cirad

IRD

DynaGeV

Dynamique des Génomes Végétaux

Workshop on Plant Genome Dynamics and Evolution

CIRAD Campus de LAVALETTE - MONTPELLIER

Le 8 et 9 Juin 2017

GenomeHarvest
diversity, organization and dynamics

agropolis fondation

INVESTISSEMENTS D'AVENIR

Jeudi 8 Juin 2017

8h30-8h50: Welcome of participants / Accueil des participants
8h50-9h00: Introduction - practical information / détails pratiques

SESSION 1- (ANIMATRICE MALIKA AINOUCHE)

9h00-9h40: **GRIMANELLI Daniel** Dynamics and functional roles of DNA methylation during plant reproductive development.
9h40-10h05: **RICHARDS Christina** Transcriptomic responses to challenging environments in Spartina alterniflora populations
10h05-10h30: **KELLER Jean** *Lupinus* (Fabaceae) nodulome: A RNA-seq approach of differentially expressed genes in a context of symbiont switch in the plant/Bradyrhizobium nitrogen-fixing association
10h30-10h55: **ROUSSEAU-GUEUTIN Mathieu** Endosymbiosis and the evolutionary fate of genes duplicated in different cellular compartments.

10h55-11h15 Coffee Break

SESSION 2 - (ANIMATRICE ANNE MARIE CHÉVRE)

11h15-11h40: **STAVRINIDES Anna** Highly duplicated Medium-Chain Dehydrogenase/Reductases increase chemical diversity and hint at the creation of biosynthetic gene clusters
11h40-12h05: **KHEDIM Thinhinan** Karyotype diversity and evolution in the disjunct *Allium* subgenus Amerallium Traub (Amaryllidaceae)
12h05-12h30: **AISSIOU Fella** *Brassica rapa* genetic resources from Algeria: diversity and population structure based on SSR markers
12h30-12h55: **FARHAT Perla** Exceptional high rate of polyploidy screened in *Juniperus* genus

13h00-14h15: Lunch at CIRAD

SESSION 3 - (ANIMATEUR MATHIEU ROUSSEAU)

14h15-14h30: **D'HONT Angélique** GenomeHarvest project
14h30-14h55: **AHMED Dalel** A phylogenomic study based on Genotyping By Sequencing unravels the interspecific mosaic structures of the cultivated *Citrus* genomes
14h55-15h20: **GARCIA SANTOS Joao** Using the 3k genomes to decipher the mosaic structure of genome diversity in Rice
15h20-15h45: **COTTIN Aurélien** Evaluation of methodologies for the characterization of plant mosaic genomes

15h45-16h05: *Coffee Break*

16h05-16h30: **MARTIN Guillaume** Evolution of the banana genome is impacted by large chromosomal translocations
16h30-17h55: **DUPOUY Marion** Characterization of two large chromosomal translocations in Musa acuminata ssp. burmannicoides "Calcutta4"

17h55-17h20: **BAURENS Franc-Christophe** Large chromosomal structural variations between M. acuminata and M. balbisiana and their consequences on chromosome recombination and segregation in a AAAB polyploidy context.

17h20-17h:45 **PELE Alexandre** New insights on homologous recombination in polyploids: the striking case of *Brassica* allotriploids

20h RESTAURANT in " Bouchon Saint Roch" 14 Rue du Plan d'Agde MONTPELLIER

Vendredi 9 Juin 2017

SESSION 4- (ANIMATEUR PHILIPPE LASHERMES)

9h00-9h40: **GALTIER Nicolas** Comparative population genomics in animals: genetic diversity, adaptive rate, species barrier

9h40-10h05: **De OLIVEIRA Romain** Identification et analyse des variations structurales du génome chez le blé

10h05-10h30: **SOURDILLE Pierre** Overview of actual and ancestral recombination in relationship with the sequence in wheat: focus on chromosome 3B

10h30-10h55: **SABOT François** Evolution of African rice pangenomes through domestication and post-domestication

10h55-11h15 Coffee Break

SESSION 5 - (ANIMATEUR OLIVIER CORITON)

11h15-11h40: **PANAUD Olivier** Crops, man and TEs. The 3000 rice genomes unravel TE-driven genome dynamics in the field

11h40-12h05: **DUPEYRON Mathilde** Distribution of Divo in Coffea genomes, a poorly described family of angiosperm LTR-Retrotransposons

12h05-12h30: **FRUCHARD Cécile** Assembling a large plant Y chromosome using long read sequencing

12h30 – 13h00: **Round table**- Exchanges – End of the conference / Table Ronde-Echanges-Clôture de la conference

13h00: Lunch at CIRAD

RESUMES/ABSTRACTS

SESSION 1

(JEUDI 8 JUIN 2017)

Transcriptomic responses to challenging environments in Spartina alterniflora populations.

Christina L. Richards[1], Mariano Alvarez[1,2], Julie Ferreira de Carvalho[3,4], Armand Cavé-Radet[3], Abdelhak El Amrani[3], Tammy E. Foster[5], Sydney Moyer[1], Armel Salmon[3], Malika L. Ainouche[3]

[1]Department of Integrative Biology, University of South Florida, 4202 East Fowler Avenue, Tampa FL 33620 USA.

[2]Present address: Department of Biology, Duke University, 130 Science Drive, Box 90388, Durham, NC 27708 USA.

[3]Université de Rennes I, UMR CNRS 6553 EcoBio, Campus Beaulieu, Rennes, France

[4]Present address: INRA IGEPP, Biodiversity and Polyploidy group, Domaine de la Motte, 35653 Le Rheu France

[5]Kennedy Space Center Ecological Program, SR 405, Titusville, FL 32899 USA.

Abstract

Despite the severe impacts of the 2010 *Deepwater Horizon* oil spill which occurred in the Gulf of Mexico, the hexaploid foundation species *Spartina alterniflora* proved resilient to heavy oiling, providing an opportunity to identify mechanisms of response to the anthropogenic stress of crude oil exposure. We hypothesized that a population-scale transcriptomic approach in recovering *S. alternflora* plants, combined with reverse genetic approaches, would identify transcripts that regulate the response of plants recovering from dieback due to oil exposure. We identified differentially expressed transcripts between oil affected and unaffected populations across the Gulf of Mexico using a custom microarray. We used comparative methods to show that xenobiotic response pathways, or the xenome, responded to oil in *S. alterniflora* in a manner distinct from that of the model plant *A. thaliana*, and was significantly enriched with alpha/beta hydrolase and glycosyltransferase genes. We constructed gene interaction networks and used T-DNA insertion lines of the model grass *Brachypodium distachyon* to identify and confirm the fitness effects of candidate genes under oil stress. Our results suggest a significant role for homologs of SUVH5 and ATTPS21 in modulating the response to crude oil stress in *B. distachyon* and *S. alterniflora*.

Lupinus (Fabaceae) nodulome: A RNA-seq approach of differentially expressed genes in a context of symbiont switch in the plant/*Bradyrhizobium* nitrogen-fixing association.

Keller Jean[1], Imperial Juan[2], Ruiz Tomàs[2], Privet Kaïna[1], Lima Oscar[1], Salmon Armel[1], Aïnouche Abdelkader[1], Cabello-Hurtado Francisco[1]

[1] *UMR CNRS 6553 Ecobio, OSUR (Observatoire des Sciences de l'Univers de Rennes), Université de Rennes 1, 35042 Rennes, France;* [2] *Departamento de Biotecnología (ETSIA) and Centro de Biotecnología y Genómica de Plantas (CBGP), Universidad Politécnica de Madrid, 28040 Madrid, Spain*

jean.keller@univ-rennes1.fr

Biological nitrogen fixation is one of the major source of nitrogen in natural ecosystems and has a critical importance for environment-friendly agriculture (Graham and Vance, 2003). The association between legumes (Fabaceae) and bacteria collectively termed Rhizobia is one of the most widespread mutualistic symbiosis. Nitrogen fixation occurs in specialized root structures called nodules. Root infection and nodule organogenesis are regulated by a complex genetic pathway that has been extensively studied since the last thirty years (Desbrosses and Stougaard, 2011; Oldroyd et al., 2011). However, different aspects of the legume root-nodule symbiosis (RNS) remain unclear. This is the case of the symbiotic specificity. Indeed, some legume species are functionally nodulated by diverse rhizobium species whereas other are restricted to one strain of rhizobia. In addition, host and symbiont can exhibit different gradient of compatibility (Wang et al., 2012). To study this process, we performed a cross-inoculation between three lupine species (*Lupinus* genus, belonging to the under-investigated tribe of Genistoids) and two *Bradyrhizobium* strains presenting contrasted degrees of compatibility. RNA was extracted from nodules and roots and sequenced on an Illumina platform. We constructed, annotated and compared the three first lupine nodulomes. Results revealed significant expression changes in the genetic pathway controlling the infection and the nodule organogenesis. In addition, a functional analysis of the differentially expressed genes highlighted the central role of hormones and of secondary, carbon and nitrogen metabolisms, as well as the implication of plant defences in the response to compatible or non-compatible *Bradyrhizobium.*

Desbrosses, G.J., Stougaard, J., 2011. Root Nodulation: A Paradigm for How Plant-Microbe Symbiosis Influences Host Developmental Pathways. Cell Host Microbe 10, 348–358. doi:10.1016/j.chom.2011.09.005

Graham, P.H., Vance, C.P., 2003. Legumes: Importance and Constraints to Greater Use. PLANT Physiol. 131, 872–877. doi:10.1104/pp.017004

Oldroyd, G.E.D., Murray, J.D., Poole, P.S., Downie, J.A., 2011. The Rules of Engagement in the Legume-Rhizobial Symbiosis. Annu. Rev. Genet. 45, 119–144. doi:10.1146/annurev-genet-110410-132549

Wang, D., Yang, S., Tang, F., Zhu, H., 2012. Symbiosis specificity in the legume - rhizobial mutualism: Host specificity in root nodule symbiosis. Cell. Microbiol. 14, 334–342. doi:10.1111/j.1462-5822.2011.01736.x

Endosymbiosis and the evolutionary fate of genes duplicated in different cellular compartments.

M. Rousseau-Gueutin[1,2], J. Keller[1], G.E. Martin[3], J. Morice[2], J. Boutte[1], E. Coissac[4], M. Ourari[5], M. Aïnouche[1], A. Salmon[1], F. Cabello-Hurtado[1], A. Aïnouche[1].

[1] UMR CNRS 6553 Ecobio, OSUR, Université de Rennes 1, 35042 Rennes, France, [2] IGEPP, INRA, Agrocampus Ouest, Université de Rennes 1, BP35327, 35653 Le Rheu Cedex, France, [3] CIRAD, UMR AGAP, F-34398 Montpellier, France, [4] Laboratoire d'Ecologie Alpine, CNRS - Université de Grenoble 1 - Université de Savoie, 38041 Grenoble, France, [5] Département des Sciences Biologiques, Université Abderrahmane Mira, 06000 Bejaia, Algeria. E-mail: mathieu.rousseau-gueutin@inra.fr

Eukaryotic cells arose more than a billion years ago through endosymbiotic engulfment of free-living bacteria that were gradually converted into mitochondria and chloroplasts respectively (Margulis, 1970, Deusch *et al.* 2008). From these ancestral prokaryotes, eukaryotes acquired the novel biochemistry of oxydative phosphorylation and photosynthesis. Soon after these endosymbiotic events, there has been a continuous deluge of organellar DNA entering the nuclear genome (Huang *et al.* 2003). This DNA transfer was and still is a major driving force in eukaryote evolution as it led to the functional transfer of more than a thousand organellar genes to the nucleus and also to the creation of novel nuclear genes (approximately 4,500 of Arabidopsis genes were acquired from the cyanobacterial ancestor of plastids) (Martin *et al.* 2002). Following to the functional transfer of an organellar gene to the nucleus, two functional copies (with the same function) coexist in two separate genetic compartments until one becomes defunct. If the nuclear and organelle-encoded copies were equally efficient, loss of functionality is presumably the result of chance mutation silencing one or the other copy, thus favoring the retention of the organellar copy. If the nuclear copy becomes defunct, the whole process can be repeated again. However, if the functionality of the organellar copy is lost, then the nucleus becomes the permanent location of that gene as there is no DNA transfer from the nucleus to the chloroplast. This explains the net decrease of cytoplasmic organellar genomes (now only encode approximately 80 different genes compared with several thousand genes in the bacterial ancestor) and the increased genetic influence of the nucleus. The functional transfer/replacement of chloroplast genes is still ongoing in plants (Rousseau-Gueutin *et al.* 2013), most particularly in the Fabaceae family. From the assembly of five *Lupinus* chloroplast genomes (Fabaceae), we found that the chloroplast *rps16* gene was pseudogeneized in some *Lupinus* species and was replaced by a nuclear gene of organellar origin (Ueda *et al.* 2008). By searching the *rps16* gene in all chloroplast, mitochondrial and nuclear plant genomes available to date, we found that this organellar gene was transferred to the nucleus more than a billion years ago and never lost its functionality since then. Thus, two functional copies of this gene have been coexisting in two different cellular compartments for all this time and the chloroplast gene copy has only been recently lost in some plants. Here, the evolutionary fate of this duplicated gene (chloroplast and nucleus) and the putative reasons explaining this long cohabitation will be presented (Keller *et al.* 2017).

- Deusch *et al.* (2008) Molecular Biology and Evolution 25: 748–761.

- Huang, Ayliffe and Timmis (2003) Nature 422: 72–76.

- Keller, Rousseau-Gueutin *et al.* (2017) DNA Research. doi: 10.1093/dnares/dsx006

- Margulis (1970) Yale University Press, New Haven.

- Martin *et al.* (2002) PNAS 99: 12246–51.

- Rousseau-Gueutin *et al.* (2013) Plant Physiology 161 (4) 1918-1929.

- Ueda, Fujimoto, Arimura, Tsutsumi, Kadowaki (2008) Molecular Biology and Evolution 25: 1791–1793

SESSION 2

Highly duplicated Medium-Chain Dehydrogenase/Reductases increase chemical diversity and hint at the creation of biosynthetic gene clusters

Anna Stavrinides[1], Sarah O'Connor[2]

1.UMR DIADE, Institut de Recherche pour le Développement, Montpellier, France

2. Biological Chemistry, John Innes Centre, Norwich, UK

anna.stavrinides@ird.fr

Plant natural products are a rich source of bioactive compounds useful in medicine and agriculture. These compounds often are the products of long and complicated biosynthetic pathways involving many different enzymes. However, often the compounds of interest are produced in very low amounts which makes extraction costly and inefficient. In order to increase yields and reduce costs elucidation of these pathways is necessary for plant genetic engineering and for transfer of the pathway to a heterologous host.

The medicinal plant *Catharanthus roseus* produces hundreds of monoterpene indole alkaloids and has evolved countless specialized enzymes. Recently, a group of homologous Medium-Chain Dehydrogenase/Reductases (MDRs) was shown to act at a branch point of this biosynthetic pathway and each presented a different product profile [1]. A phylogenetic tree of MDRs of *C. roseus* indicated that half of these genes (including the secondary metabolism MDRs) are derived from recent duplications and subsequent divergence leading to neofunctionalisation.

Thanks to the genome sequence of *C. roseus* [2] the genomic context of these enzymes was investigated. It was found that these MDRs were in tandem duplication blocks, some more recent than others. One cluster (10 MDRs) contained multiple enzymes with diverse functions. Surprisingly, similar clusters with tandemly-duplicated MDRs are found in most sequenced plant species. Phylogenetic analysis of these homologous clusters revealed that the clusters are not conserved but rather are constantly in flux and contain different tandem copies between distantly related species. This supports the hypothesis that duplications are the source of new enzyme functions and is a dynamic process in plants.

Unlike bacteria and fungi, which conveniently possess clusters of biosynthetic genes (operons and operon-like clusters), plants rarely present such clusters. Notable examples include the avenacin biosynthetic cluster in oat and the thalianol biosynthetic cluster in Arabidopsis [3]. In *C. roseus* two of the MDRs found to be involved in the MIA pathway are located next to an upstream P450 enzyme. This, together with the phylogenetic information, suggests active recruitment of enzymes common to the MIA pathway to an operon-like cluster has been taking place in *C. roseus* and could, through subsequent evolution, recruit more MIA biosynthetic genes.

[1] Stavrinides, Tatsis et al. Nat. Commun. 2016

[2] Buell and Kim, in press

[3] Field and Osbourn, Science, 2008

Karyotype diversity and evolution in the disjunct *Allium* subgenus *Amerallium* Traub (Amaryllidaceae)

Khedim Thinhinan[1], Aïnouche Abdelkader[2], Amirouche Nabila[1], Aïnouche Malika Lily[2], Amirouche Rachid.[1]

[1] *Université des Sciences et de la Technologie Houari Boumediene, Faculté des Sciences Biologiques, LBPO, BP n° 32 El-Alia, Bab-Ezzouar, 16111 Alger, Algérie.*
[2] *Université de Rennes 1, Campus de Beaulieu, UMR, Ecobio, Bat. 14, 35042 Rennes, France.*
tinakhedim@yahoo.fr ; tkhedim@usthb.dz, ramirouche@usthb.dz.

Allium subgenus *Amerallium* Traub is a large group containing about 140 species with disjunct distribution in the Mediterranean region, North America and Western Asia. This subgenus constitutes a monophyletic group, which is considered as the most ancient evolutionary lineage within genus *Allium*. Molecular markers revealed an Old World origin of the North American species. This is in agreement with cytogenetic evidence which indicates that all species of the *Amerallium* group have only *x*=7 in the New World, while a wide range of chromosome base numbers is observed in the Old World, x=7, 8, 9, 10 or 11.

In North Africa and especially in Algeria, species belonging to subgenus *Amerallium* are poorly explored and often taxonomically confused. This study aims to investigate karyotype diversity and phylogenetic relationships in order to clarify inter and intraspecific circumscription and to identify the major mechanisms involved in their evolution. Karyological and molecular analyzes were applied to 40 natural populations sampled in Northern Algeria throughout the bioclimatic transect from the littoral to the arid zones. The populations belong to six taxa: *A. subhirsutum* L., *A. subvillosum* Salzm., *A. roseum* L., *A. odoratissimum* Desf., *A. triquetrum* L. and *A. chamaemoly* L.

The karyological screening revealed a high dysploidy and polyploidy with six cytotypes, 2n=14, 28; 2n=16, 32; 2n=18 and 2n=22, and four base numbers *x*=7, 8, 9 and 11 respectively. Multivariate analysis of karyotype parameters shows significant difference in the asymmetry indexes values, particularly within the polyploids. Polyploidy concerns only the base numbers *x*=7 and *x*=8. Molecular phylogenetic analyzes were based on ITS rDNA and chloroplast trnL-trnF and trnD-trnT sequences. Tree topologies showed resolved relationships between the North African clade and the rest of Old Word and New Word groups. Results highlight the originality of the North African *Allium* species characterized by an exceptional karyotypic diversity. Additional multigenic approach should help elucidating the origin of polyploidy within this enigmatic lineage.

Keywords: *Amerallium,* evolution, karyotype, dysploidy, polyploidy, Algeria.

Brassica rapa genetic resources from Algeria: diversity and population structure based on SSR markers

Aïssiou F[1], Falentin C[2], Laperche A[2], Rousseau-Gueutin M[2], Abrous O[1], Chèvre AM[2], Hadj-Arab H[1]

[1] *University of Science and Technology Houari Boumediene, USTHB, Faculty of Biological Sciences FSB, Bab-ezzouar, BP 32 El-Alia, 16111 Algiers, Algeria*

[2] *UMR IGEPP, INRA Agrocampus Ouest, University Rennes1 - Domaine de la Motte, BP 35327 - F - 35653 Le Rheu cedex - France*

e.mail : hhadj_arab@yahoo.fr anne-marie.chevre@inra.fr

In a context of global warming, it is particularly important to preserve the genetic diversity of spontaneous and cultivated species that grow in contrasting environments and to elucidate the conditions for maintaining and exploiting this diversity. *Brassica rapa* is an economically important crop that contains spontaneous forms of potential genetic resources for pathogen resistance and environmental stress tolerance that can be introduced into cultivated Brassicaceae. In Algeria, *B. rapa* is represented by many cultivated and spontaneous forms. To evaluate the genetic diversity of this species, we used 69 microsatellite markers to assess the genetic variation of wild and cultivated accessions from Algeria, Europe and Asia. We identified more than 700 alleles with an average of 10 alleles per marker. Population structure analysis revealed three groups; the first contains the Algerian wild accessions of *B. rapa*, the second includes populations from Algerian cultivated accessions, this group seems to be close to *B. napus* lines and *B. rapa* populations subsp. *trilocularis*, var. *rapifera*. The third group includes accessions from Europe and Asia. The data can be used for defining a *B. rapa* Algerian collection useful for further investigations for agronomic traits.

Key words: *Brassica rapa*, genetic diversity, SSR.

Exceptional high rate of polyploidy screened in *Juniperus* genus

Farhat Perla [1,2], Hidalgo Oriane[3], Adams Robert,[4], Siljak-Yakovlev Sonja [2], Leitch Ilia [3], Robert Thierry [2], Bou Dagher Kharrat Magda [1]

[1] *Laboratoire 'Caractérisation Génomique des Plantes', Faculte´ des Sciences, Universite´ Saint-Joseph, Campus Sciences et Technologies, Mar Roukos, Mkalles, BP: 1514 Riad el Solh, Beirut 1107 2050, Lebanon.*

[2]*Université de Paris-Sud, Laboratoire Ecologie, Systématique, Evolution, UMR CNRS 8079, Bât. 360, 91405 Orsay Cedex, France.*

[3]*Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK*

[4]*Biology Department, Baylor University, Box 97388, Waco, TX 76798.*

Abstract:

The genus *Juniperus L.* belongs to the Cupressaceae family. It is the most diverse genus of this family and the second largest genus of living conifers after *Pinus*. It has ca. 75 species belonging to three sections: *Caryocedrus*, *Juniperus* and *Sabina*. While polyploidy is considered rare in gymnosperms, and more particularly in conifers, recent investigations have revealed unexpected polyploidy cases in *Juniperus*. This led us to undertake a global ploidy screening, seeking to evaluate the extent of polyploid diversity within *Juniperus* and its impact on the evolution of the genus.

We have determined by flow cytometry the ploidy level for 108 *Juniperus* species and subspecies, using the collection of silica gel-preserved material of Robert Adams (the author of the book "Junipers of the World").

Our results showed that ca. 15% of studied taxa present C-values suggesting tetraploid level and one species, *J. foetidissima*, is a putative hexaploid. All polyploids detected so far belong to two clades within the *Sabina* section and their geographical distributions restrict to the old world. In addition, it was revealed some variations in genome size for the same ploidy level inside this genus.

In conclusion, *Juniperus* genus showed relatively high number of polyloids comparing to other conifer groups, and, together with *Ephedra*, represents exceptions within gymnosperms. These results open new venues for future research aimed at providing insights on whether polyploidy could translate into differences in population structure, genetic diversity and response to stress in a conifer lineage.

Keywords: *Juniperus,* gymnosperms, conifers, ploidy screening, silica gel-preserved material, flow cytometry.

SESSION 3

A phylogenomic study based on Genotyping By Sequencing unravels the interspecific mosaic structures of the cultivated *Citrus* genomes.

Dalel Ahmed[1], Franck Curk[1], François Luro[1], Gilles Costantino[1], Hélène Vignes[2], Pierre Mournet[2], Yann Froelicher[3] and Patrick Ollitrault[4]

[1]UMR AGAP, INRA, 20230, San-Giuliano, France, [2]UMR AGAP, Cirad, 34398, Montpellier, France, [3]UMR AGAP, Cirad, 20230, San-Giuliano, France, [4]UMR AGAP, Cirad, Station de Roujol, 97170, Petit-Bourg, Guadeloupe, France

dalel.ahmed@inra.fr

Reticulate evolution, including hybrid speciation, introgression and lateral gene transfer is frequent in plant species. When coupled with vegetative propagation, it results in mosaic genomes of large genomic fragments, from different species or sub species. These complex inter(sub)-specific genomic structures can support a major part of the phenotypic diversity organization. Cultivated *Citrus* are a good example of such genepool issued from reticulate evolution with limited number of further interspecific recombination. It is generally agreed that four ancestral taxa (*Citrus maxima*: pummelos, *Citrus reticulata*: mandarins, *Citrus medica*: citrons and *Citrus micrantha*: papedas) are the ancestors of all the cultivated citrus species. These four species, have undergone allopatric evolution. This has led to a strong genetic differentiation which is also found for many phenotypic characters. The so-called secondary species (sweet and sour oranges, grapefruits, lemons and limes) are the result of a reticulate evolution between these four ancestral taxa. Facultative apomixis (nucellar polyembrionny) then limited the number of interspecific meiosis cycle.

The aplication of NGS on reduced genome representation with methods such as GBS (Genotyping By Sequencing) coupled with the availability of a clementine reference sequence open the way for pangenomic studies of large populations. The objective of this work was to validate a GBS approach on citrus in order to identify a pangenomic panel of diagnostic markers (DSNPs) of each ancestral taxa and to decipher the phylogenomic structures of 56 citrus varieties representative of the ancestral taxa and secondary species. The DSNP panel was also used to analyse the phylogenomic structures of diploid and triploid recombining populations of the Cirad-Inra breeding programs.

GBS library were prepared with ApekI and a selective PCR to improve the depth of the analysis. 56 accessions were pooled and sequenced in one line of Illumina HISEQ-2000 (single reads). Diversity structure analysis showed that the varieties are distributed among the four ancestral taxa in perfect consistency with the previous studies carried out with SSRs, Indels and SNPs markers. The GBS approach is thus validated. 14926 DSNPs were identified. These diagnostic markers allowed to infer efficiently the majority of the phylogenomic karyotypes of the 56 *Citrus* accessions and revealed the interspecific recombination point in the diploid hybrids of breeding populations. For polyploid germplasm and hybrids, needing the evaluation of allelic doses, we were not able to infer these doses from the relative reads number at individual locus level. The potential of analysis at genome fragment level, covering numerous DSNPs, is under study.

This work demonstrate the potential of GBS for deciphering the phylogenomic structure of the modern citrus varieties and recent hybrids of breeding populations. It bring new insights on the origins of citrus fruit and open the way for genetic associations studies and QTLs analysis based on phylogenomics and further to genomic selection.

Key words: *Citrus*, evolution, interspecific structure, GBS, SNPs, phylogenomic inference

Using the 3k genomes to decipher the mosaic structure of genome diversity in rice

Santos J.[1], Billot C.[1], Glaszmann J. C.[1]

Introgression among rice populations is worth studying for its role as a path to adaptation. Along human migrations, gene flow between cultivars and wild or primitive cultivated forms have generated new types which thus harbor admixed genomes with distinct components traceable to early crop history. Recent analyses based on massive sequencing efforts have enabled detailed studies of crop evolution in rice that revealed introgressions allowing the spread of domestication factors across varietal groups as well as the secondary hybrid origin of some varietal clusters. Yet distinct views still coexist as to the global interpretation of the data, featuring one *vs* multiple domestication events and diverse scenarios for the origin of secondary varietal groups. Contrasting hypotheses make it difficult to draw conclusions from the data given their weight on the premises underlying any analysis. Nonetheless, by recognizing the existence of more or less cohesive groups of rice varieties, important information can still be extracted as to their interactions. The results presented here are a still frame of our analysis of the exchanges between the major clusters of rice genetic diversity, Japonica, Indica and *circum*Aus, as can be determined by their relative differentiation, and the concomitant retrieval of the cryptic *circum*Basmati genetic signature.

Keywords: *rice, gene flow, admixed genomes, cBasmati, mosaic*

[1] CIRAD UMR AGAP, Montpellier

Evaluation of methodologies for the characterization of plant mosaic genomes

Aurélien Cottin[1], Benjamin Penaud[1], Guillaume Martin[1], Joao Santos[1], Franck Curk[2], Angélique D'Hont[1], Manuel Ruiz[1], Jean-Christophe Glaszmann[1], Nabila Yahiaoui[1], Mathieu Gautier[3]

[1]CIRAD, UMR AGAP, F-34398 Montpellier, France

[2]INRA, UMR AGAP, F-20230 San Giuliano, France

[3]INRA, UMR CBGP, F-34988 Montferrier-sur-Lez Cedex, France

CIRAD - TA A-108 / 03 Avenue Agropolis - 34398 Montpellier Cedex 5 - France

aurelien.cottin@cirad.fr

Hybridization events between species and subspecies are considered as major evolutionary steps, possibly contributing to the advent of new phenotypes. These events are widespread in several crop species and are expected to produce genomes with a mosaic structure of sequence blocks from different ancestry. Characterizing the inter(sub)specific mosaic structure of crop plant genomes that result from recent hybridization events can help understanding how they were formed, their domestication history, and possibly the ancestral origin of phenotypic traits.

With the development of NGS genotyping technologies, several population genomics approaches have been proposed to infer the ancestry of genome segments, by comparing polymorphism patterns across individuals along chromosomes. However, these Local Ancestry Inference (LAI) methods have mainly been developed for applications in animal models, and human most particularly. They are based on assumptions which do not always fit plant models due to more complex genome structures (e.g. different ploidy levels, variable heterozygosity levels within species) or different reproductive systems (e.g., vegetative propagation, selfing). In this context, there is a need to evaluate available methods on plant models.

To that end, we developed a small and flexible R program to simulate data under a wide variety of scenarios representative of plant model characteristics. We use this tool to evaluate two main types of LAI methods: exploratory approaches (based on multivariate analysis) and full probabilistic approaches (based on Hidden Markov Model). First results will be presented and discussed.

Evolution of the banana genome is impacted by large chromosomal translocations

Guillaume Martin[1], Françoise Carreel[1], Olivier Coriton[2], Catherine Hervouet[1], Céline Cardi[1], Paco Derouault[1], Danièle Roques[3], Frédéric Salmon[4], Mathieu Rouard[5], Julie Sardos[5], Karine Labadie[6], Franc-Christophe Baurens[1], Angélique D'Hont[1]

[1]. *CIRAD, UMR AGAP, F-34398 Montpellier, France*

[2]. *Institut National de la Recherche Agronomique, Unité Mixte de Recherche 118, Amélioration des Plantes et Biotechnologies Végétales, BP 35327, F-35653 Le Rheu cedex, France*

[3]. *CIRAD, UMR AGAP, F-97130 Capesterre Belle Eau, Guadeloupe, France*

[4]. *CIRAD, UMR AGAP, F-97170 Petit-Bourg, Guadeloupe, France*

[5]. *Bioversity International, Parc Scientifique Agropolis II, 34397 Montpellier Cedex 5, France*

[6]. *Commissariat à l'Energie Atomique (CEA), Institut de Genomique (IG), Genoscope, 2 rue Gaston Cremieux, BP5706, 91057 Evry, France*

guillaume.martin@cirad.fr

Most banana cultivars are triploid derived from *Musa acuminata* (2n=2x=22), sometimes combined with *Musa balbisiana* (2n=2x=22). These species and subspecies diverged following geographical isolation in distinct Southeast Asian continental regions and islands. Contact between them was made possible by human migration and led to the selection of seedless parthenocarpic hybrids. *M. acuminata* subspecies were suggested to differ by a few large chromosomal rearrangements based on chromosome pairing configurations in inter-subspecies hybrids. We searched for large chromosomal rearrangements in a seedy *M. acuminata* ssp. *malaccensis* banana accession through mate-pair sequencing, BAC-FISH, targeted PCR and marker (DArTseq) segregation in its progeny. We identified a heterozygous reciprocal translocation involving two distal 3 Mb and 10 Mb segments from chromosomes 01 and 04, respectively, and showed that it locally generated high segregation distortions and reduced recombinations in its progeny. The two chromosome structures were found to be mutually exclusive in gametes and the rearranged structure was preferentially transmitted to the progeny. The rearranged chromosome structure was frequently found in triploid cultivars but within the wild accessions, it was only found within *malaccensis* sub-species accessions, thus suggesting that this rearrangement occurred in this sub-species. We propose mechanisms for the spread of this rearrangement in *Musa* diversity and propose that this structure may have played a role in the emergence of triploid cultivars.

Characterization of two large chromosomal translocations in *Musa acuminata ssp. burmannicoides* "Calcutta4"

Marion Dupouy[1], Guillaume Martin[1], Paco Derouault[1], Catherine Hervouet[1], Céline Cardi[1], Frédéric Salmon[2], Lyonel Toubi[3], Karine Labadie[4], Franc-Christophe Baurens[1], Angélique D'Hont[1]

[1]. *CIRAD, UMR AGAP, F-34398 Montpellier, France*
[2]. *CIRAD, UMR AGAP, F-97130 Capesterre Belle Eau, Guadeloupe, France*
[3]. *CIRAD, UMR AGAP, CRB Plant tropical, F-97170 Petit-Bourg, Guadeloupe, France*
[4]. *Commissariat à l'Energie Atomique (CEA), Institut de Genomique (IG), Genoscope, 2 rue Gaston Cremieux, BP5706, 91057 Evry, France*

marion.dupouy@cirad.fr

A few *Musa acuminata* subspecies are involved in banana cultivars. Theses subspecies diverged following geographical isolation in distinct Southeast Asian continental regions and islands. Observation of chromosome pairing irregularities in meiosis of hybrids between theses subspecies suggested the presence of large chromosomal structural variations.

We analyzed accession "Calcutta4" from the subspecies *burmannicoides* to search for chromosomal structural variations relative to the reference sequence assembly obtained with a *M. a.* ssp *malaccensis* accession.

A self progeny of Calcutta4 was genotyped by GBS (genotyping by sequencing). We observed linkage between markers from reference chromosome 2 and 8 and reference chromosome 1 and 9 suggesting the presence of two reciprocal translocations involving these two pairs of chromosomes.

Large insert size paired reads (5 and 8kb) from "Calcutta4" were mapped on the reference sequence to confirm the presence of these translocations and precisely locate the translocation breakpoints. Analysis of discordant read mapping suggested a first reciprocal translocation involving a 240Kb distal region of acrocentric chromosome 2 and a 7.2 Mb distal region of chromosome 8. A second reciprocal translocation involves a 20.8Mb distal region of acrocentric chromosome 1 and a 11.6Mb distal region of chromosome 9, with intricate events of duplication and deletion at the breakpoint.

We are currently using BAC-FISH to validate these structural variations.

Perspectives are to develop PCR markers at the breakpoint to analyze the presence of these rearrangements in Musa germplasm and to analyze of the impact of these rearrangements in heterozygous accessions.

Large chromosomal structural variations between *M. acuminata* and *M balbisiana* and their consequences on chromosome recombination and segregation in a AAAB polyploidy context

BAURENS Franc-Christophe[1], MARTIN Guillaume[1], ROUARD Matthieu[2], SALMON Frédéric [3], NJEMBELE Célestin[4], HABAS Remy[5], RICCI Sébastien[3,4], and D'HONT Angélique [1]

[1]: *CIRAD, UMR AGAP, F-34398 Montpellier, France*
[2]: *Bioversity International, Montpellier France*
[3]. *CIRAD, UMR AGAP, F-97130 Capesterre Belle Eau, Guadeloupe, France*
[4]. *CARBAP, Rue Dinde, N°110, Bonanjo, BP 832 Douala – Cameroun*
[5]: *CIRAD, UMR BGPI, F-34398 Montpellier, France*

Many banana cultivars are triploid interspecific hybrids between M. acuminata (Genome A, 2n=22) and M. balbisiana (Genome B, 2n=22). They included the important group of Plantain cooking bananas classified as AAB.

We compared the global chromosome structure of A and B genomes through the construction of a *M. balbisiana* genetic map and its comparison with the *M. acuminata* reference sequence assembly. We identified a large reciprocal translocation involving chromosome 1 and 3 and a large inversion on chromosome 5.

We analyzed the A/B chromosomes composition of plantain cultivars revealing a few chromosome segments with AAA composition and one entire chromosome with ABB composition instead of the supposed general 'AAB' composition.

We analyzed the A/B chromosomes composition of a progeny from a 'AAAB' tetraploid plantain derived breeding accession. We observed recombination between A and B genomes along most of the chromosomes. A few exceptions were observed with no recombination in the inverted segment between A and B on chromosome 5 and a reduced recombination near the translocated breakpoints on chromosome 1 and 3. We also observed a very important number of aneuploids in the progeny (62%) with 6% involving chromosome 5 and 46% involving chromosome 1 and 3, the three chromosomes that differed in their global structure between A and B genomes.

Implication of these results on the origin of banana cultivars and on breeding of banana allopolyploids will be discussed.

New insights on homologous recombination in polyploids: the striking case of *Brassica* allotriploids

Alexandre Pelé[1], Matthieu Falque[2], Gwenn Trotoux[1], Frédérique Eber[1], Sylvie Nègre[1], Marie Gilet[1], Virginie Huteau[1], Maryse Lodé[1], Thibaut Jousseaume[1], Jérôme Morice[1], Olivier Coriton[1], Olivier C. Martin[2], Mathieu Rousseau-Gueutin[1] and Anne-Marie Chèvre[1]

1. IGEPP, INRA, Agrocampus Ouest, Université de Rennes 1, 35653 Le Rheu, France
2. GQE-Le Moulon, INRA, Université Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay, 91190, Gif sur Yvette, France

Meiotic recombination by crossovers (COs) is tightly regulated, limiting its key role in producing genetic diversity. While one obligate CO occurs per pair of homologs, ensuring their proper segregation during meiosis, rarely more than three are formed and their distribution is not homogenous along chromosomes. In plants, Whole Genome Duplication (WGD) was highlighted to result in a boost of the number of COs between homologous chromosomes, which may contribute to the success of allopolyploid species by generating enhanced allelic combinations. However, the consequences on COs distribution as well as the regulation of this phenomenon are poorly understood. We investigated here the striking case of *Brassica* allotriploids (AAC, 2*n*=3x=29), resulting from crosses between *B. napus* (AACC, 2*n*=4x=38) and its *B. rapa* progenitor (AA, 2*n*=2x=20), showing far higher CO rates between A homologs than diploids and allotetraploids due to specific additional C chromosomes. From several populations developed, we assessed the homologous recombination in AA diploids and hybrids carrying either an additional complete C genome (9 C chromosomes) or specific C chromosomes, through the genotyping of 204 SNP markers well distributed along A chromosomes (one SNP each 1.2 Mb). Compared to what was previously known, we showed that the presence of the C genome in AAC allotriploids leads to a very substantial increase of COs all along the A chromosomes, especially in the vicinity of centromeres that are normally deprived of COs. We also demonstrated that the addition of a C09 chromosome originating either from *B. oleracea* (CC, 2*n*=2x=18) or *B. napus* have contrasted effects on recombination. This latter result points out that regulation of homologous recombination in AAC allotriploids may have changed in the allopolyploid *B. napus*, as a result of the 7,500 years of genome coevolution in a polyploid context. Together, our findings provide new insights on homologous recombination in allopolyploids as well as the opportunity to break the linkage disequilibrium in the rapeseed breading programs by using allotriploids.

SESSION 4

(VENDREDI 9 JUIN 2017)

Identification et analyse des variations structurales du génome chez le blé

Romain De Oliveira, Etienne Paux, François Balfourier, Frédéric Choulet

GDEC, INRA, UCA, UMR1095 Clermont-Ferrand

Le blé tendre est une espèce d'intérêt agronomique majeur au génome complexe : hexaploïde, composé de 16 Gb dont 85-90% d'éléments répétés. L'ampleur de la variabilité génomique chez cette espèce et, plus largement, au sein des *Triticeae*, reste encore mal connue du fait de cette complexité. Nous cherchons à mieux caractériser la diversité structurale et à mieux comprendre la dynamique évolutive du génome. Les travaux que nous avons entrepris visent à mettre en évidence les variations structurales (CNV et PAV, variations du nombre de copies de gènes et présence/absence) à partir de données de re-séquençage sur différents matériels : le génome entier (16 variétés), des échantillons d'ADN issus de capture d'exome (112 variétés), des chromosomes triés pour 44 variétés cultivées et sauvages. Nous avons développé des stratégies d'analyse pour identifier les variations structurales affectant les gènes mais aussi les éléments transposables. Les premiers résultats montrent un taux accru de polymorphisme dans les extrémités des chromosomes, confirmant les hypothèses émises précédemment sur la dynamique évolutive. Les méthodologies mises en place et les premiers résultats seront présentés et discutés.

Overview of actual and ancestral recombination in relationship with the sequence in wheat: focus on chromosome 3B

Darrier Benoît[1,2], Loussert Alain[1,2], Rimbert Hélène[1,2], Navarro Julien[1,2], Choulet Frédéric[1,2], Philippe Leroy[1,2], Lise Pingault[1,2], Aurelie Evrard[1,2], IWGSC[3], Balfourier François[1,2], **Sourdille Pierre**[1,2]

[1]*INRA UMR1095 Genetics, Diversity and Ecophysiology of Cereals, 5, Chemin de Beaulieu, 63039 Clermont-Ferrand, France*
[2]*University Blaise Pascal UMR1095 Genetics, Diversity and Ecophysiology of Cereals, 5, Chemin de Beaulieu, 63000 Clermont-Ferrand, France*
[3]*IWGSC, http://www.wheatgenome.org/*
pierre.sourdille@inra.fr

Meiotic recombination (crossovers) is a mechanism that largely contributes to shape the genome structure in all species. However, in bread wheat (Triticum aestivum L.), recombination occurs almost exclusively in distal telomeric regions. Thus, many genes contained in pericentromeric regions which represent 80% of the chromosomes are not admixed during meiosis. We exploited the chromosome-survey data from IWGSC to study at the genome scale 596 CO events located on 445 contigs. Among these, more than 250 CO mapped on 25 scaffolds of the 3B pseudomolecule on features smaller than 50kb. This number was increased to reach ~500 COs on 3B which were correlated to sequence features (epigenetic landmarks, genes, transposable elements, specific motifs, RNAseq data of meiosis) available for chromosome 3B. We thus showed that COs mainly occur in the vicinity of genes. Moreover, we used two col
lections of 90 lines corresponding to Asian and European genetic pools to study the variation of linkage disequilibrium (LD) as well as the ancestral recombination at a finer scale in the regions covering our previously detected hot spots. Analysis revealed high difference in LD structure between the two populations underlying the complex history of each genetic pool. However historical mapping of recombination events using the same SNPs showed a common location of recombination breakpoints with variable intensity. Impact of these results on the improvement of recombination in wheat will be discussed.

Evolution of African rices pangenomes through domestication and post-domestication

SABOT François

1.UMR DIADE, Institut de Recherche pour le Développement, Montpellier, France

Cultivated African rice O. glaberrima and its wild ancestor O. barthii were investigated for the structure of their pangenomes.

For that we developped a new method to detect presence/abscence of gene based on Deth of Coverage in population data.

Using massive Illumina sequencing upon 162 and 86 samples respectively, we were able to identify major variations in their pangenomic structure.

Indeed, a massive loss of genes was observed in the cultivated species, especially in the core genome. Thus, two cultivated individuals are less similar in terms of pangenome than two wild ones.

More precisely, when focusing on GO and gene families, we identified massive loss in stress responses, signaling and cell wall development.

Crops, man and TEs. The 3000 rice genomes unravel TE-driven genome dynamics in the field.

Olivier PANAUD, Laboratoire Génome et Développement des Plantes, UMR Université de Perpignan/CNRS 5096. Institut Universitaire de France.

Plants have in common with other eukaryotes to harbour a large amount of transposable elements in their genome. These mobile entities have been shown to contribute to their biological diversity. First considered as mutagenic agents that could impede the function of genes upon their insertion into coding sequences and then as selfish replicators (the majority of TE insertions being neutral, without any effect on the phenotypes), TEs have been indeed shown to be involved in the emergence of biological novelty, mainly through the modification of the regulation of genes and/or gene networks. In this regard, one of the main challenges in plant genomics is to understand the relationships between the structure and the function of genomes, especially in the context of transposition. One possible strategy for such study is to conduct large scale comparative genomic surveys when sufficient data is available. The recent release of sequence data for 3000 rice varieties provides a unique opportunity to characterize the transpositonal landscape at the whole species level. We developed a new bioinformatic pipeline to perform this task. In this presentaton, we will provide new insights about TE-driven rice genome dynamics, on the domestication of the crop and on the possible functional impact of transposition in the species.

Distribution of *Divo* in *Coffea* genomes, a poorly described family of angiosperm LTR-Retrotransposons

Dupeyron M, Fernandes de Souza R, Hamon P, de Kochko A, Crouzillat D, Couturon E, Domingues DS, Guyot R.

Institut de Recherche pour le Développement (IRD), UMR DIADE, Montpellier, France

mathilde.dupeyron@ird.fr

LTR-retrotransposons (LTR-RTs) are omnipresent components of plant genomes. Numerous lineages and families have been described, leading to a well-established classification of elements[1]. With the availability of bioinformatics tools dedicated to the LTR-RTs identification and analysis, it became reasonable to perform annotation of such transposable elements at whole-genome scale.

During the process of *Coffea* genomes annotation[2,3], we characterized and studied a new *Copia* LTR-RT family in diploid and allotetraploid *Coffea* genomes called *Divo*. It is a complete and relatively compact LTR-RT element (~5 kb), carrying typical Gag and Pol *Copia* type domains. Reverse Trancriptase (RT) domain-based phylogeny demonstrated that *Divo* is a new and well-supported family in the *Bianca* lineage, but strictly restricted to dicotyledonous species. In *C. canephora*, *Divo* showed a genomic distribution along gene rich and gene poor regions. In coffee trees, as well as in Arabidopsis and grapevine, *Divo* was present in relatively low copy numbers. In *Coffea* genomes, the presence of recently inserted and complete copies and the detection of RNAseq transcription suggest that *Divo* might be active.

Coffea arabica (the Arabica coffee) is an allotetraploid species originating from a recent hybridization between two diploid species: *C. canephora* and *C. eugenioides*. The copy number, the molecular estimation of insertion time and the analysis at orthologous locations of insertions in these three species suggest that *Divo* underwent a different and recent transposition activity in *C. arabica* and *C. canephora* when compared to *C. eugenioides*. The analysis of this novel LTR-RT family represents an important step toward uncovering the genome structure and evolution of *C. arabica* allotetraploid genome.

1 - Llorens C. et al. (2009) Network dynamics of eukaryotic LTR retroelements beyond phylogenetic trees. Biol Direct **4**:41

2 - Denœud F. et al. (2014) The coffee genome provides insight into the convergent evolution of caffein biosynthesis. *Science* **345**(6201):1181-1184.

3 - Arabica Coffee Genome Consortium (ACGC)

Assembling a large plant Y chromosome using long read sequencing

Cécile Fruchard[1], Nelly Burlet[1], Radim Čegan[2], Mohamed Amine Chebbi[3], Richard Cordaux[3], Roman Hobza[24], Gabriel Marais[1]

[1] Université de Lyon, F-69000, Lyon ; Université Lyon 1 ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, France.

[2] Department of Plant Developmental Genetics ; Institute of Biophysics of the CAS, v.v.i. ; Kralovopolska street 135, 612 00 Brno, Czech Republic

[3] Université de Poitiers ; UMR CNRS 7267 Ecologie et Biologie des Interactions ; Equipe Ecologie Evolution Symbiose ; 86073 Poitiers, France

[4] Institute of Experimental Botany ; Laboratory of Molecular Cytogenetics and Cytometry ; Centre of the Region Hana for Biotechnical and Agricultural Research, Sokolovska 6, CZ-77200, Olomouc, Czech Republic.

Cecile.fruchard@univ-lyon1.fr

Abstract:

Although rarer than in animals, separate sexes (dioecy) have evolved in ~15,600 angiosperm species (~6% of all species, ~20% of crops). Only two master sex-determining genes have been identified in dioecious plants so far, namely in persimmons and asparagus. In the vast majority of dioecious species, those genes are unknown, even in the well-studied papaya and Silene latifolia. Sequencing Y chromosomes is a necessary step towards identifying sex-determining genes. However, this sequencing of Y chromosomes remains one of the greatest challenges of current genomics. In sharp contrast with the >300 fully sequenced eukaryotic genomes, only a handful of Y chromosomes (mostly animal ones) have been sequenced to date. The non-recombining Y chromosome tends to accumulate repeats (transposable elements and amplicons) which makes virtually impossible the assembly using only short-read sequencing technologies. We aim at sequencing the Y chromosome of S. latifolia, a well-studied dioecious plant, in which three large sex-determining regions have been described but which represents a real challenge as the Y, in this species, is 550 Mb-long and probably comprises a very large fraction of accumulated repeats.

We obtained ~95% pure S. latifolia Y DNA using flow cytometry. Sequencing of this chromosome relies on combining both short-read Illumina paired-end sequencing and PacBio plus MinION, two 3rd generation sequencing technologies providing long (>1Kb) to extra-long (>50Kb) reads to improve assembly. We present our attempts at assembling and identifying a catalogue of genes present on the S. latifolia Y chromosome.

Further scaffolding will be performed using different independent data sources (RNA-seq, Y-linked BACs, physical map of the Y). Annotation will be done with the help of a reference transcriptome that we have assembled previously and in which a number of Y-linked genes have already been identified. Detailed analysis of the three sex-determining regions will help short-listing master sex-determining gene candidates, which we will test using different approaches and resources including a Y-mutant collection.