applied Biosystems, Foster City, CA) and a LightCycler 480 Real-Time PCR System (Roche Diagnostics Ltd, Lewes, United Kingdom) according to the manufacturer's instructions. We compared the the frequency of rs1800795 in patients with favorable treatment outcomes (>30% pain scores after 1 month) (Group RESP) (from literature data up to 57% of patients) with the frequency obtained in patients without favorable treatment outcome (Group non RESP). Genotype/phenotype correlations was performed in collaboration and showed no association of chronic LBP with the SNP genetic variant in *IL6* gene.

A look from one of the partners in this project, OSM, enrolled 73 patients in this prospective project, 52 of whom did not go on to develop chronic low pain. Epiduroscopy was performed to aid in the diagnosis of the pain generator and to see visual observations of inflammation. Initial analysis did not show any association of IL-6 levels and the development of chronic low back with or without BMI as a covariate. Further genetic analysis in this group thus far suggest that the levels of inflammation in the epidural space as reflected in the serum samples do not appear to reflect whether acute LBP becomes chronic or not, nor indeed do they reflect the genotype of the *IL6* gene. It might have been interesting to examine the levels of these inflammatory markers in the cerebrospinal fluid on epiduroscopy but because our aim was to produce useful biomarkers, and epiduroscopy is not a standard procedure, this was not performed because there was little possibility of this becoming standard practice in back pain care.

To date, 1697 patients have been enrolled in the prospective controlled trial to identify biomarkers involved in the transition from acute to persistent chronic LBP. These constitute patients who initially developed acute LBP episode. At the present time n=1371 patients and their samples have follow up data at 3 months. Three months of persistent back pain following an acute episode defined the presence of chronic back pain for the purposes of the current study. Of the n=1371 patients with data from 3 month follow up, n=861 had developed chronic LBP from their acute back pain episode. The remaining 510 patients did not progress to chronic LBP. At the time of this report n=473 patients have had 6 month follow ups and 192 have been followed for 12 months. There have been n=4335 samples sent to HMGU (n=821), GENOS (n=1934), and Photeomix (n=1580), respectively.

From the original number of patients enrolled in the prospective controlled trial there have been 1624 patients with samples that could be analysed. These come from 7 clinical sites in Europe, USA, and Australia, the vast majority from European sites. Data continues to be collected and tabulated for patients who have not achieved the 3 month or 6 month thresholds.

The following table reports the findings in this cohort of acute back pain patients:

	OSM	ZOL	St-Cat	UPR	CPI	ECU	POLY	KCL	Sum
Zygapophyseal/Facet Joints	22	20	25	83	0	0	0	0	150
Sacroiliac Joint	13	6	16	36	0	0	0	0	71
Stenosis	0	6	2	10	0	0	0	0	18
Widespread Degeneration	20	1	5	6	0	0	0	0	32
Radiculopathy	2	5	81	405	1	0	0	0	494
Discal (Discogenic) Pain	2	49	130	19	0	0	0	0	200
Unknown/Cannot say	14	123	0	65	0	6	0	0	208
NA	0	10	6	10	0	0	0	425	451
Sum	73	220	265	634	1	6	0	425	1624

The 425 patients from KCL were not evaluated by a physician because they were recruited, for the most part, at the British School of Osteopathy, so these patients were not considered for determination of the pain generator.

A further question focused on stratification of acute low back pain patients versus chronic low back patients with respect to the pain generator. Of the 1624 patients the following was noted:

	no	yes	NA	Sum
Zygapophyseal/Facet Joints	88	46	16	150
Sacroiliac Joint	36	29	6	71
Stenosis	7	9	2	18
Widespread Degeneration	20	10	2	32
Radiculopathy	260	180	54	494
Discal (Discogenic) Pain	164	30	6	200
Unknown/Cannot say	141	38	29	208
NA	145	168	138	451
Sum	861	510	253	1624

OSM recruited n=73 patients in the prospective study. Of these, n=52 patients did not develop chronic low-back pain, while n=21 developed it, three months after the enrolment. Seven patients dropped out, all of them developed chronic pain. We obtained the bio-samples after 3 months on n= 3 of them.

In conclusion, the prospective study successfully enrolled n=1624 patients with evaluable samples. It is believed by the group that the prospective study will add valuable and novel information to our understanding of low back pain. Very little is understood why many patients with non-specific low back pain injuries heal without sequelae while others go on to develop chronic low back pain.

In the current project we were able to stratify low back patients in several ways that will be important when looking for biomarkers. At the simplest level, we could stratify patients by gender and age. More importantly, perhaps, we stratified patients by pain generators. This may prove important in looking at biomarkers. Furthermore, we were able to stratify low back pain patients into two important groups: those that heal and do not progress to chronic pain versus those, whose acute injuries appear similar but go on to develop chronic low back pain.

WP8

Validation of biomarkers in twins: heritability of biomarkers

The TwinsUK NIHR BRC Bioresource provided an assessment of predictive capacity of biomarkers revealed in other studies within PainOMICS as well as to providing ground for establishing the genetic basis of susceptibility of low back pain via analysis of heritability of the biomarkers and genome-wide association study.

We examined the predictive capacity of plasma and IgG N-glycome for the development of chronic LBP. IgG glycome data were available for ~4500 twins and plasma glycome data were available for ~2500 twins. Thus, this provides highly powered sample for association studies and discriminant analyses. Also, heritability of these biomarkers was established using classical twin design (see

below) and by fitting ACE models to estimate the relative contribution genetic and environmental factors on their variation.

Penalized linear discriminant analysis (LDA) based on Fisher's linear discriminant was applied using using either IgG N-glycome or plasma N-glycome or both. All the discriminant models exhibited relatively high specificity (61.7%-63.7%) and low sensitivity (38.7%-46.5%). Accordingly, there were high negative predictive values and low positive predictive values of the glycomic markers (see table below).

Prevalence of LBP in the population, %	Sensitivity, %	Specificity, %	Negative predictive value, %	Positive predictive value, %
10	46.5 (40.1-53.0)	61.7 (57.7-65.7)	91.2 (90.1-92.2)	11.9 (10.2-13.8)
20			82.2 (80.2-84.1)	23.3 (20.4-26.5)
25			77.6 (75.2-79.8)	28.8 (25.5-32.4)
30			72.9 (70.2-75.5)	34.3 (30.6-38.2)

This implies that the models could be used reliably in the identification of patients at low risk of developing chronic LBP (good negative predictive value), but less effective in prediction of individuals with high risk of chronic LBP (poor positive predictive value). In the model that included IgG N-glycans alone, the following glycans provided the most remarkable contribution: IGP25, IGP49, IGP62, IGP64, IGP70-73. The majority of these glycans belong to the co-regulated modules identified by us as associated with LBP. The results point towards a lack of fucosylation as a possible mechanism of chronic LBP development or exacerbation and have been published in Scientific Reports (doi: 10.1038/srep26815). The most contributing glycans in the model of plasma glycans were: GP1, GP66, GP73, GP76-78, GP81, GP84-87, and GP91. The overall pattern of the changes in these glycan traits also points at the decrease of fucosylation in people with chronic LBP and provides further evidence that the findings are real and provide mechanistic insight into the development of chronic LBP.

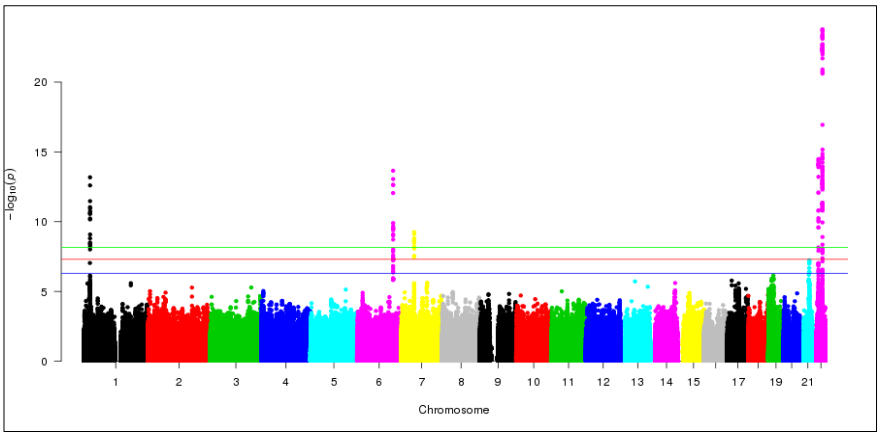
The heritability of plasma glycans was estimated using structural equation modelling to decompose the observed phenotypic variance into three latent sources of variation: additive genetic variance (A), common environment variance (C), and unique environment variance (E). Heritability was estimated from the most parsimonious model as a proportion of the observed phenotypic variation attributable to genetic factors. For the vast majority of plasma glycans the most suitable model was provided by

the AE model which describes a combination of additive genetic and non-shared (among twins) environmental factors as providing the major sources of variation. The lack of influence of shared environmental factors contributing to plasma glycome variation suggests that non-genetic family contribution for plasma glycans variability is very limited and that most of the contribution from the environment is not shared by members of a twin pair. In contrast, for two plasma glycans GP29 and GP36, the best model was CE. This provides no contribution from genetic factors and suggests that the environment (both common and non-shared) is the main source of variation. The heritability estimates varied from 17% to 74% with a majority of cases around 60%. These estimates are high – much higher than was expected - and suggested that significant associations in a genome-wide association study (GWAS) of plasma glycans would be detected, if adequately powered. Overall, these data suggest a high contribution of genetic factors to the inter-individual variation of many plasma glycans.

A GWAS was carried out for clusters of correlated IgG N-glycans found by us to be associated with LBP (doi: 10.1038/srep26815). Genome-wide significant associations have been identified for all the clusters within or near the genes responsible for glycosylation, immune response and inflammation (eg. *FUT8*, *B4GALT1*, *ST6GAL1*, *HLA-DR*, *RUNX*, *IKZF1*).

These results were consistent with glycome GWAS data published earlier or obtained within PainOMICS (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5587582/).

Figure showing GWAS of turquoise cluster of glycans in TwinsUK



While the associations that reached genome-wide statistical significance are of great interest and may provide information about pathways, it is recognised that a vast amount of useful genetic information exists in the strata of GWAS results that do not reach genome-wide significance. Recent direction in this field has turned towards combining these results in order to mine the associations

for information of potential use. This approach is called a polygenic risk score (PRS) – literally, using the results of many gene variants to determine likely pathways of influence or to provide a measure of genetic risk. With this in mind, we used the results of GWAS for three clusters associated with chronic LBP to make lists of SNPs from which to calculate PRS in the UK Biobank dataset. While we found that the PRS was statistically significant, it offered a poor prediction of the risk of back pain (R^2 = 2.6e-5, p = 0.012 for SNPs with p<5e-8 in glycome GWAS). We also used the results of meta-GWAS carried out using UK Biobank data for back pain to calculate PRS in TwinsUK dataset. We found that PRS calculated using SNPs with p<1e-3 (N = 7269) explained almost 1% heritability for disabling LBP and about 0.5% heritability for any LBP in TwinsUK, suggesting relatively high predictive potential. We also found that stratification of people into low and high risk groups by percentiles of the PRS was statistically significant (OR ~ 0.7 for low risk group and OR ~ 1.5 for high risk group vs medium risk group, p<0.05 in both cases).

These results suggest the potential for the development a predictive test for the risk of back pain becoming chronic using PRS based on GWAS data. This may become a reality in the future as the cost of genotyping continues to fall. In addition, a full GWAS would not be required, a prognostic testing kit would only need to contain the SNPs that were used in the generation of the PRS. Another approach would be to use this information to improve the current screening tool used to predict those at risk of developing chronic LBP. One such tool in widespread use is the STarT Back screening tool developed by Keele university (https://www.keele.ac.uk/sbst/startbacktool/) which uses a few simple questions around pain and catastrophising of pain to identify patients at risk of non-resolving back pain. We envisage using this tool in combination with a PRS to see whether we can improve prediction. A study is already planned and will be subject of future grant application, in collaboration with Jaro Karppinen of the Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland. He is performing studies comparing a Finnish screening tool Örebro Musculoskeletal Pain Screening Questionnaire with STarT Back and, more importantly, has recruits for whom DNA has been collected and GWAS performed. Thus we will be able to observe, in a prospective study, whether either of these screening tools in combination with the PRS, can be improved in their predictive capacity by including genetic information. Based on these results, and other common complex traits (see above), the accurate identification of patients at high risk for future chronic LBP – those who might contribute to the world's leading cause of disability according to the Lancet – is a very real possibility in the next decade or so.

WP9

Ethical considerations of pain

UNIPR has coordinated a special group that has addressed all possible ethical concerns of this project involving not only the main sections of the PainOMICS project (the retrospective and prospective patient collections, and the epidurscopy study held at UNIPR) but also the broader consideration of clinical research into a subjective symptom such as pain. Furthermore, we have given consideration to the nature and value of informed consent and the role of placebo in the management of chronic pain. This latter subject has become a very hot topic in recent months, with a widespread public debate (BBC Horizon October 2018 https://www.bbc.co.uk/news/health-45721670)
about the potential usefulness of placebos in clinical practice and some high profile journal publications calling for greater consideration of their use (https://www.bmj.com/content/363/bmj.k3889).

The design of clinical trials was evaluated firstly by each group and the final version of the clinical trial protocols was submitted to Ethical Committees of the involved clinical participants of the Consortium. Ethics and regulatory approvals and training certificates, trial registration numbers are available in clinicaltrials.gov in accordance with current recommended practice in this field. We also published, in collaboration with leading pain ethicist Dr Michael Schatman of Foundation for Ethics in Pain Care, Bellevue, Washington, USA, a review of the ethics of pain entitled (https://onlinelibrary.wiley.com/doi/full/10.1111/papr.12485) in which the challenges of performing studies to high ethical standards, and of biobanking, were explored across the world. In this paper the governance in the EU countries and the USA were compared and contrasted, with particular emphasis on the role of informed consent. Discussion was made relevant by example to the PainOMICS study.

Furthermore, the management of the samples across the PainOMICS study was coordinated, from the collection to the dispatch to the specialized laboratories, creating the documents in order to provide a reproducible model to use the same conditions of the biological samples from each patient enrolled in the retrospective and prospective studies. For this purpose three SOPs were defined to

(i) validate that the personalized documents already in existence for samples collection are compatible with the specific analyses and

(ii) provide reliable samples from all the centresby recording all of the collected data in a special database, to use all the information in the maximum protection of the privacy.

In summary, multiple steps in the execution of the PainOMICS study were designed with ethical considerations uppermost in mind. The protection of the identity and biological uniqueness of the kind individuals who generously shared their clinical story and samples with the consortium was considered at all points in the study design. Separate databases of clinical information and personal details were maintained by different groups with access to only one source. Samples were transferred between centres using 2 anonymised identifiers which could not lead back to the individual except by password-protected database interrogation.

SOCIO-ECONOMIC IMPACT AND SOCIETAL IMPLICATIONS

The PainOMICS consortium is composed of four clinical, three academic and four SME partners characterized by a comprehensive vertical and horizontal integration of the clinical, academic and industrial attributes. Our partners are drawn from the leading European researchers in research and development with a wide variety of backgrounds. All aspects of chronic pain relating to the low back are covered in this project, from clinicians drawn from anaesthesia, rehabilitation and rheumatology actively treating patients to geneticists and omics experts who provide detailed molecular characterisation and innovative data analysis. In coming together our research-intensive analytical SME partners (Genos and PHOTEOMIX) have significant exposure to commercial markets and experience with diagnostic applications. Their collaboration with experienced clinical partners, with insight into the greatest problems facing patients today, facilitated translation of useful research results into potential the diagnostic products of the future. Collaboration with prominent statistical geneticists (HMGU, YuriiA, KCL) who have demonstrated global leadership in genome-wide association studies and the development on new innovative tools for multidimensional data analysis greatly increased the chances of successful scientific impact and progress beyond the current state of the art. Our clinical partners are potential end users of the developed technology which will help to optimize decisions about research direction and enable successful translation of research results into clinical use.

We have produced a panel of biomarkers of adequate sensitivity and specificity to predict progression of acute LBP to chronic LBP and with it raise intriguing questions about the nature of this change. We anticipated, at the outset of the project that the results would point to abnormalities in pain processing and possible central nervous system mechanisms. To our surprise, the glycome results indicate that unsuspected levels of local inflammation may be mediating the chronicity of pain. Furthermore the GWAS results show that rather than pain mechanisms we have strong association with genes involved in musculoskeletal and spine development. This points once again to the degenerative process of the spine – in the intervertebral disc and facet joints – that has been shown previously in epidemiological studies to be one of the greatest risk factors for chronic LBP. Our

results raise the very real possibility of a comprehensive screening test for chronic pain risk without the patient having to fill out lengthy questionnaires (with some degree of subjectivity involved) and would hint at a possible primary care blood test to enable GPs to stratify those at high risk of chronic pain into a group for intervention. Alternatively, the PRS would be used in conjunction with the existing questionnaire methods of stratification (such as the STarT Back tool) and studies are being set up to examine this in future. Clinical trials of such an intervention would be required with accurate comparison of existing methods with our newly proposed ones, but being able to target the most vulnerable patients would make such trials very much more cost effective and therefore more likely to be performed. The rising recognition of the burden of disability provided by back pain, not just in the EU but across the world, means we cannot afford *not* to do such clinical trials of patient stratification.

Our key results have been widely disseminated not only at scientific conferences (such as the World Congress on Pain, Boston 2018 and the Society for Back Pain Research, Groningen November 2018) but more broadly through visits to patient organisations (Kensington and Chelsea forum, London 2018), partner organisations (British School of Osteopathy, London 2016 and 2017) schools (Tormead School, UK 2018) and through digital (BBC radio 4 Inside Health, October 2018) and social media (Youtube animation film to be uploaded once Freidin et al accepted for publication; reviews thus far are positive). We will continue to raise awareness of the many biological mechanisms implicated in low back pain. Our publicity will seek to put back pain firmly on the public agenda, raise awareness of the huge costs of back pain to society, and promote the use of existing tools to stratify patients into low and high risk. We continue to disseminate our findings at international scientific meetings, at national charity level and to the lay public, by engaging both face to face and using social media. The wider societal implications of the project will be realised by continuing to disseminate our findings. Importantly, any scientific findings given a platform in dissemination has been and will continue to be used as a way to remind the public the best way known to manage episodes of acute low back pain (as in BBC radio 4 interview). This is a symptom that almost everyone has at some time in their life and normalising it, and helping people manage it without recourse to the many unvalidated interventions available, is an important part of the PainOMICS message.

Impact on our understanding of the low back pain

PAIN-OMICS project is expected to significantly expand the level of knowledge how low back pain is generated, propagated and treated. We have mobilized significant human and material resources in seven European countries and the USA and for the first time enable comprehensive characterization of large cohorts of patients with chronic low back pain. Biomarkers related to different aspects of chronic LBP, as well as potential new targets for therapy were identified.

Translating basic research into clinically useful biomarkers and targets

The PainOMICS consortium has benefited from synergy between leading European and US scientists in working on pain, genomics and glycomics. Following the successful identification of a number of genetic factors associated with chronic LBP, as well as the development of a PRS, we have also focused on the environmental influences. These are captured by some of the risk factors in UK Biobank genetic study and by analysing epigenetic, glycomic, glycoproteomic and Activomic aspects of chronic LBP.

Impacts of the new discoveries on patients, their clinicians and treatment of LBP

Chronic pain is a particularly unpleasant condition of very high prevalence and patients suffer in many ways. In this field LBP is one of the most common causes of chronic neuropathic pain affecting up to 15% of general population with a huge social and economic impact. The current lack of diagnostic tests means that many patients are often given inadequate therapy or surgery and remain mistreated for years. Meanwhile, their condition progresses and their symptoms can get increasingly severe. There is clearly an urgent need for better tools for uncovering the molecular basis of LBP, developing reliable diagnostic tests and biomarkers that could better stratify patients and direct the individual patient towards more personalized and effective therapy. The ideal would be to use treatments that go beyond just alleviating symptoms. The technologies developed in this programme address this need and, together, form a powerful new system for studying the molecular aetiology of chronic LBP. We hope the next stages of research following on from this programme will have a profound, beneficial impact on these patients and the clinicians responsible for their treatment. Firstly, the availability of non-invasive clinical tests to stratify patients with low-back pain would allow patients to be quickly put on suitable clinical management programmes to alleviate their symptoms. A realistic prospect is the combination of questionnaire screening with the PRS we have developed in the genetic study of chronic LBP. Secondly, tests for early stage diagnosis of progressive pain could eliminate years of unnecessary, inappropriate medical intervention (or lack of intervention) which recent work suggests may exacerbate disability. Thirdly, targeted therapeutics that might arise from the greater understanding of the molecular aetiology of chronic LBP could potentially be safer and more effective than the poor drugs with considerable side effects currently used to treat chronic LBP.

Impact on competitiveness of Europe and participating SMEs

The reduction of the days lost through disability caused by LBP is one of the areas in which Europe could rapidly increase its competitiveness. It has been shown in Australia that public health campaigns to keep moving during an episode of LBP has impact on use of medical services and a

reduction in long term LBP. Thus there are many strategies which could improve EU competitiveness by taking a population approach to the management of LBP – such is the impact of this condition on productivity. Personalized medicine with efficient therapeutic strategies tailored to specific molecular pathology is one of the most important goals for improving the quality and costs of healthcare in the developed world. Diagnostic tests which could stratify patients with complex conditions such as LBP have huge market potential.

Several factors are important for competitiveness of the SMEs. The PainOMICS project used and promoted novel technologies (Glycomics through GENOS and Activomics through Photeomix). The SME partners in the project are in a favourable position to achieve financial gain with production of screening technology components, manufactured for wide distribution both on the markets in their own country, as well as on the global market. Another benefit may be the development of the "pain" network on a long-term basis. These are all factors that will assist in reaching competitiveness in this area of both financial and medical importance.

Synergy with other national or international research activities

The PainOMICS consortium gathers together scientists from different disciplines: geneticists, biologists, biochemists, clinicians and statisticians and have facilitated the relations and exchanges also with other initiatives, such as the EU FP7 project HighGlycan, and the exchange and sharing of data from the CHARGE consortium and the UK BioBank. In addition, PainOMICS partners have decided to go beyond FP7 rules in terms of data management. Namely, in our efforts to publish the data produced by the consortium we took into consideration the H2020 requirements of FAIR (Findable, Accessible, Interoperable, Reusable) data. To this end we have deposited the Activomics and glycomics data generated for the retrospective study on Zenodo as part of the PAIN-OMICS community we have created at https://zenodo.org/communities/painomics/. We have also shaped the development of the GWAS archive, a browsable interface to GWAS results which can be found at http://www.gwasarchive.org, where one can find summary statistics of three GWAS of back pain. The genotype data generated by this consortium together with the related clinical information will be uploaded to the European Genome-phenome Archive.

Impact on society in general

The situation with patients who have LBP has been more and more serious in past years, as reflected in its rise up the Lancet ranking of global burden of disease (2016). It is estimated that the total cost of LBP in Europe (including lost wages and reduced productivity, as well as medical costs) are over Euros 100 billion per year, which presents an enormous financial burden. In fact, this burden is measurable in terms of gross domestic product, so large it is (1-2% for most EU countries).

PainOMICS aims to develop new diagnostic tools for early and reliable identification of patients at high risk for developing chronic LBP and guidance regarding the most suitable forms of treatment. In addition, PainOMICS has enabled insight into molecular events involved in the development of these conditions. These results can be used to produce targeted therapeutics that will slow down or arrest the progression of acute LBP into a chronic condition. If done effectively, this could translate into a meaningful reduction of healthcare costs related to both the diagnosis and the therapeutics treatment. With the costs to society of chronic LBP being so high, any effective intervention which is efficiently rolled out to the general population will have a huge financial impact. Such work will require a coordinated effort from multiple stakeholders as well as a much raised profile of LBP in the minds of both the public and politicians who determine the spend. The EU has led the way by funding this project and we have generated insights into the mechanisms of LBP. Individual countries now need to give priority to research implementation if the benefits of our work are to be translated into practical results of benefit to society.