Societat Catalana de BIOLOGIA

XVIII Jornada de Biologia Evolutiva

Organitzada per la Secció de Biologia Evolutiva de la SCB

INSTITUT D´ESTUDIS CATALANS

Carrer del Carme 47

Barcelona

25 de juny de 2018

XVIII Jornada de Biologia Evolutiva

PROGRAMA

Organitzadors:

David Juan (UPF) i Sara Guirao-Rico (IBE)

Secretaria de la SCB:
scb@iec.cat

9:00-9:20 **Registration**
9:20-9:30 **Welcome**
9:30-10:20 Sex chromosome evolution in non-model organisms. **Beatriz Vicoso** (Institute of Science and Technology-Austria).
Moderator: Marta Riutort (UB).
10:20-10:35 Screening for positive selection signals on the X chromosome in human populations. **Pablo Villegas Miron** (IBE-UPF-CSIC).
10:35-10:50 PRDM9 diversity at fine geographical scale in chromosomal races of house mice reveals contrasting evolutionary patterns in continental versus island populations. **Covadonga Vara González** (UAB).
10:50-11:05 Dynamics of de novo gene emergence in baker's yeast. **William Blevins** (UPF-IMIM).

11.05-11.30 **Coffee break**

Moderator: Cristian Cañestro (UB).
11.30-12:00 De novo gene evolution: fact or fiction?. **Mar Albà** (UPF-IMIM).
12:00-12:15 Tracing human adaptation through population genomics data. **Jesús Murga Moren** (UAB).
12:15-12:30 Transposable elements affect the transcriptional regulation of stress response gene. **Vivien Horvath** (IBE-UPF-CSIC).
12:30:12:45 Recent evolution of the epigenetic regulatory landscape in human and other primates. **Raquel García-Pérez** (IBE-UPF-CSIC).
12:45-13:00 Role of Myoglianin in metamorphosis of Blattella germanica. **Orathai Kamsoi** (IBE-UPF-CSIC).
13:00-13:15 A key role of DNA-methyltransferase 1 in early embryogenesis of the German cockroach Blattella germanica. **Alba Ventos-Alfonso** (IBE-UPF-CSIC).
13:15-13:30 Network analyses of metabarcoding data unravel molecular novelty and diversity in global marine Opisthokonta. **Alicia S. Arroyo** (IBE-UPF-CSIC).

13:30-14:45 **Lunch break**

Moderator: Elena Bosch (IBE-UPF-CSIC).
14.45-15:15 Applying New-old methods to old-new questions: using Deep Learning to studying archaic introgression in the human genome. **Oscar Lao** (CNAG).
15:15-15:30 Population genomics of marine species: a pilot study to improve laboratory protocols. **Anna Barbanti** (UB).
15:30-15:45 The colonization of a global invader studied genomewide. **Maria Casso** (CEAB-CSIC).
15:45-16:00 Decrypting the West-Eurasian ancestry of Roma people. **Erica Bianco** (IBE-UPF-CSIC).
16:00-16:15 People from Ibiza: an unexpected isolate in the Western Mediterranean. Simone **Andrea Biagini** (IBE-UPF-CSIC).
16:15-16:30 Double or nothing in the Fungal Kingdom. **Miguel Ángel Naranjo Ortiz** (CRG).
16:30-16:45 Next generation sequencing data sheds light on the phylogeography of two land planarian species in Brazil. **Ricard Sabaté Gascón de Gotor** (UB).
16:45:17:00 Host-gut microbial symbiosis in islands: the case of Podarcis lifordi lizards from the Menorcan coastal islets. **Laura Baldo** (UB).

17:00-17:20 **Break/Prevosti Prize Committee Meeting**
17:20 **Announcement of the Prevosti Prize winner. End of the meeting**

XI Premi Antoni Prevosti de Biologia Evolutiva, any 2018

Amb l'objectiu de fomentar la participació i la discussió de la recerca dels joves investigadors, pre i postdoctoral, en tots els camps de la Biologia Evolutiva, l'any 2007 fou instaurat per primera vegada el premi Antoni Prevosti de Biologia Evolutiva per premiar la millor comunicació a la jornada presentada per un jove investigador (que faci menys de 3 anys que ha llegit la tesi).

El premi consisteix aquest any en 250 € que es lliuraran al finalitzar la darrera sessió de la Jornada. El receptor del premi haurà de ser present a la sala per rebre'l. Si no és així, el premi passarà a la comunicació que hagi quedat en segon lloc, si n'hi hagués, o podria ser declarat desert.

La comissió que decidirà el premi de l'edició de l'any 2018 estarà formada per.

President: Marta Pascual (UB)
Secretari: David Comas (UPF)
Vocal1: Aurora Ruiz-Herrera (UAB)
Vocal2: Mario Cáceres (IBB)
Vocal3: Sebastián Ramos-Onsins (CRAG)

La comissió atorgarà el premi en base a la qualitat científica i presentació del treball, així com a les respostes a les intervencions dels assistents.
Els membres de la comissió no podran votar a les persones del seu grup.

SCREENING FOR POSITIVE SELECTION SIGNALS ON THE X CHROMOSOME IN HUMAN POPULATIONS

Pablo Villegas-Mirón, Jessica Nye, Jaume Bertranpetit and Hafid Laayouni

Adaptive (positive) selection in the last 100,000 years has been one of the most important evolutionary forces on the *Homo sapiens* genome. Selective sweeps are the genomic footprint of recent selection and can be detected by statistics such as the integrated haplotype score (iHS). Previously reported sweeps in the human genome have identified interesting selected targets for specific adaptations. However, few investigations into signatures of selection are on sexual chromosomes. In this study, we report a screening of recent positive selection signals in the X chromosome based on data from the third phase of the 1000 Genomes Project. We describe specific signals from each continental group (Africa, Europe, and Asia), with the strongest signature of selection in sub-Saharan Africans. Evidence suggests that in overall more than 50% of the selection footprints are non-genic, pointing to the importance of selection at a regulatory level. We find shared signals between populations in regions related to reproduction and neural development, suggesting common adaptations in dimorphic traits and spermatogenesis-related processes. Conversely, we also find signals enriched in genes that escape X chromosome inactivation, suggesting the presence of selection on expression regulation of genes; which may contribute to the dimorphic differentiation between sexes.

PRDM9 DIVERSITY AT FINE GEOGRAPHICAL SCALE IN CHROMOSOMAL RACES OF HOUSE MICE REVEALS CONTRASTING EVOLUTIONARY PATTERNS IN CONTINENTAL *VERSUS* ISLAND POPULATIONS

C. Vara[1,2], L. Capilla[1], L. Ferretti[3], R.A. Sanchéz-Guillén[1,4], S.I. Gabriel[5], G. Albert-Lizandra[1,2], B. Florit-Sabater[1,2], J. Bello-Rodríguez[1,2], J. Ventura[6] , J.B. Searle[7], M.L. Mathias[5], A. Ruiz-Herrera[1,2,*]

[1]Genome Integrity and Instability Group, Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, Barcelona, Spain. [2]Departament de Biologia Cel.lular, Fisiologia i Immunologia, Universitat Autònoma de Barcelona, Barcelona, Spain. [3]The Pirbright Institute, Woking, United Kingdom, [4] Instituto de Ecología AC (INECOL), Red de Biología Evolutiva, Xalapa, Veracruz,Mexico. [5]CESAM – Centre for Environmental and Marine Studies, Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal. [6]Departament de Biologia Animal, Biologia Vegetal i Ecologia, Universitat Autònoma de Barcelona, Campus UAB, Barcelona, Spain. [7]Department of Ecology and Evolutionary Biology, Corson Hall, Cornell University, Ithaca, NY, USA

One of the major challenges in evolutionary biology is the identification of the genetic basis of speciation. Krueppel genes comprise a large family with some members implicated in reproduction isolation among which *Prdm9* is best known. Given its pivotal role in speciation, here we consider the drivers that may account for the evolutionary dynamics of *Prdm9* between continental and island systems of chromosomal variation in house mice. Our dataset comprised nearly 400 wild-caught mice of Robertsonian (Rb) systems with diploid numbers ranging from 2n=40 to 2n=22 due to chromosomal fusions, from the vicinity of Barcelona (continental) and Madeira (island). We identified circa 60 newly described *Prdm9* alleles, revealing contrasting patterns between the Rb systems. *Prdm9* diversity was especially evident in Madeira, with 90% of the new alleles, contrary to the expectations of reduced variation for island populations. This was in sharp contrast with the situation in the Barcelona system, where Rb mice were mainly homozygous for a *Prdm9* allele characterised by 10 repeats, most probably as a result of a founder event. Moreover, the phylogenetic reconstruction including previously described *Prdm9* alleles from mice distributed across Eurasia was suggestive of multiple introgression events onto Madeira, with retention potentially favoured by the presence of Rb fusions. Given the contrasting patterns of allelic diversity observed between the Barcelona and Madeira systems, we anticipate that our study will provide the grounds for exploring the role of different genetic backgrounds for *Prdm9* in reproductive isolation in natural populations.

INVESTIGATING *DE NOVO* GENE EMERGENCE IN BAKERS YEAST

William R. Blevins[1,2], Xavier Messeguer[3], Jorge Ruiz-Orera[1] , Bernat Blasco-Moreno[4], José Luis Villanueva-Cañas[1,], Lorena Espinar[2], Juana Díez[4], Lucas Carey[2], M. Mar Albà[1,5,*]
1 Evolutionary Genomics Group, Research Programme on Biomedical Informatics, Hospital del Mar Research Institute and Universitat Pompeu Fabra, Barcelona, Spain
2 Single Cell Behavior Group, Experimental and Health Sciences Department, Universitat Pompeu Fabra, Barcelona, Spain
3 Computer Sciences Department, Universitat Politècnica de Catalunya, Barcelona, Spain
4 Virology Unit, Experimental and Health Sciences Department, Universitat Pompeu Fabra, Barcelona, Spain
5 Catalan Institution for Research and Advanced Studies, Barcelona, Spain

Recent advances in comparative genomics have revealed that some genes originate directly from intergenic regions, or *de novo*, rather than by the duplication or combination of preexisting genes. One fascinating aspect of *de novo* originated genes is that they can encode proteins with new sequences which are unlike any other proteins-some of which could gain useful functions. However, the dynamics of *de novo* gene evolution have yet to be explored in depth. One limitation to studying *de novo* gene emergence is that species or lineage-specific proteins are under-represented in current gene annotations; this results in important biases when trying to identify *de novo* genes and estimate the rate of turnover. In order to circumvent this problem, we used RNAseq to assemble *de novo* transcriptomes for eleven species of yeast from the phylum of *Ascomycota* in both normal and stress conditions. This new data allows us to compare all expressed transcripts, annotated or not, in the 11 different species. By combining comparative transcriptomics with genomic synteny, we were able to identify 213 putative *de novo* genes that are likely to have emerged over the past 20 million years. These genes, the majority of which are not annotated, represent 4.4% of all genes which were expressed in *S. cerevisiae* in our experiment. By using high-coverage ribosome profiling data, we find that a significant fraction of them are translated and are likely to encode functional proteins.

TRACING HUMAN ADAPTATION THROUGH POPULATION GENOMICS DATA

Jesús Murga-Moreno, Marta Coronado-Zamora, Sandra Bodelón, Sònia Casillas, Antonio Barbadilla
Departament de Genètica i de Microbiologia i Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Since the split with chimpanzees, and especially since the migrations that lead humans to colonize almost every single place on Earth, we have been exposed to environmental and social challenges that have shaped our genomes through the action of natural selection. The availability of a comprehensive worldwide nucleotide variation data set from the 1000 Genomes Project, contained in PopHuman database, provides the human lineage with an invaluable resource, allowing the testing of molecular population genetics hypotheses and eventually understand the evolutionary dynamics of genetic variation in human populations. Here we analyze a battery of population genetics metrics calculated along the human genome from PopHuman to: (i) describe the patterns of diversity and divergence for different chromosomes and genomic regions showing distinct rates of recombination, and (ii) perform a genome-wide scan of selection to identify regions of the genome that have been subjected to either recent sweeps or recurrent selection since the split between our species and chimpanzees. Our genome-wide scan of selection includes 3268 genomic regions showing signatures of positive selection at different timescales, of which, 1297 (~40%) overlap genes. Well-known examples of human genetic adaptation published elsewhere are included in the catalog, as well as hundreds of other interesting candidates that will require more thoroughly analyses. The catalog facilitates comparisons of each signature of selection with empirical distributions of the corresponding DNA diversity metric across the human genome and among populations, an estimation time to the common ancestor for the beneficial allele, and structural and functional annotations of the region. Our catalog will be freely available online and presented as a collaborative database to compile and annotate adaptation events along the human evolutionary history.

TRANSPOSABLE ELEMENTS AFFECT THE TRANSCRIPTIONAL REGULATION OF STRESS RESPONSE GENE

Vivien Horvath (IBE-UPF-CSIC)

A growing number of studies show the important role of TEs in stress-related adaptation in different organisms. However, the role of TEs in *Drosophila melanogaster* stress regulatory networks is poorly understood. *In silico* predictions of transcription factor binding sites (TFBS) and transcription factor binding motifs (TFBM) in the *Drosophila melanogaster* genome were performed in our laboratory, using position weight matrixes (PWMs), transcription factor flexible models (TFFMs), and ChIP-Seq data, in six different stress responses. Since not all of the predicted binding sites are expected to be functional other lines of evidence were used to select the best candidates, such as specific histone modifications observed in active promoters/enhancers, or the chromatin accessibility of the genomic regions. To confirm that the predicted TFBM/TFBS are functional, we tested a subset of them experimentally. We performed enhancer reporter assays for 11 candidate TEs, to test whether they affect gene expression under different stress conditions. We checked the enhancer function of TEs in immunity, heat-shock, oxidative stress, and in xenobiotic stress conditions by generating transgenic *D. melanogaster* flies. To create the mutants, we used a directional cloning method, and then we checked by qRT-PCR the level of expression of the reporter gene driven by the TE. The results showed that three of the tested TEs drive the expression of the reporter gene, suggesting that they add functional transcription factor binding sites in the promoter region of the genes. Although only three of the 11 TEs were found to drive the expression of a reporter gene, these number could be an underestimation, since the predicted binding sites were not obtained in stress conditions.

RECENT EVOLUTION OF THE EPIGENETIC REGULATORY LANDSCAPE IN HUMAN AND OTHER PRIMATES

Raquel Garcia-Perez[1,2], Gloria Mas-Martin[2,3], Martin Kuhlwilm[1,2], Meritxell Riera[1,2], Antoine Blancher[4], Marc Marti-Renom[2,3,5], Luciano Di Croce[2,3,5], , Jose Luis Gómez-Skarmeta[6], Tomas Marques-Bonet[1,2,5] , David Juan[1,2]
[1]Institut de Biologia Evolutiva , CSIC-UPF, Barcelona, Spain, [2]Universitat Pompeu Fabra , UPF, Barcelona, Spain, [3]Centro Nacional de Análisis Genómico - Centro de Regulación Genómica , CNAG-CRG, Barcelona, Spain, [4] Laboratoire d'Immunogenetique moleculaire , Faculte de Medecine Purpan, Universite Toulouse 3, Toulouse, France, [5]Catalan Institution of Research and Advanced Studies , ICREA, Barcelona, Spain, [6]Centro Andaluz de Biología del Desarrollo (CABD), CSIC-Universidad Pablo de Olavide-Junta de Andalucía, Sevilla, Spain

Evolutionary biologists have long sought to discern the molecular basis of phenotypic variation. Changes in gene regulation are thought to play a major role in evolution and speciation, particularly in primates. Over the last decade, the field has experienced a major shift towards inter-species comparative epigenomics in search of a conceptual step-forward in our understanding of evolution. However, the lack of coherent multi-omic datasets has hindered the integrative study of the interplay between epigenomic and genomic evolution in different species. Our study aims to characterize the evolutionary dynamics of regulatory elements in the primate lineage. To that end, we have comprehensively profiled lymphoblastoid cell lines (LCLs) from human, chimpanzee, gorilla, orangutan and macaque. We take advantage of our data to: 1) create a consistent dataset that defines the regulatory landscape of LCLs in primates 2) investigate mechanisms of gene expression regulation and understand the contribution of inter-species regulatory differences to gene expression variation 3) study the interplay between different regulatory elements and understand the chromatin changes underlying the emergence or repurposing of these elements.

ROLE OF MYOGLIANIN IN METAMORPHOSIS OF *BLATTELLA GERMANICA*

Orathai Kamsoi, Xavier Belles
Institute of Evolutionary Biology (UPF-CSIC)
Passeig Marítim 47, 08003 Barcelona, Spain, Tel. 93 230 95 00
orathai.kamsoi@ibe.upf-csic.es, xavier.belles@ibe.upf-csic.es

The onset of insect metamorphosis requires a burst of 20-hydroxyecdysone (20E) production from the prothoracic gland (PG) and a decrease of juvenile hormone (JH) production from the corpora allata (CA). In *Blattella germanica*, a 20E peak occurs at the transition from the fifth (penultimate) nymphal instar (N5) to the sixth (last) nymphal instar (N6), whereas JH rapidly declines at the beginning of N6. During this period, and in both PG and CA, we found a high expression of Myoglianin (Myo), a homolog of *Drosophila melanogaster* Myoglianin/vertebrate GDF8/11. By using RNAi approaches, we showed that the depletion of Myo induced PG overgrowth resulting from abnormally high cell proliferation. In turn, this hyper-proliferation resulted from the down-regulation of the expression of *dacapo*, an inhibitor of cyclin E/Cdk2 complex. Moreover, Myo-depleted nymphs failed to initiate metamorphosis. The insects delayed the molting in N5 and, after N6, they molted to a supernumerary nymph (N7). This is explained by the fact that Myo depletion in N5 up-regulated the expression of Juvenile hormone acid methyl transferase (a key enzyme in JH synthesis) in the CA. Our results suggest that high Myo expression in N5 promotes the initiation of metamorphosis through the large rise of 20E production in PG and the cessation of JH synthesis in CA. Thus, Myoglianin is a crucial factor in the regulation of metamorphosis in *B. germanica* and probably in other hemeimatabolan insects.

NETWORK ANALYSES OF METABARCODING DATA UNRAVEL MOLECULAR NOVELTY AND DIVERSITY IN GLOBAL MARINE OPISTHOKONTA

Alicia S. Arroyo[1], Romain Iannes[2], Eric Bapteste[2,3], Iñaki Ruiz-Trillo[1,4,5]

[1] Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra). Barcelona, Spain

[2] CNRS, UMR7138, Institut de Biologie Paris-Seine. Paris, France

[3] Sorbonne Universités, UPMC Université Paris 06, Institut de Biologie Paris-Seine. Paris, France

[4] Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona. Barcelona, Spain

[5] ICREA. Pg Lluís Companys 23, 08010. Barcelona, Spain

Opisthokonta is a clade of eukaryotes, which contains animals, fungi and several unicellular lineages. This group is essential to address evolutionary questions, such as the origins of animals or fungi. They are also ecologically relevant because of their wide range of lifestyles and ecological niches. However, the diversity and geographical distribution of some Opisthokonta lineages, specially the unicellular ones, remains poorly described.

In this paper we overcome this issue by analysing v9 18S rDNA metabarcoding data from the TaraOceans expedition with a very powerful methodological approach: similarity networks. Our first goal was to check if the environmental diversity from TaraOceans would change our current knowledge of this group. Secondly, we looked for potential novel diversity within or between some Opisthokonta groups.

We built our similarity network by blasting reference 18S sequences and environmental Operational Taxonomic Units (OTUs) from TaraOceans Opisthokonta. First, we analyzed the general structure of the diversity by computing network properties, such as closeness, betweenness, node degree and diameter of Connected Components (CCs). Our preliminary results suggest a different structure of the environmental dataset compared to the reference one, which implies that the environmental Opisthokonta expands the known diversity of this group. We also found hidden environmental diversity, not previously detected in the reference databases, by calculating the assortativity and the shortest paths of the graph.

Our research describes for the first time the marine ecology of Opisthokonta from a global perspective. Moreover, it proves that there is still a lot of unknown diversity. Finally, it provides a framework for future metabarcoding studies to assess molecular novelty and diversity using graphical networks, especially when the phylogenetic signal of the marker is low.

POPULATION GENOMICS OF MARINE SPECIES: A PILOT STUDY TO IMPROVE LABORATORY PROTOCOLS.

Anna Barbanti, Hector Torrado, Luca Bargelloni, Rafaella Franch, Luis Cardona, Jesús Tomás, Abdulmaula A. Hamza, Enrique, MacPherson, Carlos Carreras and Marta Pascual
Department de Genètica, Microbiologia i Estadística and IRBio, Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Spain. Tel: 934024850. Email: anna.barbanti.es@gmail.com.

Genomic 2b-RAD libraries represent a powerful tool to identify SNPs across the whole genome and to perform population genomic studies even for species without reference genome. However, its effectiveness is highly dependent on the selection of the restriction enzyme and depth sequencing and thus adequate pilot studies are crucial to optimize the laboratory protocols. In this study, we tested the efficiency of 2b-RAD on two different non-model species with different genome sizes, *Caretta caretta* (2.4 Gb) and *Diplodus puntazzo* (0.9 Gb). We designed a pilot study including 15 samples from Libya and 9 samples from Valencia for *C.caretta* and 12 samples from Blanes and 12 samples from Javea for *D.puntazzo*. All samples were analysed following a 2b-RAD protocol using both AlfI and CspCI restriction enzymes. Furthermore, we used bioinformatics tools to simulate a selective base selection to predict the number of loci that would be obtained and their impact on genetic differentiation. The number of loci found after filtering was very variable (71,000 – 16,886), being higher in *D.puntazzo* than *C.caretta,* and higher in AlfI than CspCI. *C.caretta* not only showed less loci, but also lower mean depth due to its larger genome. The percentage of retained loci by applying the base selection simulation was very variable, indicating that bioinformatic simulations are beneficial prior to laboratory base selections. Genetic distances among individuals calculated using all data were highly correlated with those computed with AT and GC base selection simulations. Overall, the 2b-RAD technique produced good results in both species and provided data to adjust the protocol to obtain the desired number of loci but maintaining enough mean depth not to compromise loci quality. This is a case study to use pilot studies to improve laboratory protocols for genomic projects on non-model organisms with different genome sizes.

THE COLONIZATION OF A GLOBAL INVADER STUDIED GENOMEWIDE

Maria Casso1,2, Marta Pascual2, Xavier Turon1

1 Center for Advanced Studies of Blanes (CEAB-CSIC), Blanes, Spain
2 Department of Genetics, Microbiology and Statistics, and IRBio, University of Barcelona, Barcelona, Spain.

Biological invasions are a great threat to biodiversity around the world. In the marine environment, ascidians are a major group of invader species. *Didemnum vexillum* is a colonial ascidian native from Japan which has colonized most of the world temperate regions: New Zealand, East and West of USA, and Atlantic and Mediterranean coasts of Europe. This species usually lives on hard substrates, mainly on artificial substrates such as boat hulls and marinas, but also fouling bivalves in aquaculture facilities causing a strong ecological and economic impact.

Genetic studies about *D. vexillum* using COI (Cytochrome Oxidase I) gene revealed the existence of at least 2 differentiated mitochondrial clades of which only one seemed to be invasive. In order to better know the population structure and the mechanisms that led this species to colonize all temperate regions we used Genotyping-by-Sequencing (GBS) technique in a total of 293 individuals from 12 populations distributed worldwide including native and invaded areas.

As many colonial ascidians can form chimeric colonies, we decided to perform DNA extractions from a single zooid. To get enough DNA for GBS from a single zooid (ca. 1 mm) we used the Whole Genome Amplification (WGA). To test the reliability of WGA we cut in half 8 zooids and sequenced them separately. The results show a percentage of shared alleles (PSA) higher than 98% between halves from the same individual and a PSA between 55 and 65% when comparing different individuals.

The ca. 2,000,000 reads per individual obtained from GBS were filtered and analysed with GIbPSs. Clones (PSA > 98%) and non-invasive clade individuals (PSA < 50%) were deleted and a final dataset of 577 loci was obtained and used to carry out Multidimensional Scaling and STRUCTURE analyses. The preliminary results show 3 independent colonizations: New Zealand with East USA, West USA and Europe.

Key words: invasive species, *Didemnum vexillum*, WGA, GBS, GIbPSs.

DECRYPTING THE WEST-EURASIAN ANCESTRY OF ROMA PEOPLE

Erica Bianco[1], Carla Garcia-Fernandez[1], Begoña Dobon-Berenguer[1], Mihai G. Netea[2], Jaume Bertranpetit[1], David Comas[1]

[1]Departament de Ciències Experimentals i de la Salut, Institut de Biologia Evolutiva (CSIC-UPF), Universitat Pompeu Fabra, Barcelona, Spain; [2]Department of Internal Medicine, Radboud University Medical Center, Nijmegen, the Netherlands.

Roma people are the largest minority in Europe. They are among the last populations to settle in Europe, arriving ~1kya. We analyzed the whole genome sequence of 45 Roma people, belonged to 4 migrant groups (Balkan, Romungro, Vlax and North and Western), to understand recent admixture events that shaped Roma people genomes. Classical clustering methods (PCA and admixture) confirmed Roma people have two main genetic components: West-Eurasian and South-Asian. The West-Eurasian component had two sources: an ancient source (the Ancestral North Indian component of proto-Roma) and a recent source (from the recent admixture with Europeans). Using the admixture graph approach, we found ~60% of Roma West-Eurasian component came from the recent European admixture (recent source). To test for differences in the admixture pattern between migrant groups, we used the total length of IBD fragments between Roma and Europeans. We found North and Western had a larger part of the genome IBD with Europeans (366±83Mb in IBD), but it was only significantly higher than Vlax(211±80Mb). Finally, we used the f3-outgroup approach to identify the European population that shared the more drift with Roma, likely the closest to the source of admixture. Irrespective of the migrant group, all the tested European populations shared similar amount of drift with Roma, and there were no differences between migrant groups. Our results show that recent admixture between Roma and Europeans increased Roma West-Eurasian component of ~60%. The increase of West-Eurasian component was different between migrant groups, being significant between Vlax and North and Western Roma.

PEOPLE FROM IBIZA: AN UNEXPECTED ISOLATE IN THE WESTERN MEDITERRANEAN

Simone Andrea Biagini, Neus Solé-Morata, Elizabeth Matisoo-Smith, Pierre Zalloua, David Comas, Francesc Calafell

According to history, Ibiza's ancestry finds its roots in the Middle East, North Africa and Europe: before the Catalans conquered the island in 1235, Ibiza already had experienced many different cultures. It was the Phoenician Iboshim, the Carthaginian Ibosium, the Islamic Yabisah, up to the Catalan Eivissa. How all these different civilizations affected the modern genetic structure of the islanders is still unexplored. In this genome-wide study, we dug into the genetic structure of a group made up of individuals coming from different autonomous communities of Spain. Our results pointed to a clear split in two major groups, clearly separating Ibiza from the rest of the samples. We aimed to find the historical reasons behind this result: is Ibiza separating because of recent historical events, or because of some more well-established historical reasons? We explored the possibility that the modern samples from Ibiza had something to share with the ancient culture from Phoenicia using a sample retrieved in a Phoenician necropolis on the island of Ibiza. Mostly, our analyses pointed out different aspects that seem to link the genetics of the modern samples with the history of the area they lived in, more than to any ancient genetic echo from the past. According to history, Ibiza experienced a series of dramatic demographic changes due to several moments of famine, wars, up to malaria and plague. Interestingly, the ROH analysis, together with the analysis of the changes in effective population size (Ne) through time, support the idea that a founder effect, followed by a strong recent bottleneck, together with the perpetuated practice of endogamous unions, have shaped the genetic structure of current Ibizans.

DOUBLE OR NOTHING IN THE FUNGAL KINGDOM

Miguel Ángel Naranjo-Ortiz. Comparative genomics group, CRG (Barcelona)

Whole genome duplication (WGD) is a drastic genomic accident that resides at the heart of many macroevolutionary transitions in both plants and animals. However, fungal evolutionary history remains suspiciously calmed, with just a few described examples. Two approaches have been used in this work to try to unearth these events. We hypothesized that the absence of WGD is apparent, caused by a insufficient lineage representation, incomplete phylogenomic exploration and methodological artifacts related to the very nature of polyploid genomes. Here, we use a phylogenomic approach in a dataset within the 1000 Fungal Genomes Project to try to pinpoint undescribed ancient WGD events. At the same time, we have tried to identify undetected allopolyploids by applying *de novo* genome assembly and an specialized pipeline on public SRA libraries of highly fragmented genome assemblies, under the assemption that, at least in some cases, poor assembly quality is caused by unsuspected genomic traits.

NEXT GENERATION SEQUENCING DATA SHEDS LIGHT ON THE PHYLOGEOGRAPHY OF TWO LAND PLANARIAN SPECIES IN BRAZIL

Ricard S. Gascón-de-Gotor, Marta Riutort and Marta Álvarez-Presas

*Department of Genetics, Microbiology and Statistics, Universitat de Barcelona

The Brazilian Atlantic Forest (BAF) is one of the most diverse and endangered biomes in our planet. Knowing the evolutionary history of such biodiversity is critical when planning conservation policies. Low dispersal animals, such as land planarians (Platyhelminthes), have been proposed as an accurate indicator of the history and conservation state of the region. Previous studies with flatworms in the BAF used few genetic markers with low resolution, pointing to the need of new molecular data. Here, we have increased the number of molecular markers and avoided the lack of resolution of previous studies. We have used Next Generation Sequencing (NGS) in two different approaches, the first one to obtain new variable nuclear markers (DOM4, DOM5 and DOM6) with the use of software designed for this purpose, and the second, obtaining ddRADseq data using Illumina sequencing. We test the validity of these two approaches and compare the results obtained in a phylogeographic study of th e two land planarian species Cephaloflexa bergi and Imbira marcusi, used in previous studies, in a region around the city of São Paulo. In that scenario both natural (climatic, geographic) and human derived (roads, cities, agriculture) factors affect present distribution and diversity of populations.

Results using few markers vs the ones obtained with a big number of loci are congruent despite the different conditions and bioinformatic analytical requirements. Our phylogeographic analyses show the existence of genetic structure in the region and are consistent with the previous ones, which suggest an ancient origin for the diversity in the forest that may have been shaped by events older than last glaciations. Our results also point to the existence of putative glacial microrefugia in the lowland areas of the sampling region and to several possible secondary contacts occurring in the central regions of the distribution.

HOST-GUT MICROBIAL SYMBIOSIS IN ISLANDS: THE CASE OF *PODARCIS LIFORDI* LIZARDS FROM THE MENORCAN COASTAL ISLETS

Baldo L; Riera JL; Mitsi K; Pretus JL

Departamento de Biología Evolutiva, Ecologia y Ciencias Ambientales. Universidad de Barcelona, Barcelona, España

Islands are wonderful systems to explore eco-evolutionary processes occurring at rapid time scales, with innumerous studies looking at trait variation (largely morphological) under both strong neutral and selective forces following population isolation. The gut microbiota represents a still unexplored host trait in the contest of island evolution.
Here we present the trajectories of the gut microbiota diversification in seven allopatric populations of the lizard *Podarcis lilfordi* from the Menorcan coastal islets. These populations descend from an ancestral mainland population, now extinct, and became isolated following the sea level rise during the past 10 000 years (Holocene) which led to coastal fragmentation and progressive isolation of islet populations up to date.
Processes shaping the conservation and diversification of the ancestral gut microbial community during the repetitive host population splits and bottlenecks were investigated through Illumina amplicon metagenomics. Findings are discussed in light of the host phylogeography and ecology.