Societat Catalana de BIOLOGIA

XVI Jornada de Biologia Evolutiva

Organitzada per la Secció de Biologia Evolutiva de la SCB

INSTITUT D´ESTUDIS CATALANS

Carrer del Carme 47

Barcelona

27 de juny de 2016

XVI Jornada de Biologia Evolutiva

PROGRAMA

Organitzadora:

Marta Pascual (UB)

Secretaria de la SCB:
scb@iec.cat

9 -9:20 Registration

9:20- 9:30 Welcome

9:30- 10:20 Invited speaker talk. Studying connectivity and the structure of marine populations with Lagrangian Flow Networks. **Vincent Rossi** (CSIC-UIB)

Moderator: Tomas Marques

10:20-10:35 Detecting positive selection in admixed populations from Ethiopia. **Sandra Walsh** (UPF-CSIC)

10:35-10:50 Gene complexity is the proximate explanation for the molecular hourglass model in *Drosophila melanogaster.* **Marta Coronado** (UAB)

10:50-11:05 Purifying selection is acting on human long non-coding RNAs. **Cinta Pegueroles** (CRG-UPF)

11:05-11:30 Coffee break

Moderator: Barbara Negre

11:30-11:45 North African Populations: Insights from whole Y chromosome sequences. **Neus Solé-Morata** (CSIC-UPF)

11:45-12:00 Impact of misidentified samples on population genomics of non-model organisms. **Marc Vega** (UB)

12:00-12:15 . *Capsaspora owczarzaki* as a unicellular model to study co-option of the ancestral integrin adhesome. **Helena Parra** (UPF-CSIC)

12:15-12:30 Admixture, Genetic Diversity and Relatedness in Captive Chimpanzee Populations. **Clàudia Fontserè-Alemany** (UPF-CSIC)

12:30-12:45 Differential expression of transcription factor E93 is involved in neoteny of Strepsipteran *Xenos vesparum* and *Xenos peckii*. **Silvia Chafino** (UPF-CSIC)

12:45-13:00 Integrative taxonomy gives light to an unexpected diverse and structured world underground: the case of land planarians in two Spanish national parks. **Marta Álvarez-Presas** (UB)

13:00-13:15 Native American pre-colonization genetic history through admixed Brazilians. **Alex Mas-Sandoval**. (UPF-CSIC)

13:15-14:45 Lunch break

Moderator: Marta Riutort

14:45-15:00 Demographic history of the genus Pan inferred from whole mitochondrial genome reconstructions. **Irene Lobon** (UPF-CSIC)

15:00-15:15 The evolutionary role of polymorphic inversions in the human genome. **Carla Giner-Delgado** (UAB)

15:15-15:30 The adaptive transposable element *Bari-Jheh* in *Drosophila melanogaster*. **Lain Guio** (UPF-CSIC)

15:30-15:45 Evaluating the reintroduction program of green turtle (*Chelonia mydas*) in Cayman Islands. **Clara Martín** (UB)

15:45-16:00 Chimpanzee genomic diversity reveals ancient admixture with bonobos. **Marc De Manuel Montero** (UPF).

16:00-16:15 New genes and functional innovation in mammals. **José Luis Villanueva-Cañas** (GRIB-IMIM-UPF)

16:15- 16:45 Break/Prevosti Prize Committee Meeting

16:45 Announcement of the Prevosti Prize winner. End of the meeting

IX Premi Antoni Prevosti de Biologia Evolutiva, any 2016

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 2016 estarà formada per.

President: Julio Rozas (UB)
Secretari: Cèsar Blanché (UB)
Vocal 1: Mª Pilar García Guerreiro (UAB)
Vocal 2: David Comas (IBE-UPF)
Vocal 3: Mª Inés Roldán (UdG)

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.

Studying connectivity and the structure of marine populations with Lagrangian Flow Networks

Vincent Rossi

Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus Universitat de les Illes Balears

Assessing the effect of larval dispersal on the structuring and dynamics of marine populations is still problematic, despite the fact that connectivity has crucial implications for the design of reserves and the evaluation of exploited stocks. The Lagrangian Flow Network (LFN) methodology provides a systematic characterization of multi-scale dispersal and connectivity for the dispersive stages of marine organisms, offering new opportunities in population genetics, fisheries and conservation biology. This modelling approach consists in subdividing the basin into sub-regions which are interconnected through the transport of larvae by ocean currents. Post-processing of connectivity matrices permits the identification of hydrodynamical provinces, the computation of various connectivity proxies measuring retention and exchange of larvae and the estimation of multi-generational connectivity. Those diagnostics are used to better understand how dispersal processes control the genetic structure of marine populations. They should also help geneticists to formulate improved hypotheses of gene flow patterns and to design their sampling strategy accordingly. Another project using the LFN framework explores the sensitivity of management units' boundaries to connectivity processes. Focusing on Hake, we study the links between the mainland spawning areas and the recruited stock surrounding the Balearic Islands. Numerical experiments and long-term observations are combined to test if connectivity can explain inter-annual variations of recruitment. Last, this methodology shows great promise as future developments (e.g. implementing habitat patchiness and modelling the abiotic controls of larval production and survival) are handily accessible.

Integrative taxonomy gives light to an unexpected diverse and structured world underground: the case of land planarians in two Spanish national parks

Marta Álvarez-Presas[1,*], Eduardo Mateos[2], Ronald Sluys[3] and Marta Riutort[1]

[1] Dept de Genètica, Facultat de Biologia and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain.
[2] Dept de Biologia Animal, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain.
[3] Naturalis Biodiversity Center, Leiden, The Netherlands.
[*] onaalvarez@ub.edu

There is a belief that the number of species awaiting description on the planet is much higher in the tropics than in temperate zones. However, this does not mean that most species of the temperate regions such as Europe have already been described and that there is not more work to do. A clear example are land planarians, a little known group of interesting creatures that inhabit forest soils and can provide useful information on the effects of past climate change as reflected in their current diversity distribution. For this reason, we have conducted several sampling campaigns in a comparative study focused on two national parks in northern Spain: Picos de Europa (Asturias, León and Cantabria) and Ordesa y Monte Perdido (Huesca) in order to analyse distribution patterns of genetic diversity of the most common European species, *Microplana terrestris*. Contrary to our expectations, we found a large number of new cryptic species under the external appearance of *M. terrestris*. In the comparison between parks, Ordesa National Park has less diversity of species, only 4 compared to the 13 that are in the National Park Picos de Europa. Only one of these species is new to science in Ordesa, while in the Picos de Europa have been detected 7 new species. The integration of molecular data with anatomical information allows us to learn more about these species and we have also been able to determine that the level of genetic diversity among them is quite high. Another difference between the two parks is the fact that the species found in Ordesa are also found in other remote parts of Europe, while most species in Picos de Europa present a restricted distribution to the park or nearby regions in the north of the Iberian Peninsula. Probably Picos de Europa acted as a refuge of forests and their associated fauna during the Pleistocene, and then remained isolated or expanded only to nearby regions. By contrast Ordesa may have been re-colonized by species of Eastern Europe or acted as one of the refugia and points of origin of recolonization to northern Europe. The present study shows that diversity of the cryptic soil fauna is higher than imagined beforehand and, although land planarians can be difficult to find in temperate regions, the study of their diversity and biogeography can lend fruitful knowledge on the past history of the forests they inhabit.

Differential expression of transcription factor E93 is involved in neoteny of Strepsipteran *Xenos vesparum* and *Xenos peckii*.

Silvia Chafino[1,*], David Lopez[1], Jeyaraney Kathirithamby[2], Elena Casacuberta[1], Xavier Franch-Marro[1], and David Martin[1]

[1] Institute of Evolutionary Biology (IBE) CSIC- Pompeu Fabra Universitat, Barcelona Biomedical Research Park (PRBB), Passeig Marítim de la Barceloneta, 37-49, 08003, Barcelona (Spain)
[2] Department of Zoology, University of Oxford, Oxford, United Kingdom.
[*] silvia.chafino@ibe.upf-csic.es

Endoparasitic females from derived families of the order Strepsiptera are neotenic, and reach sexual maturity within their hosts, the wasps. This developmental strategy is characterized for the lack of complete metamorphosis and the retention of juvenile morphology in the adult form. In fact, only the cephalothorax of Strepsiptera females presents adult features, such as the genitalia, whereas the abdomen remains larviform. In holometabolous insects, metamorphosis is triggered by the steroid hormone 20-hydroxyecdysone (20E) and the specific up-regulation of the adult-specifier transcription factor E93 in the last juvenile stage as well as during the pupal stage. In this study, we investigated the role of *E93* in neoteny in Strepsipteran species. Using *Xenos peckii* and *Xenos vesparum* as Strepsipteran species models, we cloned and characterized the expression of *E93* in different developmental stages, and in two differentiated body sections of neotenic females, the cephalothorax and the abdomen. In addition, we checked the expression of two 20E-responsive genes, *HR3* and *EcRA*, in adult and juvenile parts of Strepsiptera females. Our results show that the expression of *E93* is higher in the cephalothorax than in the abdomen of neotenic females. This *E93* differential expression can be caused by differential sensibility to the action of 20E in the abdomen and cephalothorax regions. Overall, our data suggest that *E93* is involved in female neoteny in Strepsipteran species.

Gene complexity is the proximate explanation for the molecular hourglass model in *Drosophila melanogaster*

Marta Coronado[1,*], Irepan Salvador[2], David Castellano[3], Antonio Barbadilla[1] and Isaac Salazar[1,2]

[1] Universitat Autònoma de Barcelona, Institut de Biotecnologia i Biomedicina/Departament de Genètica i Microbiologia, 08193 Cerdanyola del Vallès, Spain.
[2] University of Helsinki, Center of Excellence in Experimental and Computational Developmental Biology. Institute of Biotechnology, FI-00531 Helsinki, Finland.
[3] Bioinformatics Research Centre, Aarhus University, 8000 Aarhus C, Denmark.
[*] marta.coronado@uab.cat

The hourglass model of embryonic evolution is revisited at the molecular level using a recent and powerful approach, the DFE-alpha[1], which allows the estimation of both conservation and adaptation at the level of protein-coding genes. Genome data on polymorphism within a species (*Drosophila melanogaster*, from DGRP[2]) and divergence between two species (*D. melanogaster* and *Drosophila yakuba*) are used to estimate the fraction of non-synonymous fixations occurred from the separation of both species that are actually adaptive (α). We calculate α for genes expressed through the complete life cycle using data from RNA-seq experiments (modENCODE database[3]). In addition, we estimate for each stage genomic determinants such as codon usage bias, intron length, expression bias or number of exons.

We find that pupa and adult male stages exhibit the highest levels of adaptive change while mid and late embryonic stages show higher conservation. The earliest stages of embryonic development are the most divergent but, in contrast to the pupa and adult stages, this lack of conservation is partially explained by relaxation of natural selection. The highest conserved states, mid and late embryonic development, show the most complex gene structure: larger genes, more exons, more transcripts and longer introns. We conclude that: (i) adaptation along the D. melanogaster life-cycle occurs in two discrete and well differentiated periods and; (ii) gene structure complexity associated to each developmental stage seems the likely proximate explanation for the observed hourglass pattern.

References

[1] Eyre-Walker A, Keightley PD. Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. Mol Biol Evol. 2009;26: 2097–108.

[2] Mackay TFC, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D, et al. The Drosophila melanogaster Genetic Reference Panel. Nature. 2012;482: 173–178.

[3] Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, et al. The developmental transcriptome of Drosophila melanogaster. 2011;471: 473–479.

Chimpanzee genomic diversity reveals ancient admixture with bonobos

Marc de Manuel[1], Martin Kuhlwilm[1], Peter Frandsen[2], Chimpanzee Diversity Consortium, Christina Hvilsom[2], Tomas Marques-Bonet[1,4,5]

[1] Institut de Biologia Evolutiva, (CSIC-Universitat Pompeu Fabra), PRBB, Doctor Aiguader 88, Barcelona, Catalonia 08003, Spain.
[2] Department of Biology, Bioinformatics, University of Copenhagen, 2200 Copenhagen, Denmark.
[3] Research and Conservation, Copenhagen Zoo, 2000 Frederiksberg, Denmark.
[4]CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain
[5] Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia 08010, Spain.

Due to an almost complete absence of fossils, the evolutionary history of chimpanzees can best be explored using genetic data from present-day populations. Here, we analyzed the complete genomes of 75 wild-born chimpanzees and bonobos from ten different countries in Africa to decipher the complex demographic history of our closest living relatives. We find that the central chimpanzees carry the largest amount of ancestral variation and that population structure at regional scale makes genetic diversity a good predictor of the geographic origin, making it possible to re-assign chimpanzees of unknown provenance. Multiple lines of evidence from allele sharing, properties and ages of haplotypes, as well as demographic models based on the allele frequency spectrum suggest ancient gene flow from bonobos into the ancestors of central and eastern chimpanzees more than 200 thousand and less than 500 thousand years ago, probably with subsequent spread to Nigeria-Cameroon chimpanzees. Additionally, more recent gene flow from bonobos into central chimpanzees implies at least two phases of secondary contact, contributing at least ~1% to the central chimpanzee genome. Admixture thus appears to have been widespread during hominid evolution.

Admixture, Genetic Diversity and Relatedness in Captive Chimpanzee Populations

Clàudia Fontserè-Alemany[1], Jéssica Hernández-Rodríguez[1], Marc de Manuel Montero[1], Peter Frandsen[2,3], Christina Hvilsom[2,3], David Hughes[1], Tomàs Marquès-Bonet[1,4].

[1] *Institut de Biologia Evolutiva, (CSIC-Universitat Pompeu Fabra), PRBB, Doctor Aiguader 88, Barcelona, Catalonia 08003, Spain*
[2] *Department of Biology, University of Copenhagen, DK-2200 Copenhagen N, Denmark*
[3] *Copenhagen Zoo, DK 2000 Frederiksberg, Denmark*
[4] *Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia 08010, Spain*

Conservation programs for endangered species, such as Chimpanzees (*Pan troglodytes*), are of major importance, particularly when hunting, deforestation and disease are provoking a continuous decline of such wild populations. Here, we genotyped 95 European captive chimpanzees and 13 individuals from African Sanctuaries with the aim of using genetic diversity and chimpanzee population structure to help *ex-situ* conservation efforts. To do so we performed targeted capture re-sequencing for genome-wide distributed ancestry informative markers (AIMs) and estimated genetic diversity, ancestry and relatedness.

We were able to assign a genetically inferred subspecies or admixed status to each individual and compare this information with the recorded data in the Studbook of Chimpanzee. We reported that 40% of the zoo captive chimpanzees exhibit some degree of admixed ancestry. We also conducted relatedness characterization and we observed high levels of kinship among captive chimpanzees. Our results highlight the importance of using a genetic approach in management programs to avoid outcomes counter to conservation efforts such as inbreeding or hybridization events. This is particularly important if there are plans to reintroduce individuals into the wild.

All this new available information will help improve and complete the Studbook of Chimpanzees to design precise breeding and transfer programs.

The evolutionary role of polymorphic inversions in the human genome

Carla Giner-Delgado[1,2], David Castellano[1], Sergi Villatoro[1], Isaac Noguera[1], David Izquierdo[1], Paul F O'Reilly[3], Mario Cáceres[1,4]

[1] Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain
[2] Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, Bellaterra (Barcelona), Spain
[3] MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[4] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Inversions are balanced structural variants that inhibit recombination between inverted and non-inverted alleles. This special feature led to the hypothesis that they may play a key role in processes such as local adaptation, evolution of sex chromosomes and speciation. However, despite their potential effect as evolutionary drivers, their role in the human genome is still poorly understood. In this work, we used the latest population-scale data from the INVFEST project to assess inversion importance in recent human evolution. The data set consists of 45 high-frequency inversions (average global minor allele frequency ~0.3), with sizes ranging from ~100 bp to 0.4 Mb, which have been experimentally validated and genotyped in 550 individuals from seven populations of the 1000 Genomes Project. By analysing their frequencies in different populations and the nucleotide divergence between inverted and non-inverted haplotypes, we studied the general patterns of human inversions and also selected candidates likely to have a greater functional or evolutionary impact. First, we compared inversion frequencies and geographical distribution to those of SNPs under the same discovery and genotyping process. This allows us to detect differences in the distribution of selection coefficients or mutation rate. Next, we tested for signals of positive or balancing selection in each variant by modeling population differentiation and estimating their age from nucleotide variation data. We find inversions with an estimated recent origin that have rapidly increased in frequency in a continent or population, indicative of a beneficial effect. We also find old inversions that have been maintained at high frequency in all populations, showing low population differentiation, which may be under balancing selection or recurrent mutation equilibrium. Finally, we provide several examples to illustrate the potential of inversions as genomic modifiers and report cases that deserve further molecular and phenotypic characterization.

The adaptive transposable element *Bari-Jheh* in *Drosophila melanogaster*

Lain Guio, Josefa González

Institute of Evolutionary Biology (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta, 37-49, 08003 Barcelona, Spain. Tel: +34-93 230 9507.

Understanding how organisms survive to their environments is one of the current major challenges in evolutionary biology. Elucidating the interplay between genotype, phenotype and environment is necessary to make accurate predictions of important biological outcomes such as stress resistance or yield in economically important plants and animals, and disease in humans. Choosing the appropriate tools is key to shed light in the process of adaptation. In this case, we chose *Drosophila melanogaster* as a model organism, and transposable elements as a source of adaptive mutation. Based on population frequency patterns and footprints of selection at the DNA level, we identified the transposable element *Bari-Jheh* as a candidate insertion. *Bari-Jheh* is inserted in the intergenic region of *Juvenile Hormone Epoxy Hydrolase* (*Jheh*) genes. However, the adaptive effect of this mutation remained elusive. Using transcription factor binding site prediction software, and modENCODE available data, we showed that *Bari-Jheh* adds extra antioxidant response elements upstream of *Jheh1* and *Jheh2* genes. Accordingly, we find that *Bari-Jheh* is associated with upregulation of *Jheh1* and *Jheh2* and with resistance to oxidative stress induced by two different compounds, both of them relevant for natural *Drosophila melanogaster* populations. Finally, we found that *Bari-Jheh* is also associated with downregulation of *Jheh3*. To shed light in the molecular mechanisms underlying this downregulation, we studied how *Bari-Jheh* affects the chromatin structure of the region. We showed that *Bari-Jheh* changed the histone mark enrichment in stress and nonstress conditions by adding H3K9me3 and or H3K27me3 histone marks. Overall our results highlight the role of transposable elements in environmental adaptation, describing two different molecular mechanisms that contribute to the understanding of resistance to oxidative stress in natural populations of *Drosophila melanogaster*.

Demographic history of the genus Pan inferred from whole mitochondrial genome reconstructions

Irene Lobon[1], Serena Tucci[2], Marc de Manuel[1], Silvia Ghirotto[2], Andrea Benazzo[2], Guido Barbujani[2], Christina Hvilsom[3], Tomas Marques-Bonet[1,4,5]

[1]Institut de Biologia Evolutiva (CSIC-UPF), Universitat Pompeu Fabra, Barcelona, Spain
[2]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy
[3]Research and Conservation, Copenhagen Zoo, Frederiksberg, Denmark
[4]Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain
[5]CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

The genus Pan is the closest genus to our own and it includes two species, Pan paniscus (bonobos) and Pan troglodytes (chimpanzees). The later is constituted by four subspecies, all highly endangered. The study of the Pan genera has been incessantly complicated by the intricate relationship among subspecies and the statistical limitations imposed by the reduced number of samples or genomic markers analysed.

Here, we present a new method to reconstruct complete mitochondrial genomes (mitogenomes) from whole genome shotgun (WGS) datasets, mtArchitect, showing that its reconstructions are highly accurate and consistent with long range PCR mitogenomes. We used this approach to build the mitochondrial genomes of 20 newly sequenced samples which, together with available genomes, allowed us to analyse the hitherto most complete Pan mitochondrial genome dataset including 156 chimpanzee and 44 bonobo individuals, with a proportional contribution from all chimpanzee subspecies. We estimated the separation time between chimpanzees and bonobos around 1.15 Mya [0.81-1.49]. Further, we found that under the most probable genealogical model the two clades of chimpanzees, Western+Nigeria-Cameroon and Central+Eastern, separated at 0.59 Mya [0.41-0.78] with further internal separations at 0.32 Mya [0.22-0.43] and 0.16 Mya [0.17-0.34], respectively. Finally, for a subset of our samples, we compared nuclear vs. mitochondrial genomes and we found that chimpanzee subspecies have different patterns of nuclear and mitochondrial diversity, which could be a result of either processes affecting the mitochondrial genome, such as hitchhiking or background selection, or a result of population dynamics.

Evaluating the reintroduction program of green turtle (*Chelonia mydas*) in Cayman Islands

Clara Martín*, Marta Pascual, Carlos Carreras

Department of Genetics, Microbiology and Statistics and IRBio, University of Barcelona, Av.Diagonal 643, E-08028 Barcelona, Spain
*claramarpez@hotmail.com

Reintroduction programs have been widely used as a tool for conservation of endangered species against population declines. However, the effectiveness of these programs and their impact on wild populations has not been always assessed. The endangered green turtle (*Chelonia mydas*) has been declining due to many different threats, including the poaching of eggs, the use of adult turtles for meat consumption and fisheries interaction. A reintroduction program began in 1970 in the Cayman Islands with the aim to repopulate its natural population, which was considered extinct, using the captive stock of the Cayman Turtle farm started with individuals from Mexico, Nicaragua, Costa Rica, Guyana, Suriname and Ascension Island. To evaluate the effect of the farm reintroduction in the wild green turtle population we analyzed breeding adult females from the farm (n= 70) and from the wild population (n= 58) by sequencing an 800 bp fragment of the control region of the mitochondrial DNA. The results showed no significant differences between the two sites (F_{ST}= -0.0035, p- value>0.05). Furthermore, they both presented haplotypes of the two major lineages of the Atlantic, in accordance to the mixed origin of the founding farm stock. In addition, we trimmed the sequence to 380 bps as used in previous studies of other locations in the Atlantic. The comparison with all other available populations resulted in generally high levels of differentiation among the Atlantic populations and both farm and wild sample sets clustered with the Caribbean populations. The whole 800 bp sequence data was available for some of these populations but no improvement of the resolution of population differentiation was observed in comparison to the 380 bp fragment, despite having two new polymorphic sites. Our results agree with the known origin of the founder individuals of the farm and are consistent with a scenario of a successful reintroduction, although alternative hypothesis could not be fully discarded. In summary our results suggest that the reintroduction program resulted in a high diverse wild population but additional analyses, using other genetic markers such as microsatellites or mtDNA STRs, are needed to confirm the success of the reintroduction project.

Native American pre-colonization genetic history through admixed Brazilians

Alex Mas-Sandoval[1,2,*], Lara R. Arauna[2], Eduardo Tarazona-Santos[3], Maria-Cátira Bortolini[1], David Comas[2]

[1] Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil.
[2] Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciències de la Salut i de la Vida, Universitat Pompeu Fabra, Barcelona, Spain.
[3] Departamento de Biologia Geral, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.
[*] alex.mas.sandoval@upf.edu

Brazilian urban admixed populations are build of three main genetic components: Native American, European and African. Although the proportions between and within populations are different, the Native American component is the one found in lowest proportions (6-8%) and was mostly admixed in east Brazil soon after the European Conquest. From 6487 urban admixed Brazilians genotyped at 2,5M SNPs (Kehdy et al 2015) in three locations of eastern Brazil (North-East, South-East and South regions) we extract the Native American haplotypes to rearrange them in *Assembled Individuals* with full Native American ancestry, recreating a historical scenario where Native Americans had not admixed with Europeans and Africans. This method allows us to analyze the Native Americans population structure in a region where most of Native American populations disappeared after European colonization. The structure of the Native American component in admixed individuals has ben described in a fine scale in the Caribean, Mesoamerican, and Andean regions (Moreno-Estrada et al 2013, Moreno-Estrada et al 2014, Homburger et al 2015) but few approaches have analyzed it in South American east coast. We observe that population structure of Native American Assembled Individuals is modulated by the geography of the sampling location while the populations structure of urban admixed Brazilians is driven by the European or the African ancestry. In order to explore in depth the demographic history and the origins of the Brazilian Native American component we analyze the Native American Assembled Individuals together with a dataset of non-admixed Native Americans of Brazil to see differential ancestry of Native American main genetic groups in Brazilian urban admixed populations.

References

1. Kehdy, F. S. G. *et al*. Origin and dynamics of admixture in Brazilians and its effect on the pattern of deleterious mutations. *Proc. Natl. Acad. Sci.* **112,** 8696–8701 (2015).
2. Moreno-Estrada, A. *et al*. Reconstructing the Population Genetic History of the Caribbean. *PLoS Genet*. **9,** (2013).
3. Moreno-Estrada, A. *et al*. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science (80-.).* **344,** 1280–1285 (2014).
4. Homburger, J. R. *et al*. Genomic Insights into the Ancestry and Demographic History of South America. *PLoS Genet*. **11,** 1–26 (2015).

Capsaspora owczarzaki as a unicellular model to study co-option of the ancestral integrin adhesome

Helena Parra[1], Matija Harcet[1,2], Alberto Pérez Posada[1], Nicholas H. Brown[3], Iñaki Ruiz-Trillo[1,4,5]

[1] Institut de Biologia Evolutiva (CSIC-Universitat Pompeu Fabra), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Catalonia, Spain.
[2] Division of molecular biology, Ruđer Bošković Institute, Bijenicka cesta 54, 10 000 Zagreb, Croatia
[3] Dept of Physiology, Development and Neuroscience,University of Cambridge,Downing St,Cambridge CB2 3DY,UK
[4] Departament de Genètica, Universitat de Barcelona, Barcelona, Catalonia, Spain
[5] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain

Adhesion systems and signaling networks are both essential in multicellular organisms. Some elements of the adhesion and signaling pathways of metazoans, such as proteins from the integrin adhesome are conserved in their closest unicellular relatives. This means these proteins already existed in the unicellular ancestor of metazoans and that they were co-opted for a multicellular lifestyle. To understand how the integrin adhesome was co-opted at the onset of Metazoa, we aim to unravel its function in a close unicellular relative of animals, the filasterean *Capsaspora owczarzaki*. This protist is the closest unicellular relative to metazoans that contains in its genome the basic core of proteins that constitutes the integrin adhesome. The expression of these proteins is upregulated during the aggregative stage in culture conditions. In order to understand its role, we are developing some molecular and genetic tools, such as immunostaining, transfection, and CRISPR/cas9. We will discuss preliminary data on the localization of several cytoskeletal and adhesion proteins of the integrin adhesome in *C. owczarzaki*, obtained by overexpression and by immunostaining with antibodies raised against our proteins of interest. We will also discuss the development of CRISPR system in this organism with the aim to develop a complete model system to analyze the origin of animals.

Purifying selection is acting on human long non-coding RNAs

Cinta Pegueroles[1,2], Toni Gabaldón[1,2,3]

[1] Bioinformatics and Genomics Programme. Centre for Genomic Regulation (CRG). Dr. Aiguader, 88. 08003 Barcelona, Spain
[2] Universitat Pompeu Fabra (UPF). 08003 Barcelona, Spain.
[3] Institució Catalana de Recerca i Estudis Avançats (ICREA), Pg. Lluís Companys 23, 08010 Barcelona, Spain.

Long non-coding RNAs' (lncRNAs) functionality is still under debate. Despite they are transcribed in all organisms studied so far, previous studies showed that they are expressed at low levels, that they are poorly conserved across species, and suggested that in those species having an small effective population size, selection may be not enough to counteract the effect of drift. Here we evaluate signatures of selection in lncRNAs annotated in human using three types of analysis: conservation across species, patterns of polymorphism within a species, and relationships between sequence constraints and secondary structure. In all analyses, we included a curated dataset of experimentally characterized human lncRNAs, as a reference for truly functional lncRNAs. We found evidence of purifying selection acting on lncRNAs and we describe that RNA secondary structure constrains sequence variation in lncRNAs. Importantly, this relation is independent of the GC content and the presence of splice-related motifs. We conclude that numerous predicted lncRNAs may be functional and may play key roles that remain to be discovered.

North African Populations: Insights from whole Y chromosome sequences

Neus Solé-Morata[1], Carla García-Fernández[1], Vadim Urasin[2], David Comas[1], Francesc Calafell[1]

[1] Institute of Evolutionary Biology (CSIC-UPF), Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, Catalonia, Spain.
[2] YFull – Research Group, Russia.

The analysis of uniparental markers demonstrated that E-M81 is the most important North African lineage (Fadhlaouid et al. 2013). It has been found at high frequencies (71%) in Noertheastern Africa, with its presence decreasing from east to west. Moreover, low frequencies of E-M81 (7%) have been found in the Iberian Peninsula, suggesting limited gene flow (Bosch et al. 2001). However, little is known about the phlyogeographic structure of this haplogroup. For the first time, we analyze whole Y chromosome sequences from 32 males selected for belonging to the E-M81 haplogroup. The analysis of whole Y chromosome sequences has enabled the discovery of new variants defining new subclades within the E-M81 branch. Those variants have been validated using Taqman probes in more than 200 North African samples, from which we also have information of the 17 Y-STRs contained in the AmpFlSTR®YFiler® PCR Amplification kit. In conclusion, the present work has increased the knowledge of this important but ignored Y chromosomal branch, by refining the phylogeographic structure of the North African paternal lineage.

Impact of Misidentified Samples on Population Genomics of Non Model Organisms

Marc Vega[1,*], Marta Pascual[1], Enrique Macpherson[2], Carlos Carreras[1]

[1] Department of Genetics, Microbiology and Statistics and IRBio, University of Barcelona, Av. Diagonal 643, E-08028 Barcelona, Spain.
[2] Centre d'Estudis Avançats de Blanes (CEAB-CSIC), Car. Acc. Cala St. Francesc 14, Blanes 17300 Girona, Spain.
[*]vegagalcera@gmail.com

High-throughput sequencing has opened new research opportunities on non-model organisms. For instance, Genotyping by sequencing (GBS) allows population genomics analysis based on genotyping Single Nucleotide Polymorphisms (SNPs) genome wide. However, these organisms often lack a reference genome for SNP calling and thus, obtaining SNPs could be potentially very sensitive to the presence of reads from other species. We analysed 380 individuals for a study targeting *Symphodus tinca* across the Mediterranean and the Black Sea. The results were not consistent with the expected population structuring and we thus screened the 20 outlier samples by sequencing the Cytochrome Oxidase I gene (COI). We detected them as having been morphologically misidentified and belonging to four different species, three *Symphodus* species and one *Serranus* species. The removal of these samples resulted in the expected population structuring of the target species. In order to evaluate the effect that these misidentified samples had on the number of SNPs found in the whole dataset we performed two different sets of analyses. On one hand, we evaluated the effect of the phylogenetic distance between misidentified and target species by analysing each species independently and by combining them in groups of 2, 3 and 4 different species with similar number of samples per species. On the other hand, we evaluated the effect of the proportion of misidentified samples on the identification of shared SNPs by changing the proportion of samples of two *Symphodus* species with equal number of individuals analysed with these techniques in the Western Mediterranean. As a general effect, the inclusion of misidentified samples reduced the number of SNPs shared among samples but increased the total number of SNPs identified. Greater phylogenetic distances and higher proportion of misidentified samples increased this effect. Our results highlight the importance of detailed analysis to detect outlier samples that may interfere with the population genomic study and for this reason we developed a protocol of filters to detect and remove misidentified samples when working on non-model organisms.

New genes and functional innovation in mammals

José Luis Villanueva-Cañas[1,*], Jorge Ruiz-Orera[1], Isabel Agea[1], M. Mar Albà[1,2]

[1] Evolutionary Genomics Group, Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Universitat Pompeu Fabra (UPF), Barcelona, Spain
[2] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain
*jlvillanueva84@gmail.com

The birth of new genes has led to remarkable physiological adaptations. Examples are the anti-freeze proteins that confer resistance to cold temperatures in Antarctic fishes or the caseins that transport calcium in the milk of mammals. Such lineage-specific genes have undergone radical sequence modifications from an ancestor gene or emerged *de novo* from previously non-coding regions of the genome. Despite their important role in adaptive processes, the majority of lineage-specific genes remain poorly characterized at the functional level. Here we have used the proteomes and transcriptomes of 30 mammalian species to build a comprehensive census of mammalian-specific genes families. By combining gene annotations and transcriptome assemblies we identify 4,749 multi-species genes families plus 4,009 species-specific genes. The set is significantly enriched in genes previously reported to have originated *de novo*. Consistent with previous studies we find that the genes tend to be short and often expressed in testis. However, we discover that they are also significantly overrepresented in salivary gland, spleen, fallopian tube, skin and breast. The proteins encoded by the new genes are often secreted and/or have roles in the immune system. These features, together with their strong depletion in acidic residues, suggest that a fraction of them probably has antimicrobial activity.

Key words: *de novo* gene, species-specific genes, orphan gene, lineage-specific gene, evolutionary innovation, adaptive evolution, mammals

Detecting positive selection in admixed populations from Ethiopia.

Sandra Walsh*, Hafid Laayouni, Jaume Bertranpetit.

Institut de Biologia Evolutiva. Dr. Aiguader 88 08003 Barcelona.
*sandra.walsh@upf.edu

Ethiopia is one of the most interesting places for human evolutionary history since we find most of the fossil evidence for human origins as well as a big linguistic and cultural diversity. In this context, the study of Ethiopian populations is fundamental in order to understand African genetic variability, the demographic processes that this area has undergone and the genetic adaptations that appeared during modern human evolutionary history. Here, we analysed 120 Ethiopian whole genome sequences from five Ethiopian populations, covering a wide range of ethnolinguistic groups (four Afro-asiatic and one Nilotic speaking populations) and geographical locations in Ethiopia. Ethiopians hold one of the highest genetic diversities in Africa, there is a clear population structure that correlates with linguistic affiliation and evidence of European genetic introgression among the Afro-asiatic speaking groups. Only a few studies revealed signals of positive selection in Africa and moreover the detection of regions under selection in admixed populations is not trivial. In this study, we detected positive selection in admixed African populations and show that there is an enrichment of Eurasian haplotypes in the top regions under selection.

Nom	Cognoms	Institució
Montserrat	Aguadé	UB
Marta	Álvarez	UB
Mªcristina	Aresté Calero	IBE (CSIC-UPF)
Margarita	Asensio Casero	INS La LLauna(Badalona)
Antonio	Barbadilla	UAB
Maite Garazi	Barrón Aduriz	IBE
Simone Andrea	Biagini	UPF
Cesar	Blanché	UB
Elena	Bosch Fuste	UPF
Mario	Cáceres Aguilar	UAB
Gemma	Calabria	
Francesc D'assis	Calafell Majo	UPF
Juanma	Calvo Martín	UB
Carlos	Carreras Huergo	UB
Sònia	Casillas	UAB
Silvia	Chafino Aixa	IBE (CSIC-UPF)
David	Comas Martinez	UPF
Marta	Coronado Zamora	UAB
Pere	Crosas Farrés	Ensenyament secundari
Josefa	Cruz	IBE- CSIC
Marc	De Manuel Montero	UPF
Begoña	Dobon Berenguer	UPF
Paula	Escuer Pifarré	UB
Domènec	Farré	UB
Albert	Ferrer Mata	UB
Xavier	Florenza Garcia	IBE (CSIC-UPF)
André	Flores Bello	UPF
Neus	Font Porterias	UPF
Clàudia	Fontserè Alemany	IBE (CSIC-UPF)
Víctor	Gámez Visairas	UAB
Sònia	Garcia	IBB-CSIC
Carla	García Fernández	UPF
Mª Del Pilar	Garcia Guerreiro	UAB
Teresa	Garnatje	IBB-CSIC
Carla	Giner-delgado	UAB
Xavier	Grau-bové	IBE (CSIC-UPF)
Lain	Guio Leiman	IBE (CSIC-UPF)
Sergi	Hervás	UAB
Vanessa	Jacovas	UPF
Aleksandra	Kozyczkowska	IBE
Hafid	Laayouni	UPF
Irene	Lobón Garcia	UPF
David	López Escardó	IBE (CSIC-UPF)
Jesus	Lozano Fernandez	University of Bristol

Carles	Llorca	
Tomàs	Marquès Bonet	UPF
David	Martin	IBE (CSIC-UPF)
Clara	Martín Muñoz	UB
David	Mas Ponte	UPF
Alex	Mas Sandoval	UPF
Miriam	Merenciano	IBE (CSIC-UPF)
Francesc	Mestres	UB
Rafael Tomoya	Michita	FEDERAL UNIVERSITY OF RIO GRANDE
Mayukh	Mondal	UPF
Rafael	Montiel	Langebio, Cinvestav-IPN, Mexico
Antoni	Moreno	UB
Bàrbara	Negre	UAB
Isaac	Noguera	UAB
Jessica Marie	Nye	UPF
Eduard	Ocaña	IBE (CSIC-UPF)
Víctor	Ojeda Martín	UB
Dorcas Juana	Orengo	UB
Helena	Parra Acero	IBE
Marta	Pascual	UB
Cinta	Pegueroles	CRG
Alberto	Pérez Posada	IBE (CSIC-UPF)
Laia	Planella Albi	UDG
Sergio	Pous Sanchez	UB
Eva M.	Puerma	UB
Marta	Riutort León	UB
Mª Inés	Roldan	UDG
Valèria	Romero-soriano	UAB
Núria	Ros i Rocher	IBE (CSIC-UPF)
Vicent	Rossi	CSIC-UIB
Julio	Rozas	UB
Lara	Rubio Arauna	UPF
Iñaki	Ruiz-trillo	IBE
Alicia	Sánchez Arroyo	IBE (CSIC-UPF)
José Francisco	Sanchez Herrero	UB
Carmen	Segarra Robert	UB
Gerard	Serra Vidal	UPF
Neus	Solés Morata	UPF
Juan Antonio	Subirana	UPC-IMIM
Apostolia	Topaloudi	UPF
Anna	Ullastres Coll	IBE (CSIC-UPF)
Joan	Vallès Xirau	UB
Marc	Vega Galcerà	UB
José Luis	Villanueva Cañas	FIMIM / IBE
Pablo	Villegas Mirón	UPF
Sandra	Walsh Capdevila	UPF