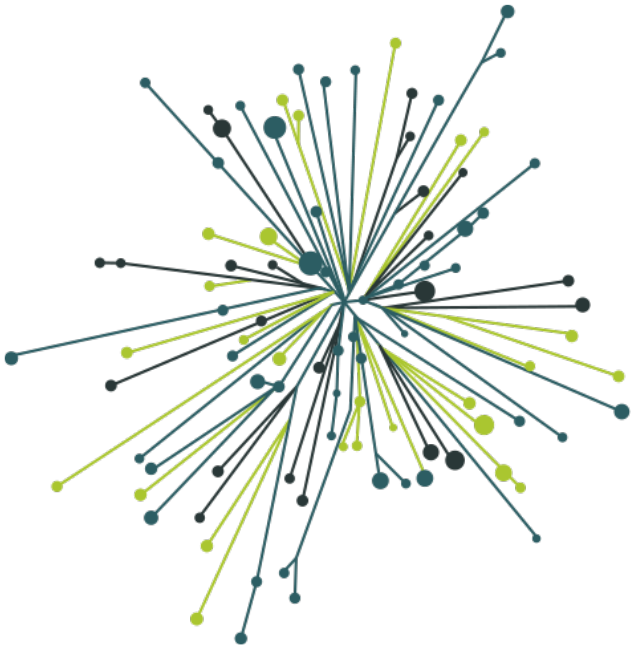

18º meeting

portugaliæ genetica

22 & 23 March 2018 | Porto

Abstract Book

Speaker´s Abstract

Evolution of gene dosage on Z chromosomes

Beatriz Vicoso

Institute of Science and Technology Austria, IST Austria

The haploidy of sex-linked genes in the heterogametic sex should lead to reduced gene dosage, and cause imbalances in gene networks that involve sex-linked and autosomal genes. The presence of complex mechanism of dosage compensation in fruitflies, mammals and nematodes, which globally target the X and reestablish optimal gene expression, supports the idea that these X/autosome imbalances are generally deleterious. Curiously, many female-heterogametic species (including birds, snakes, schistosome parasites and some moths) appear to lack a global mechanism to compensate the expression of the single Z-chromosome of females. Why reduced Z dosage in females should be permissible is still under debate. We have revisited two of these cases, Lepidoptera and Schistosoma, and used comparative genomics and transcriptomics to show that compensation is more widespread than previously suggested. Our results suggest a more complex scenario than a ZW versus XY dichotomy, and offer new avenues of research into the evolution of dosage compensation.

The feedback loop between aging and aneuploidy

Elsa Logarinho

i3S/ IBMC, Porto, Portugal

Aneuploidy, an abnormal chromosome number, has been linked to aging and age-associated diseases, but the underlying molecular mechanisms remain unknown. Through direct live-cell imaging of young, middle-aged, and old-aged primary human dermal fibroblasts, we found that aneuploidy increases with aging due to general dysfunction of the mitotic machinery. Increased chromosome segregation defects in elderly mitotic cells correlated with an early senescence-associated secretory phenotype (SASP) and repression of Forkhead box M1 (FoxM1), the transcription factor that drives expression of most G2/M genes. By restoring FoxM1 levels in elderly and Hutchison-Gilford Progeria Syndrome fibroblasts we prevented aneuploidy and, importantly, ameliorated cellular phenotypes associated with aging. Moreover, we found senescent fibroblasts isolated from elderly donors' cultures to be significantly aneuploid, and aneuploidy to be a key player in the progression into full senescence phenotypes. Based on this feedback loop between cellular aging and aneuploidy, we propose modulation of mitotic efficiency through FoxM1 as a potential strategy against aging and progeria syndromes.

When secondary comes first: the role of DNA secondary structures in mitochondrial genomes

Filipe Pereira

CIIMAR, Porto, Portugal

In addition to the famous double helix, DNA can fold into various other inter- and intramolecular secondary structures that are known to regulate diverse cellular functions and contribute to genetic instability. The non-B DNA structures are deviations to the orthodox right-handed B-DNA, often called non-canonical, unusual, alternative or secondary DNA structures. The most common types of non-B DNA structures are hairpin, cruciform and cloverleaf-like elements, G-quadruplexes, triplexes, slipped structures, left-handed Z-DNA, bent and sticky DNA. Although long thought to be in vitro artefacts, recent studies demonstrate that non-B DNA structures are conserved throughout evolution, suggesting their existence in vivo. In this talk, I will describe our investigations on the formation of non-B DNA in the human mitochondrial genome. Our data suggest that non-B DNA conformations are an important piece in the complex puzzle of mitochondrial DNA evolution, regulation and instability.

Copy number variations associated with changes in post-translational regulation in cancer

Pedro Beltrão

EMBL-EBI, Wellcome Genome Campus, Cambridge, UK

Chromosomal rearrangements, despite being detrimental, are ubiquitous in cancer and often act as driver events. The effect of copy number variations (CNVs) on the cellular proteome of tumours is poorly understood. Large scale characterization of patient tumours and cancer cell lines now allow for the study of how CNVs affect post-translation regulatory networks of the cell. We have analysed recently generated proteogenomic data-sets for tumour samples to investigate the impact of CNVs on protein abundances. We found that CNVs are post-transcriptionally buffered in 23-33% of proteins with enrichment for protein complexes. The protein abundances of complex subunits are highly co-regulated and some act as rate-limiting steps of complex assembly, indirectly controlling the abundance of other complex members. We identified 48 such regulatory interactions and experimentally validated AP3B1 and GTF2E2 as controlling subunits. Based on cell line data we found that a gene-signature of protein attenuation is associated with increased resistance to chaperone and proteasome inhibitors. In addition, we have identified CNVs that are associated with steady state changes in kinase signalling of cancer patient samples. Given that cancer cell lines tend to become “addicted” to changes in kinase signalling, as observed in shRNA data from cell lines, we suggest that some CNVs can be used as markers for differential sensitivity to kinase drug treatment. These results highlight the importance of studying post-translational regulation in cancer.

Exploring the role of non-coding mutations in pancreatic diseases through the analysis of the zebrafish pancreas regulome

José Bessa

i3S/ IBMC, Porto, Portugal

Defining when, where and how much genes are transcriptionally expressed is one of the major mechanisms underlying organogenesis and organ function. Non-coding cis-regulatory elements (CREs), spread over large genomic distances, are key elements in gene transcriptional regulation. In recent years, large-scale Genome Wide Association Studies, for diabetes and pancreatic cancer, have uncovered risk alleles that overlap with non-coding sequences. Many of these non-coding sequences are enriched for epigenetic marks usually found in CREs. This highlights a potential association between pancreatic disorders, as diabetes and cancer, with mutations in non-coding CREs. How these mutations can contribute to disease is still unclear.

Using the zebrafish as a model, we are applying genome wide techniques to uncover the chromatin architecture of pancreatic cells, identify pancreatic CREs and perform mutation analysis, to translate the “non-coding genome” to phenotype.

Chasing the elusive origins of novel enhancers

Kohtaro Tanaka[1], Marília Santos[1], Luís Baudouin-Gonzalez[1], Élio Sucena[1,2]

[1] Instituto Gulbenkian de Ciência, Oeiras, Portugal
[2] Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

Over the past decades, it has become well established that pairs of duplicated genes generally show both overlapping and divergent expression domains. However, important questions regarding the mechanisms underlying this pattern remain: what are the evolutionary trajectories leading to divergent gene expression and what are the underlying cis-regulatory changes? The increasing number of genome sequences from closely related species now allows detailed studies of recent duplication events to address these issues. We have been examining the gene-regulatory evolution of two tandem duplicates, the *Drosophila* Ly6 genes CG9336 and CG9338, which arose at the base of the drosophilids between 40 and 60 Mya ago. Despite their relatively recent origin, the paralogs in *D. melanogaster* have diverged in embryonic tissue-specificities from each other as well as from the unduplicated ortholog found in related species. In this talk, we focus on the neofunctionalization of CG9338, which exhibits a unique expression profile in hemocytes. We will show how this novel expression domain arose through independent emergence of enhancer elements. We also present our data aiming at identifying the causative sequence changes underlying its novel expression pattern and speculate on the mechanism behind its emergence.

Genetic and evolutionary determinants of immune variability: a population genetics view of the response to infection

Maxime Rotival

Institut Pasteur, CNRS URA3012, Paris, France

Infectious diseases have been a major cause of human mortality, so natural selection is expected to act strongly on host defence genes. This is particularly expected for innate immunity genes, as they represent the first line of host defence against pathogens. Here, I will present different cases of how some of these genes and the pathways they trigger have been targeted by natural selection, in its different forms and intensities, helping to delineate genes that fulfil essential functions in host defence, with respect to those exhibiting higher immunological redundancy. I will further discuss how population-specific genetic variation can impact immune-related molecular phenotypes, such mRNA expression upon infection. Using expression quantitative trait loci (eQTL) mapping and RNA-sequencing, we identify multiple cis- and trans-eQTL that contribute to the marked differences in immune responses detected within and between populations of African- and European-ancestry, including a TLR1 master regulator that decreases expression of pro-inflammatory genes in Europeans only. These results show that regulatory variants have been privileged targets of natural selection, uncovering evolutionarily advantageous mechanisms, such as attenuated inflammation. Finally, I will present population genetic analyses showing how admixture with populations presenting distinct demographic and adaptive histories has shaped the contemporary diversity of innate immunity genes, through preferential introgression of specific genes and loci associated with advantageous phenotypes, including increased resistance to pathogens. Together, these studies increase our understanding of immunological and evolutionary mechanisms that have been crucial for our past and present survival against infection.

An interactomics approach to DJ-1 function and role in oxidative stress

Sandra Anjo[1,2], Matilde Melo[1], Andreia Ferreira[1,3], Nuno Jordão[1,3], Liliana R. Loureiro[1,3], Lúcia Sabala[1,3], Pedro Castanheira[4], Mário Grãos[1,4], Bruno Manadas[1]

[1] Center for Neuroscience and Cell Biology, University of Coimbra, Coimbra, Portugal
[2] Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal
[3] Department of Chemistry, University of Aveiro, Aveiro, Portugal
[4] Biocant, Biotechnology Transfer Association, Cantanhede, Portugal

Corresponding author: bmanadas@gmail.com

The PD- associated protein DJ-1 has been recognized as a redox response protein with an important role in the protection against the oxidative stress insults. Moreover, several functions have been reported for DJ-1 which have contributed with significant insights into the mechanisms of oxidative stress-related neurodegeneration observed in PD patients. Considering all these evidences, DJ-1 was considered as a multifunctional protein although its exact mechanism of action and regulation remains largely unknown. It is therefore expected that by elucidating DJ-1 physiological function important insights into PD will also be achieved. Moreover, oxidative stress conditions seem to be crucial for DJ-1 activity, being associated with the regulation of the majority of its mechanisms of action and the interactions established by it. However, all the exact mechanisms behind oxidative stress regulation of DJ-1 and many of the interaction partners of DJ-1 remain unidentified.

This project proposed to elucidate the mechanisms of action of DJ-1, more specifically, aimed to clarify the mechanisms by which DJ-1 modulates short-term and long-term neuroprotection, by identifying the binding partners of DJ-1 during oxidative stress and the dynamics of these interactions under oxidative stress.

A comprehensive interactomic study of DJ-1 under oxidative stress conditions was conducted, allowing the identification and quantification of proteins involved in the network of interactions established via DJ-1 activity. Moreover, this study was performed in a well-defined model associated with the activation of two key survival pathways central for DJ-1 neuroprotective function (ERK1/2 and PI3-K/Akt pathways), and with important implications for the study of DJ-1 role in Parkinson's Disease context, since an evident impact in the mitochondrial activity was also reflected in the experimental conditions used. More importantly, the optimized SWATH-MS based pipeline developed consisting in the use of a high reproducible and efficient in-gel digestion methods, designed as short-GeLC, combined with the use of recombinant proteins as internal standard, was applied to characterize the dynamic nature of the interactions.

In conclusion, this study resulted in the identification of several novel proteins involved in DJ-1 network contributing to the elucidation of DJ-1-mediated neuronal protection against oxidative stress and pointing out some new mechanisms. In addition, many of the proteins identified are well established in distinct cellular functions implicated in PD. Thus, these

results may also contribute to a better understanding of PD and the distinct pathways involved in the establishment and progression of the disease.

Function and dynamics of low complexity regions in proteins

Sandra Macedo-Ribeiro

Biomolecular Structure and Function Group, IBMC/i3S

Low complexity regions (LCRs) are sequences containing a reduced diversity in amino acid composition that are extremely abundant in eukaryotic proteomes. These simple sequences, often lacking a defined three-dimensional structure, play critical functions in diverse protein interaction networks as a result of their intrinsic conformational plasticity. The most studied LCRs are pure single amino acid tandem repeats that are present in ~20% of the human proteome. A number of those proteins are associated with severe human neurodegenerative diseases caused by expansion of stretches of glutamine repeats. The talk will discuss the structural features of polyglutamine repeats and their role on the structure, dynamics and aggregation of ataxin-3, the protein implicated in Machado-Joseph disease.

Genetic diversity in *Mycobacterium tuberculosis*: linking genotypes to phenotypes

Margarida Saraiva

i3S/ IBMC, Porto, Portugal

Tuberculosis (TB) is caused by bacteria belonging to the *Mycobacterium tuberculosis* complex (MTBC), which includes seven phylogenetically distinct lineages associated with specific geographic regions in the world. Thus, the genetic diversity of *M. tuberculosis* is higher than once expected. Furthermore, a growing body of evidence show a functional relevance to this diversity, both in what regards the triggering of the immune response and important clinical features of TB. However, definite studies linking *M. tuberculosis* genomes, immune phenotypes and clinical outcomes are needed. Since covering the full diversity of *M. tuberculosis* in nature is virtually impossible, we devised an experimental approach started in the study of the interactions of clinical isolates of *M. tuberculosis* obtained in the region of Porto with immune cells of geographically matched human donors, relating the obtained findings with the clinics of the TB patient and then moving to the mouse model for mechanistic studies.

In what regards the phylogeographycal structure of *M. tuberculosis* in Porto, we show a highly homogeneous phylogenetic structure of M. tuberculosis, with nearly all cases (96.7%) belonging to Lineage 4 (L4). Within the L4 clade, the most represented sublineage was LAM (65.1%). At the functional level, we found that *M. tuberculosis* isolated from patients with more severe forms of TB are generally poor triggers of the host immune response. Mechanistically, we relate this poor induction of cytokine production with a differential capacity of the different clinical isolates in activating the host inflammasome. Using the mouse model of infection, we found striking differences in the in vivo progression of infection by high versus low cytokine- inducing *M. tuberculosis* isolates, with low cytokine inducing *M. tuberculosis* isolates disseminating more than the high inducing ones.

Our studies are for the first time establishing clear bacterial genotype- immune phenotype-clinical links, offering host immune pathways and bacterial signatures to be targeted in the design of novel TB interventions.

Oral communications

Mitochondrial DNA evolutionary breakpoints are associated with non-B DNA conformations

J. Carneiro[1], M. Bernt[2,3], F. Pereira[1]

[1] Interdisciplinary Centre of Marine and Environmental Research (CIIMAR), University of Porto, Porto, Portugal
[2] Bioinformatics Group, Department of Computer Science University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany
[3] Parallel Computing and Complex Systems Group, Department of Computer Science, University of Leipzig, Augustusplatz 10, D-04103 Leipzig, Germany

Corresponding author: joaomiguelsov@gmail.com

Genome rearrangements are driving forces of evolution and are key events in the development of many diseases. Contiguous ancestral regions of mitochondrial genomes can evolve by a series of rearrangement events, such as inversions, transpositions, and tandem duplication random loss (TDRL). The exact mechanism by which such rearrangements are generated is unknown. It has been shown that non-B DNA conformations (deviations to the orthodox right-handed B-DNA) are associated with mtDNA deletion breakpoints, mutagenesis and diseases. To understand the process behind rearrangements, we studied the formation of non-B conformations in rearrangement breakpoint regions of Chordata mtDNA. To test the hypothesis of the involvement of these structures in 1621 mtDNA rearrangements, we calculated the free energy of the breakpoint regions midpoint. Control datasets were used to test how significant was the occurrence of non-B conformations. We observed that the free energy mean value of inversions, transpositions, and TDRL were respectively -13.81, -17-58 and -15.64 kcal/mol. Non-B conformations with lower free energies (i.e., more stable structures) were confirmed in a sliding window analysis that showed a decrease in the folding energy of the breakpoint regions. The analysis of 6261 breakpoint regions showed that non-B DNA conformations may be related with evolutionary rearrangement scenarios.

Funding: Investigator FCT program to FP and FCT research grant SFRH/BPD/100912/2014 to JC.

Contrasting molecular adaption in Slo1 and Slo3 channel subfamilies

A. Vicens[1,3]*, P. Vinuesa[2], M. Arenas[3], C. Treviño[1]

[1] Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología, Universidad Nacional Autónoma de México
[2] Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México
[3] Department of Biochemistry, Genetics and Immunology. University of Vigo, 36310 Vigo, Spain

*Corresponding author: avicens859@gmail.com

Slo1 and Slo3 encode close paralogues of the Slo potassium (K+) channels family. Despite their evolutionary relatedness, Slo1 and Slo3 channels show marked functional differences. However, the molecular mechanisms and selective pressures behind the evolution of Slo1 and Slo3 channels are unknown. In this study, we analyze the molecular evolution of Slo1 and Slo3 channels and we found that they were overall subjected to strong purifying selection at the global level, although we also identified more relaxed purifying selection and some positively selected sites in Slo3. Next, we found that more variable regions containing positive selection sites in Slo3 gating ring are likely to promote functional differences among Slo3 orthologues. Faster divergence of Slo3 channel might be attributed to episodes of positive selection in a background of relaxed selection promoted by its male-biased condition, as male-biased genes show high rates of evolution often caused by non-adaptive processes. These findings can be used to explain the functional diversity of sperm ion channels among species and to provide new insights into the mechanisms governing the evolution of reproductive proteins.

Accounting for protein stability in sequence evolution: Applications to Phylogenetics

M. Arenas[1], U. Bastolla[2]

[1] Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain.
[2] Centre for Molecular Biology "Severo Ochoa", Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain

*Corresponding author: marenas@uvigo.es

Phylogenetic inference from protein data is traditionally based on empirical substitution models of evolution that assume that protein sites evolve independently of each other and under the same substitution process. However, it is well known that the structural properties of a protein site in the native state affect its evolution, in particular influencing the sequence entropy and the substitution rate. In order to incorporate the protein structure into phylogenetic frameworks we developed stability constrained substitution models that explicitly consider the stability of the native state against both unfolded and misfolded states and we implemented these models into the freely available evolutionary frameworks ProteinEvolver and ProtASR. The first framework simulates protein evolution over a given phylogeny under site-dependent substitution models and the second one applies an independent sites approximation that considers structural constraints to reconstruct ancestral protein sequences. Here we present both frameworks and we describe their validation, with simulated and real proteins, with respect to empirical models that lack site-specificity.

Funding: This work was funded by the Spanish Government, grants BIO2016-79043, BFU2012-40020 and RYC-2015-18241.

Metagenomic composition analysis of ancient DNA samples

D. Pratas*, A. J. Pinho, R. M. Silva, J. M. O. S. Rodrigues, M. Hosseini, T. Caetano, P. J. S .G. Ferreira

IEETA, DETI, DMS, iBiMED, Dbio, CESAM @ University of Aveiro

*Corresponding author: pratas@ua.pt

The detection and quantification of metagenomic composition in ancient DNA are based on the alignment of the reads, according to an existing database with reference microbial whole genomes. Without proper parameterization, these methods are not suitable for ancient DNA. The quantification of somewhat dissimilar sequences by alignment methods is problematic, due to the need for fine-tuned thresholds, considering relaxed edit distances and the consequent increase in computational cost. Moreover, the appropriate choice of the thresholds emerges the problem of how to quantify similarity without producing overestimated measures.

We use a compression-based method to infer the metagenomic composition of nextgeneration sequencing samples. The unsupervised alignment-free method runs efficiently on FASTQ samples, quickly learning how to give importance to the models that cooperate to predict similarity, incorporating parallelism and flexibility.

We show substantial identification capabilities without overestimation. We give some case examples using our method, namely on a Neanderthal (gDNA), Denisova (gDNA) and a Columbian Mammoth (mtDNA). For example, in the Denisova, we show an application using human Adenovirus, while in the Mammoth mitogenome we found and authenticated an ancient Pseudomonas bacteria.

Poster sessions

P01: Genetic characterization of the male lineages present in Nigeria population

A. Campos[1,2,3], B. Martinez[4], L. Catelli[5], M. Romero[5], V.O. Okolie[6], S.O. Keshinro[7], E.F. Carvalho[8], C. Vullo[5], L. Gusmão[2,3,8], V. Gomes[2,3]

[1] FCUP, Faculty of Sciences, Department of Biology, University of Porto, Portugal
[2] IPATIMUP, Institute of Pathology and Molecular Immunology from University of Porto, Portugal
[3] i3S, Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
[4] Molecular Genetics Lab, Institute for Immunological Research, University of Cartagena, Cartagena, Colombia
[5] DNA Forensic Laboratory, Argentinean Forensic Anthropology Team (EAAF), Córdoba, Argentina
[6] Department of Cell Biology and Genetics, University of Lagos, Lagos, Nigeria
[7] FCIID Annex, Nigeria Police Force, Lagos, Nigeria
[8] DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil

Africa was the place of origin of modern humans. Therefore, the genetic study of their populations is very important for population genetics. Nigeria is the most populous African country, containing several ethnic groups, mainly Hausa, Yoruba and Igbo, and 510 living languages were estimated to exist there. The Y chromosome is exempt of recombination in most of its length, and this region was designated as Male Specific region (MSY). Hence, it is possible to trace male lineages that are transmitted unchanged from father to son, except if mutations occur. The aim of this work was to characterize the male lineages in Nigeria and to understand if this population shows genetic differences with neighbouring populations. For that, we studied 27 Y-STRs included in the Yfiler Plus kit in 358 Nigerian males (135 Yoruba, 136 Igbo and 87 Hausa). A subset of 31 were also genotyped for 28 Y-SNPs included in three multiplexes. A total of 351 unique haplotypes were found, corresponding to a very high haplotype diversity (0.9999 ± 0.0002). The Hausa ethnic group that speak an Afro-Asiatic language showed significant genetic differences with the other two Nigerian populations that belong to another language inside of Niger-Congo family. Furthermore, even if Yoruba and Igbo belong to the same language family, significant genetic distances between them were also found. 71% of the samples belong to haplogroup E, being the sub-haplogroup defined by U174 (E-U174) the most frequent.

P02: Assessing the impact of Copy Number Variation (CNVs) on severe spermatogenic impairment with exome data

A. M. Lopes[1,2]*, L. Nagirnaja[3,4], F. Carvalho[1,5], J. Gonçalves[6], S. Fernandes[1,5], I. Pereira-Caetano[6], K. Almstrup[7,8], E. Rajpert-De Meyts[7,8], S. Seixas[1,2], B. Houston[9], A. Barros[1,6], M. O'Bryan[9], K. I. Aston[10], D. F. Conrad[3,4] on Behalf of the GEMINI Consortium

[1] i3S - Instituto de Investigação e Inovação em Saúde, Porto, Portugal
[2] IPATIMUP - Institute of Molecular Pathology and Immunology, University of Porto, Portugal
[3] Dep. of Genetics, Washington University School of Medicine, St. Louis, USA
[4] Dep. of Pathology and Immunology, Washington University School of Medicine, St. Louis, USA
[5] Dep. de Genética, Faculdade de Medicina da Universidade do Porto, Portugal
[6] Dep. de Genética Humana, Instituto Nacional de Saúde Dr. Ricardo Jorge, Lisboa, Portugal
[7] Dep. of Growth & Reproduction, Rigshospitalet, Copenhagen, Denmark
[8] University of Copenhagen, Denmark
[9] The School of Biological Sciences, Monash University, Australia
[10] Dep. of Surgery, University of Utah, USA

* Corresponding author: alopes@ipatimup.pt

Azoospermia, the most severe form of male infertility, affects approximately 1% of men worldwide and in the great majority of the cases the etiology of the disease remains unidentified. Given the large number of genes involved in spermatogenesis it is likely that a proportion of cases of idiopathic azoospermia have a genetic basis. We have previously described, using SNP arrays, an excess of low frequency copy number variants (CNVs) in both the autosomes and the sex chromosomes in non-obstructive azoospermia (NOA) suggesting a heterogeneous genetic etiology for this condition.

As an ongoing effort to characterize the genetic architecture of male infertility, the Genetics of Male Infertility Initiative (GEMINI) consortium sequenced in a first phase the whole exome of 422 well phenotyped NOA cases from 3 countries and 84 controls. CNV calling was successfully performed in 315 NOA and 68 normozoospermic controls using XHMM (eXome-Hidden Markov Model). We found a similar rate of autosomal rare deletions (MAF<1%, all sizes) in NOA and normozoospermic controls (approximately 1 deletion per individual). Only a few large and highly disruptive CNVs were found in NOA, however these patients present a higher number of smaller deletions in genes involved in spermatogenesis, compared to normozoospermic controls. CNV calling from exome data has less power than from whole-genome but can contribute to the identification of genes affected in genetically heterogeneous diseases such as NOA.

P03: Taquicardia ventricular polimórfica catecolaminérgica associada a variações nos exões 3, 93 e 103 do gene RyR2

A. Abreu[1], B. Silva[3], A. Santos[2,3], M. J. Porto[3], M. Pancorbo[4], L. Cainé[2,3]

[1] Aluna de Doutoramento em Ciências Forenses da Faculdade de Medicina da Universidade do Porto
[2] FMUP – Faculdade de Medicina da Universidade do Porto
[3] INMLCF – Instituto Nacional de Medicina Legal e Ciências Forenses, IP
[4] BIOMICs – Bioscience & Microtechnologies Research Group, Universidad del País Vasco

A Morte Súbita Cardíaca (MSC) representa um problema de saúde pública que tem assumido uma crescente visibilidade, gerando uma preocupação crescente na população em geral e uma tomada de consciência da necessidade de um diagnóstico de autópsia preciso, com informação genética que possa ser útil na prevenção de outras mortes em familiares diretos. Quando se considera a MSC em jovens (≤ 40 anos) estima-se que 30% de todas as mortes não traumáticas são causadas por uma variedade de distúrbios genéticos. Estes não causam qualquer alteração estrutural no coração e nenhuma causa de morte é geralmente identificada na autópsia médico-legal sendo esses casos designados como casos de Morte Súbita Inexplicável (MSI). Admite-se que cerca de 15% dos indivíduos com MSI apresentam Taquicardia Ventricular Polimórfica Catecolaminérgica (TVPC) e 60% desses casos estão relacionados com mutações autossómicas dominantes no gene RyR2 no qual se estima que 65% das mutações podem ser detetadas pelo rastreio de 16 dos seus 105 exões. Neste estudo foi definido que se estudariam 3 destes 16 exões, os exões 3, 93 e 103 sendo o principal objetivo desta investigação padronizar um protocolo experimental para a análise de possíveis variações genéticas nestes exões.

P04: Forensic mixtures from crime scenes: Analysis and comparison of software results

C. Figueiredo[1,2,3,4]*, P. M. Ferreira[4], S. Costa[4], N. Pinto[2,3,5]

[1] FCUP – Faculdade de Ciências da Universidade do Porto
[2] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto
[3] IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto
[4] LPC-PJ - Laboratório de Polícia Científica, Polícia Judiciária
[5] CMUP – Centro de Matemática da Universidade do Porto

*Corresponding author: carolina_figueiredo_@hotmail.com

Crime samples are often characterized by low quantity and quality of DNA. Besides that, it is common the presence of more than one contributor, making the genetic analysis more difficult. Mainly on complex samples analysis, it is important to quantify the probative value and the reliability of the results. The generally accepted method is the computation of likelihood ratios (LR), comparing the probabilities of the observations under the alternative hypotheses of the prosecution and defense.

Currently, there are computer programs for statistical analysis of forensic samples based on LR, differing mainly on the information used: continuous models use quantitative information (peak heights) of the electropherogram, while qualitative models only use qualitative information.

Our aim is to compare the results obtained from different software to the same set of samples under similar specifications, measuring the LR differences. Real casework samples with at least two or three contributors were analyzed through Euroformix (continuous) and LRMix Studio (quantitative) computer programs.

Additionally, we weighted the impact on the LR results of varying some parameters related to number of possible contributors, Fst correction, drop-in, drop-out and detection threshold.

In this work we will present the results obtained by the two referred software when the number of possible contributors (given by the user) varies.

P05: Influence of admixture and last Ice Age on genetic diversity gradients of the Americans

C. Branco[1]*, M. Arenas[1]

[1] Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain

*Corresponding author: ciabranco94@gmail.com

Recent studies showed that genetic gradients can be influenced by diverse evolutionary processes such as population admixture and population range contractions. Some genetic gradients were observed in the Americas although their specific evolutionary causes were not explored so far. Using extensive computer simulations, here we evaluated the impact of admixture between first expansion waves of modern humans and the presence of ice-sheets derived from the last glacial maximum, on the genetic gradients (detected by principal component analysis) of the entire continent, North America and South America. The specific analysis of North and South America showed that genetic gradients are usually orthogonal to the direction of range expansions –either initial expansions or posterior expansions to recolonize northern regions after ice sheets melting– and we suggest that they result from allele surfing processes. Conversely, our results on the entire continent show a northwest-southeast gradient obtained with any scenario, which we interpreted as a main consequence of isolation by distance along the long length of this continent. These findings suggest that distinct genetic gradients can be detected at different scales of the Americas and that subcontinent regions present gradients more influenced by evolutionary and environmental processes than the whole continent.

P06: Characterization of male lineages in the Ashaninka from Peru

G. Aragão[1,2,3], V. Gomes[2,3], D. Herman Tineo[4,5], S. Loiola[6], L. Gusmão[2,3,6], M. J. Prata [1,2,3]

[1] Faculty of Sciences, Department of Biology, University of Porto, Portugal
[2] Instituto de Investigação e Inovação em Saúde, University of Porto, Portugal
[3] Institute of Molecular Pathology and Immunology of the University of Porto, Portugal
[4] Universidad Nacional Mayor de San Marcos, Lima, Peru
[5] Instituto de Medicina Legal del Perú, Lima, Peru
[6] DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil

Peru is home to many populations that experienced high admixture after the European colonization, showing variable degrees of European, African and Native American genetic contributions. Due to the vastly diverse ecosystem of Peru, most communities of native populations are fairly isolated from one another. That is the case of the Peruvian communities of Ashaninka, one of the Native American ethnic groups that suffered less admixture. The Ashaninka span across a large variety of territories, being one of the most numerous native populations in the western Amazon. In this study, the aim is to obtain a high resolution characterization of the patrilineal lineages of the Ashaninka. In order to accomplish this, a panel of 10 SNPs from the Y chromosome were included in a newly developed Multiplex system (Multiplex Q). 84% of the already tested samples (n=31) were successfully genotyped with this multiplex, which of these, all were positive for the mutation M3 contained in this multiplex. This mutation defines the sub-clade of Q that is typically found among Amerindians. Analysis of SNPs downstream of Q-M3 is revealing the presence of a few sub-clades, although at a moderate frequency; among them the most frequently found was the sub-lineage SA05. The remaining samples (16%), are being characterized with other multiplexes, which predictably will provide estimates on the relative proportions of European and African lineages that were introgressed into the Ashaninka.

P07: Are MJD proteases playing a role in MJD by compensating the partial loss-of-function of ATXN3?

I. P. D. Costa[1,2]*, S. Seixas[1,2], A. M. Lopes[1,2], J. Sequeiros[1,3,4], A. Amorim[1,2,5], S. Martins[1,2]

[1] Instituto de Investigação e Inovação em Saúde, Universidade do Porto (i3S), Porto, Portugal
[2] IPATIMUP - Institute of Molecular Pathology and Immunology of the University of Porto, Porto Portugal.
[3] IBMC – Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal
[4] ICBAS – Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, Portugal
[5] FCUP - Faculdade de Ciências da Universidade do Porto, Porto, Portugal.

*Corresponding author: icosta@ipatimup.pt

The Machado-Joseph disease (MJD) subfamily of deubiquitinases comprises ataxin-3 (ATXN3), ataxin-3 like (ATXN3L) and Josephin domain-containing proteins 1 (JOSD1) and 2 (JOSD2). Much attention has been given to ATXN3 after the identification of an exonic (CAG)n, responsible for the autosomal dominant neurodegenerative MJD when encoding a polyQ tract above 61 glutamines. The other members of this protease family are poorly explored; in particular, ATXN3L, remained in the shadow under the assumption that it would be a non-functional sequence due to its lack of introns. The partial loss-of-function mechanism suggested for MJD pathogenesis, together with the proved deubiquitinating activity of all four MJD family members led us to hypothesize that they may exert a neuroprotective role in MJD.

We carried out an analysis of ATXN3L, JOSD1 and JOSD2 expression patterns in a panel of 23 human tissues from healthy individuals. The evolutionary conserved paralogue of ATXN3 - ATXN3L, is transcribed in testis, placenta and brain. Interestingly, ATXN3L transcription seems to differ across brain regions, being highly expressed in the cortex, followed by substantia nigra and residually in the cerebellum. In contrast, ATXN3, JOSD1 and JOSD2 were found to be ubiquitously expressed. The observation of distinct ATXN3L expression levels in brain regions differently affected by MJD may be relevant for the observed large clinical variability not explained by the (CAG)n expansion size.

P08: Seeking species specific compensatory episodes on *Mus Musculus*

I. Sereno[1,2,3]*, L. Azevedo[1,2,3], A. Amorim[1,2,3]

[1] FCUP – Faculdade de Ciências da Universidade do Porto, Portugal
[2] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[3] IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Portugal

*Corresponding author: isereno@ipatimup.pt

Missense mutations are the most common cause of human diseases. Assuming that the genes involved in genetic diseases are maintained conserved to preserve function, it would be reasonable to expect that these would also cause disease in close species. However, a significant proportion of these mutations are found as wild-type allele in non-human species. It has been assumed that the tolerance to these mutations may be possible due to epistatic interactions between the pathogenic mutation and other residues of the same or interacting proteins. Such compensatory episodes generally occur within the same protein although the compensatory partner(s) is only known for a few cases.

We analyzed a series of human disease-associated mutations found as the wild-type allele in the mouse (*Mus musculus*) to find the most likely compensatory residues. Using homology modeling we located the most suitable compensatory candidate(s) before final validation by experimental procedures. Our preliminary results follow the general trend that compensatory episodes occur most likely in close proximity to the deleterious mutations, although the compensatory effect may not be directed in all the cases by a single residue but rather by a network of interacting residues.

This project will provide clear examples of the compensatory mechanism which is expected to be of great relevance in the development of future therapeutic strategies and in the development of animal models of severe/lethal human diseases.

P09: Mitochondrial genome of the filamentous fungus *Neurospora crassa*

J. Monteiro[1,2]*, A. Videira[2,3,4], F. Pereira[1]

[1] Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR), Universidade do Porto, Porto, Portugal.
[2] ICBAS-Instituto de Ciências Biomédicas de Abel Salazar, Universidade do Porto, Porto, Portugal.
[3] IBMC-Instituto de Biologia Molecular e Celular, Universidade do Porto, Porto, Portugal.
[4] i3S-Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto Portugal.

*Corresponding author: joaopmonteiro12@gmail.com

The filamentous fungus *Neurospora crassa* is a popular model organism, used in a wide variety of research areas. Its utility derives from its easy maintenance and handling, rapid growth on simple media, tractable life cycle of haploid nature, is an obligate aerobe and has more than twice the genes of yeast. The existence of a high-quality nuclear genomic sequence and knockout strains of almost all genes turns this organism suitable for a wide range of biochemical and genetic studies. Mitochondria are cytoplasmic organelles involved in essential cell processes, with their own genetic material, mitochondrial DNA (mtDNA). The mtDNA of fungi is highly variable in length, organization and content. Here we analyze the mitogenome of *N. crassa* and compare it with other fungi mtDNA. Fungal mtDNA can have large intergenic regions, including sequence repeats and introns, leading to length variation. The size of fungi mtDNA range from 19.4 kb in *S. pombe* to 127 kb in *A. bisporus*, whilst *N. crassa* mtDNA has about 65 kb. Variations in the number of genes can also be found. *N. crassa* mtDNA encodes 28 protein coding genes, 2 rRNAs and 28 tRNA genes, whereas *S. pombe* encodes only 8 protein coding genes, 2 rRNAs and 25 tRNA. Also, some genes are absent or inactive, such as the atp9 gene in *P. anserina* and *N. crassa*. The characterization of *N. crassa* mtDNA will help to delineate new experiments of mitochondrial dysfunction and evolution.

P10: Structural and molecular underpinnings of variable obesogen sensitivity in vertebrates

M. Lopes-Marques[1]●, A. M. F. Capitão[1,2]●, R. F. Ruivo[1], E. S. S. Fonseca[1,2], M. M. Santos[1,2]*, L. F. C. Castro[1,2]*

[1] Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIMAR), University of Porto (U.Porto), Matosinhos, Portugal
[2] Faculty of Sciences (FCUP), Department of Biology, University of Porto (U.Porto), Porto, Portugal

●Equal contribution

* Corresponding author: filipe.castro@ciimar.up.pt; santos@ciimar.up.pt

Obesity is a menacing global pandemic, linked to numerous health conditions and high mortality rates. This condition may be triggered by nutritional variations, overconsumption, individual genetic background and/or environmental cues. Of note is the role of a group of compounds identified as Obesogens, which cause fat accumulation. Obesogens are known to modify homeostatic lipid metabolism through the modulation of nuclear receptor proliferator-activated receptor (PPAR). Obesogens such as Tributyltin (TBT) exploit PPARγ by mimicking the endogenous ligands leading to the erroneous activation or repression of downstream lipid metabolizing genes. Although adverse effects of TBT exposure have been documented in many species, some species show variable sensitivities to TBT exposure. To understand what lies behind this variability we isolated and characterized PPARγ from several vertebrate species placed in key phylogenetic positions. Here, using comparative homology modelling we uncovered the critical cues in the ligand binding domain that may account for the variable sensitivity to TBT observed in several vertebrates.

Funding: This work was supported by Norte2020 and FEDER (Coral - Sustainable Ocean Exploitation - Norte-01-0145-FEDER-000036). We acknowledge Fundação para a Ciência e a Tecnologia for the support to EF (SFRH/BD/79305/2011).

P11: Insights into Fasciclin-like arabinogalactan proteins involved in plant reproduction from *Quercus* to *Arabidopsis*

M. Couto[1], M. L. Costa[1], M. Oliveira[1,2,3], S. Coimbra[1], M. I. Amorim[1]*

[1] Departamento de Biologia, Faculdade de Ciências, Universidade do Porto
[2] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto
[3] IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto

*Corresponding author: mpamorim@fc.up.pt

Plants have developed elaborate sexual reproductive structures, the flowers, where the male and female sex organs are present: the anther and the ovule. Several fasciclin-like arabinogalactan protein (FLA) genes have been involved in reproductive organogenesis in the model plant Arabidopsis thaliana. FLAs are a sub-class highly glycosylated hydroxyproline-rich glycoproteins, that have glycomodules but also one or two conserved fasciclin domains, which mediate cell-cell and cell-extracellular matrix adhesion. In *Quercus suber* a dominant tree in Portuguese Montado, male flowers develop much sooner than the female flowers providing an interesting system for comparative studies of development and sexual reproduction. In an attempt to find FLA genes differentially expressed in male and female *Q. suber* flowers transcriptome, was used a bioinformatics tools such as, BLAST against *Arabidopsis* FLAs. With this approach we identified 18 QsFLAs-like genes that clustered into four major groups. At least five QsFLAs which clustered in the same clade are differentially expressed in the male and female gametophyte. The characterization of T-DNA insertion mutant lines of Arabidopsis homologues of these QsFLAs genes putatively related to the development of reproductive tissues is been performed.

Funding: FCT/ PTDC/AGR-FOR/3356/2014

P12: A 3' untranslated region element modulates *Drosophila* Polo kinase expression

M. Oliveira[1], P. Pinto[2], A. Jesus[1], C. Lopes[3], J. Freitas[1], C. Conde[4], A. Moreira[1,5,6]*

[1] Gene Regulation, i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[2] Centre for Organismal Studies at Ruprecht-Karls-Universität Heidelberg, Germany
[3] Glial Cell Biology, i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[4] Cell Division & Genomic Stability, i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[5] IBMC-Instituto de Biologia Molecular e Celular, Universidade do Porto, Portugal
[6] ICBAS - Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, Portugal

*Corresponding author: alexandra.moreira@i3s.up.pt

The 3' untranslated region (3'UTR) of polo has two alternative polyadenylation sites (PASs: pA1 and pA2), which results in two mRNAs with the same coding sequence. We showed that the longest mRNA isoform is vital for Polo protein production [1]. Recently, we identified a 22nt U-rich sequence (USE) located upstream of pA1 that is also present in 8% of Drosophila genes. Flies lacking the USE (ΔUSE) have lower levels of Polo protein. This correlates with reduced levels and activation of Aurora B and Mps1, two mitotic targets of Polo kinase [2], at kinetochores in ΔUSE neuroblasts. These findings highlight a role for this conserved nucleotide sequence in the regulation of polo expression. Moreover, using biochemical approaches, we identified Elav/HuR and Heph/PTB, two RNA-binding proteins, as USE-binding proteins. elav and heph mutants have lower levels of polo longest mRNA and lower levels of Polo protein, suggesting a function in polo PAS selection. Taken together, our in vivo results show that polo's expression is controlled by a 22nt USE in the 3'UTR by a mechanism involving Elav and Heph.

1. Pinto, P.A., et al., (2011) EMBO J, 30:2431-44.
2. Conde, C., et al., (2013) EMBO J, 32: 1761-77.

Funding: Norte-01-0145-FEDER-000008-Porto Neurosciences and Neurologic Disease Research Initiative at i3S supported by Norte Portugal Regional Operational Programme (NORTE2020) under the PORTUGAL 2020 Partnership Agreement through the European Regional Development Fund (ERDF).

P13: Genetic characterization of the native population of Asháninka from Peru with autosomal ancestry-informative markers

P. Sousa[1,2,3], D.H. Tineo[4,5], L. Gusmão[2,3,6], V. Gomes[2,3], R. Pereira[2,3,7]

[1] FCUP - Faculdade de Ciências da Universidade do Porto, Porto, Portugal
[2] i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[3] IPATIMUP - Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal
[4] Universidad Nacional Mayor de San Marcos, Lima, Peru
[5] Instituto de Medicina Legal del Perú, Lima, Peru
[6] Laboratório de Diagnóstico por DNA, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
[7] Laboratório de Polícia Científica, Área da Diretoria do Norte da Polícia Judiciária, Porto, Portugal

*Corresponding author: rpereira@i3s.up.pt

Peru is a multi-ethnic country constituted by the combination of different cultures and ethnicities over thousands of years. Native Americans reached the South American territory more than ten thousand years ago, before the Spanish conquest in the 16th century. In fact, the native Amerindian, and the European influence brought by Spaniards, constitute the basis of current Peru genetic landscape. Ancestry-informative markers (AIMs) are useful tools to infer the biogeographic origin of individuals, to estimate genetic ancestry at individual and population level, and to assess population substructure. In this work we analyse the distribution of autosomal ancestry-informative indels in Peruvian Asháninka aiming to study their Amerindian ancestral composition, while also investigating for signs of European influx. A total of 182 samples were genotyped for 46 AIM-indels through multiplex PCR as in Pereira et al. (PLoS ONE 7(1): e29684). Amplified fragments were separated by capillary electrophoresis (ABI 3130 Genetic Analyzer) and genotyping was performed with GeneMapper v4.0. Individual and global ancestry estimates obtained with Structure v.2.3.4 using published data as reference populations put emphasis to the essentially Amerindian ancestry of Asháninka people (average=97.4%), although the European component is also noticeable in some individuals. Our study further contributes to strengthen the coverage of reference data for this set of AIMs, especially for native Americans.

P14: The influence of the different mutation models in kinship evaluation

P. Machado[1], L. Gusmão[2,3,4], E. Conde-Sousa[5], N. Pinto[2,3,5]*

[1] Faculdade de Ciências da Universidade do Porto.
[2] Institute of Pathology and Molecular Immunology from University of Porto (IPATIMUP), Portugal.
[3] Instituto de Investigação e Inovação em Saúde, I3S, Universidade do Porto, Portugal.
[4] DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil.
[5] CMUP, Centro de Matemática da Universidade do Porto, Porto, Portugal.

*Corresponding author: npinto@ipatimup.pt

Different mutation models have been developed considering the genotypic observations of parent(s)/offspring duos (or trios), even though, for autosomal transmission, only Mendelian incompatibilities, not mutations, are able to be identified. The most commonly considered mutation models are the so-called: “equal”, “proportional”, “stepwise” and “extended stepwise”, all implemented in the software familias.

In this work we simulated 100,000 profiles (duos and trios) of parent-child, full-siblings, half-siblings and unrelated individuals, as well as avuncular cases (the uncle being a brother of the father, whose parents are either unrelated, first cousins or full siblings), to be analyzed as standard parent-child. This was done assuming a specific database for 17 autosomal STRs and probabilities of incompatibility inferred from the AABB report, 2008. Using the R version of familias we calculated the likelihood ratio where the probability of the genotypic configuration of the individuals assuming each of the pedigrees was compared with the same probability assuming unrelatedness (and half-sibship, in the case of full-siblings).

The results show that for profiles generated assuming the abovementioned pedigrees, except for unrelated, the use of different mutation models does not result in relevant differences, which also indicates that the consideration of hidden mutations does not have a relevant influence in the result.

P15: Primary ciliary dyskinesia: unveiling CCDC103 expression profiles

R. Pereira[1,3]*, J. Oliveira[2,3], M. Oliveira[2,3], R. Santos[2,3], M. Sousa[1,3]

[1] Laboratory of Cell Biology, Department of Microscopy, Institute of Biomedical Sciences Abel Salazar (ICBAS), University of Porto (UP), Rua Jorge Viterbo Ferreira, 228, 4050-313 Porto, Portugal
[2] Molecular Genetics Unit, Center of Medical Genetics Dr. Jacinto Magalhães (CGMJM), Centro Hospitalar do Porto (CHP), Praça Pedro Nunes, 88, 4099-028 Porto, Portugal
[3] Multidisciplinary Unit for Biomedical Research-UMIB, ICBAS-UP

*Corresponding author: ruterpereira@gmail.com

Motile cilia occur at respiratory airways and sperm flagellum and have vital functions in mucociliary clearance and flagellar motility. These cilia possess a unique structure called the axoneme (Ax). It is composed of associated structures such as dynein arms (DA), and is critical for cilia motility. Defects in Ax gives rise to a primary ciliary dyskinesia (PCD). PCD is an autosomal recessive ciliopathy a clinically heterogeneous disease whose symptoms include chronic bronchitis, organ laterality defects, and reproductive problems. PCD is also complex from a genetic perspective with several genes involved. Previously, we have described an infertile PCD patient with absence of DA in which we found a novel homozygous variant in the coiled-coil domain-containing protein 103 (CCDC103) gene. CCDC103 was suggested to participate in DA assembly. Further, studies in *Chlamydomonas reinhardtii* found that CCDC103 binds to Ax and to MTs and that is critical to stabilizing the MT polymeric structure. Nevertheless, the knowledge about this protein is still scarce

CCDC103 expression analysis by real-time PCR and western blot in a variety of tissues. Results: We observed that CCDC103 is expressed in different reproductive cells and in leukocytes. This protein forms dimers and higher-order oligomers whose size are tissue-specific, which might reflect different assemblies and functions. Overall, our data suggest that CCDC103 might have additional roles in the cilia biology.

P16: Identification of genetic variants contributing to late-onset Alzheimer's disease risk and correlation with cerebral activity of patients

S. Martins[1,2], L. Álvarez[1,2,3], A. M. Lopes[1,2], I. Gomes[1,2], V. Oliveira[4], P. Sousa[4], M. Rodríguez[5], C. Pita[5], M. García[6], J. Poza[6], A. Maturana-Candelas[6], R. Hornero[6], A. Taborda[4], M. Figueruelo[5], M. Arenas[1,2,7], C. Gómez[6], N. Pinto[1,2,8]*

[1] Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Portugal
[2] Instituto de Investigação e Inovação em Saúde (i3S), Universidade do Porto, Portugal
[3] ONCODNA S.A, Gosselies, Belgium
[4] Associação Portuguesa de Familiares e Amigos de Doentes de Alzheimer, Delegação Norte, Portugal
[5] Asociación de Familiares y Amigos de Enfermos de Alzheimer y otras demencias de Zamora, Spain
[6] Grupo de Ingeniería Biomédica, Universidad de Valladolid, Spain
[7] Department of Biochemistry, Genetics and Immunology, University of Vigo, Vigo, Spain.
[8] Centro de Matemática da Universidade do Porto, Portugal

*Corresponding author: npinto@ipatimup.pt

Late-onset Alzheimer's disease (LOAD) is an age-related neurodegenerative disease, with symptoms appearing at 65 years or later. This common form of dementia accounts for 95% of all Alzheimer's cases and is considered to progress in four general stages: mild cognitive decline; initial stage; moderate decline; and severe decline. LOAD seems genetically complex in nature, with the presence of APOE ε4 allele identified as the most relevant genetic risk factor.

In this project, our aims are (1) to identify novel LOAD candidate genes; (2) to characterize the population regarding coding and regulatory variants within genes at the previously identified LOAD loci; (3) to associate different cerebral activities to each of the four disease stages; and (4) to correlate variants in candidate genes and disease progression (assessed by neuroimaging).

We will perform a whole-exome sequencing study in a total of 200 LOAD patients from North Portugal (n=100) and from the Spanish community of Castile and León (n=100), in addition to 50 controls from both regions. Patients will be selected according to their stage of the disease (25 per group). DNA will be extracted from saliva samples, and cerebral activity will be assessed by electroencephalography (EEG).

By correlating EEG patterns and the four stages of the disease, we expect to find a noninvasive approach to assess early stages of cognitive decline, and improving the accuracy of diagnosis.

P17: eQTL analysis of the NAPRT locus

S. Duarte-Pereira[1,2]*, S. Matos[2], J. L. Oliveira[2], R. M. Silva[1,2]

[1] Department of Medical Sciences and iBiMED – Institute of Biomedicine, University of Aveiro, Portugal
[2] IEETA – Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Portugal

*Corresponding author: sdp@ua.pt

Nicotinate phosphorybosyltransferase (Naprt) is a key enzyme in NAD salvage and a potential target in cancer therapy, as Naprt protein is silenced or downregulated in many tissues and tumors. Whether this is due to mutations in the gene is largely unknown. With the aim of characterizing the role of NAPRT variants in the regulation of gene expression, we analyzed eQTLs (expression quantitative trait loci) that alter the expression of NAPRT and surrounding genes.

We obtained RNA-seq and single-tissue cis-eQTL data from GTEx Portal. We extracted the variants that alter NAPRT expression, identified surrounding genes and analyzed the eQTLs by tissue and by gene. We performed a correlation analysis to detect genes with similar expression patterns and examined their potential transcription factors from Factorbook.

The 626 SNPs identified alter the expression levels of 46 genes, including NAPRT. The 19 SNPs located in the NAPRT gene regulate mostly expression in brain, skin and adipose tissues. MROH6 gene expression is also affected by the NAPRT located variants.

TSTA3 and ZC3H3 genes have an expression pattern similar to NAPRT, and share 28 transcription factors, which are mostly related to the JAK/STAT and other signaling pathways.

The results show that eQTL analysis may reveal regulatory circuits in NAPRT expression and suggest novel functions for NAPRT.

Funding: COMPETE/FEDER UID/BIM/04501/2013 to iBiMED, PEst-OE/EEI/UI0127/2014 to IEETA and SFRH/BD/108890/2015 to SDP.

P18: Estimation of mutation rates at Y-STRs: a bi-allele approach

S. Antão-Sousa[1], A. Amorim[1,2,3], R. Gaio[1,5], L. Gusmão[4], N. Pinto[2,3,5]*

[1] FCUP – Faculdade de Ciências da Universidade do Porto, Porto, Portugal
[2] IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto, Porto, Portugal
[3] i3S – Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[4] LDD-UERJ – Laboratório de Diagnósticos por DNA – Universidade do Estado do Rio de Janeiro
[5] CMUP – Centro de Matemática da Universidade do Porto, Porto, Portugal

*Corresponding author: npinto@ipatimup.pt

Despite its usefulness in forensic casework, the mode of transmission of the Y chromosome makes it invaluable to study the mutation phenomena due to the unambiguous knowledge of which parental allele resulted in which filial allele.

Studies on human microsatellite mutation rates have focused on average estimations per marker, through the ratio between the number of Mendelian incompatibilities and the total number of allele transfers.

We studied mutation rates at Y-STRs analyzing father-son duos in a bi-allele framework (i.e. considering both the parental and filial alleles), through complete haplotypic information (not only the ones resulting in mutation).

Allele and bi-allele mutation rates were then estimated. In an inter-marker approach, we noted that: (a.) the number of repeat gains and losses is not at equilibrium, (b.) shorter (than the modal) alleles tend to mutate to longer alleles and longer alleles tend to mutate to shorter alleles, (c.) not always longer alleles showed a higher mutation rate, and (d.) alleles have very distinct mutation rates (in some cases their confidence intervals do not intersect) which, in most cases, differ from the average marker estimate.

To increase statistical confidence, it is peremptory to collect as many haplotypic data as possible. Indeed, estimates of mutations on rare alleles from direct observations are often imprecise as available statistical models for this type of rare events are yet to be improved.

P19: Exploring Japanese quail immune repertoires for antibody discovery

T. Ochôa-Pires[1], H. Froufe[2], C. Barroso[1,2], M. Rosa[1,3], J. Laranjeira[1,3], C. Egas[1,2], R.S. Vieira-Pires[1]*

[1] Center for Neuroscience and Cell Biology and Institute for Biomedical Imaging and Life Sciences (CNC.IBILI), University of Coimbra, Portugal
[2] Next Generation Sequencing Unit, Biocant, Cantanhede, Portugal
[3] HBT – Saúde e Biotecnologia Lda, Coimbra, Portugal

*Corresponding author: ricardo.pires@biocant.pt

Avian hosts can be used in the biopharmaceutical industry for generation of unique monoclonal antibodies (mAb). Indeed, their phylogenetic distance from mammals ensures generation of robust and specific antibodies and the preparation of combinatorial phage-display libraries from chicken sources is simplified over mammals.

We have been exploring the potential of Japanese quail (*Coturnix japonica*) immune repertoires for mAb development. Multiple sequence search and analysis tools were used to revisit quail genome and identify genetic regions encoding for the light (LC) and heavy chain (HC) of quail Ig, using chicken data as reference.

We were able to identify two main regions, corresponding to genes encoding quail Ig LC and HC. These supported the design of specific oligonucleotides that were subsequently used in PCR amplification of antibody fragments from a spleen cDNA library. Preliminary NextGen sequencing (NGS) analysis of the fragments confirmed the expected distribution of conserved and variable regions.

This work has enabled the design of oligonucleotide sets to generate phage-display libraries from quail immune repertoires. Moreover, we have successfully implemented NGS methodologies to analyze and characterize quail antibody repertoires in terms of size and diversity. Together, these assets open new possibilities to explore quails as source of antibodies for diagnostics and therapy.