First workshop on Plant Genome Dynamics and Evolution

INRA CNRGV Campus d'Auzeville TOULOUSE

Le 28 et 29 Mai 2015

More information at https://colloque.inra.fr/worshop-series-dynagev

Jeudi 28 Mai 2015

8h30-8h45: Welcome of the participants / Accueil des participants

8h45-9h00: Introduction - practical information / détails pratiques

SESSION 1- GENOME PROJECTS (ANIMATRICE A. D'HONT)

9h00-9h30 Update on geneome sequencing of the allotetraploid *Coffea arabica* / Dernière avancée dans le séquençage du génome de l'allotetraploïde *Coffea arabica* (**Alexandre de KOCHKO - IRD Montpellier**)

9h30-10h00: A reference genome for Sugarcane / Une séquence de référence pour le génome de la canne à sucre (**Olivier GARSMEUR - CIRAD Montpellier**)

10h00-10h30: Oilseed rape (*Brassica napus*) allotetraploid genome sequenced / Séquençage du Génome du Colza, *Brassica napus* (**Boulos CHALHOUB - INRA Evry**)

10h30-11h00 Coffee Break

SESSION 2 - STRUCTURAL VARIATION (ANIMATRICE H. BERGES)

11h00-11h30: Structural variations of the hexaploid wheat chromosome 3B / Variations structurales sur le chromosome 3B du blé hexaploïde (**Etienne PAUX - INRA Clermont-Ferrand**)

11h30-12h00: Lessons from the genome of European maize line FV2 : structural and functionnal specificities/ Le génome de la lignée de maïs européenne FV2: spécificités structurales et fonctionnelles (**Johann JOETS - INRA Le Moulon**)

12h00-12h30: Structural evolutionary dynamics in synthetic oilseed rape / Dynamique structurale et évolutive chez des colzas synthétiques (**Mathieu ROUSSEAU-GUEUTIN - INRA Rennes**)

12h30-13h00: Do monkeys play with banana genomes? / Des singes ont-ils joué avec les génomes de bananiers ? (**Guillaume MARTIN - CIRAD Montpellier)**

13h00-14h30: Lunch at INRA

SESSION 3 - PLANT GENOME EVOLUTION (ANIMATEUR O.CORITON)

14h30-15h00: Design of high-throughput genotyping arrays to analyze structural variants in maize: first results / Premiers résultats du génotypage haut-débit de variants structuraux chez le mais (**Stéphane NICOLAS - INRA Le Moulon**)

15h00-15h30: Origin and structure of Citrus genomes / Origines et structure du génome des *Citrus* (**François LURO - INRA Corse)**

More information at https://colloque.inra.fr/worshop-series-dynagev

15h30-16h00: Evolutionary dynamics of the chloroplast genomes in the legume genus Lupinus / Dynamique évolutive des génomes chloroplastiques dans le genre Lupinus (leguminosae) (**Jean KELLER - Univ-Rennes**)

16h00-16h30: Coffea break and poster

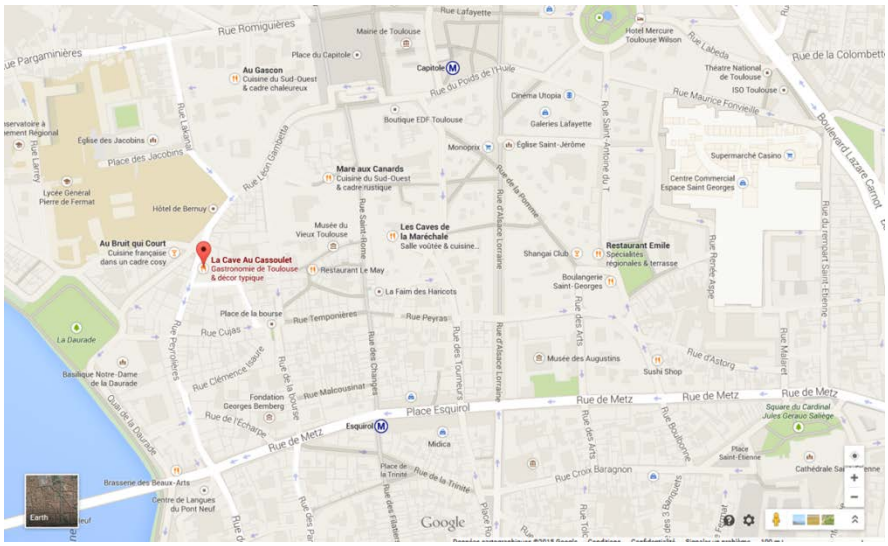
16h30-17h00: Extracting the A genome from Oilseed rape cv Darmor nain / Extraction Genome A de du Colza, Darmor Nain (**Anne-Marie CHEVRE/Gwenn TROTOUX - INRA Rennes**)

17h00-17h30: Fusion sites in wheat chromosomes / Etude des sites de fusion de chromosomes chez le blé (**Arnaud BELLEC - INRA Toulouse**)

17h30-18h00: Peculiar features of common bean subtelomeres / Spécificités des régions subtélomériques de Haricot (**Valérie GEFFROY – INRA Versailles**)

18h00 – 18h30: Round table- Exchanges / Table Ronde Echange

20h RESTAURANT in « La Cave au Cassoulet » 54 Rue Peyrolières **TOULOUSE**



More information at https://colloque.inra.fr/worshop-series-dynagev

Vendredi 29 Mai 2015

SESSION 4- MEIOTIC RECOMBINATION (ANIMATEUR E.JENCZEWSKI)

8h45-9h15: A meiotic transcriptome overview in a polyploïd species (*Brassica napus*) / Le transcriptome méiotique d'une espèce polyploïde, le colza (**Aurélien BLARY – INRA Versailles**)

9h15-9h45: Fine Mapping of recombination in BreedWheat: Example of chromosome 3B / Cartographie fine des sites de recombinaison chez le blé: l'exemple du chromosome 3B (**Benoit DARRIER – INRA Clermont Ferrand**)

SESSION 5- GENE AND GENOME DUPLICATION (ANIMATRICE A-M. CHEVRE)

9h45-10h15: Tracking duplicate copies in high-redundant *Spartina* genomes / A la recherche des copies dupliquées dans le génome hautement polyploïde des Spartines (**Julien BOUTTE - Univ-Rennes**)

10h15-10h45: Evolution of duplicated pathways in allopolyploid cotton / Evolution de voies métaboliques dupliquées chez le coton allotétraploïde (**Joseph GALLAGHER - Iowa State University**)

10h45-11h00: Coffee break and Poster

11h00-11h30: Importance of transposon-based genomic changes in natural and synthetic *Nicotiana* allopolyploids / Rôle des changements génomiques associés aux éléments transposables chez des formes allotétraploïdes naturelles et synthétiques de Nicotiana (**Corinne MHIRI - INRA Versailles)**

11h30-12h00: Persistence, dispersal and genetic evolution of recently formed Spartina homoploid hybrids and allopolyploids in Southern England / Persistence, dispersion et évolution génétique d'hybrides homoploïdes récents de Spartines dans le Sud de l'Angleterre (**Ales KOVARIK - Academy of Sciences of the Czech Republic**)

12h00-12h30: *Brassica* allohexaploids: problems, processes and potential / Potentialités et problèmes associés à la formation d'allohexaploïdes de Brassica (**Annaliese MASON - Justus Liebig University Giessen**)

12h30 – 13h00: Round table- Exchanges – End of the conférence / Table Ronde Echange Clôture de la conférence

13h00: Lunch at INRA

RESUMES/ABSTRACTS

SESSION 1- GENOME PROJECTS

Update on geneome sequencing of the allotetraploid *Coffea arabica* / Dernière avancée dans le séquençage du génome de l'allotetraploïde *Coffea arabica* (**Alexandre de KOCHKO - IRD Montpellier**)

COFFEA ARABICA GENOME SEQUENCE

Dominique Crouzillat[1], Michel Rigoreau[1], Maud Lepelley[1], Laurence Bellanger[1], Virginie Mérot-l'Anthoëne[1], Céline Vandecasteele[1], **Alexandre de Kochko[2]**, Romain Guyot[2], Valérie Poncet[2], Christine Tranchant[2], Perla Hamon[2], Serge Hamon[2], Emmanuel Couturon[2], Patrick Descombes[3], Déborah Moine[3], Lukas Müeller[4], Suzy Strickler[4], Alan Andrade[5], Luiz Filipe Pereira[5], Pierre Marraccini[6], Giovani Giuliano[7], Alessia Fiore[7], Marco Pietrella[7], Giuseppe Aprea[7], Ray Ming[8], Jennifer Wai[8], Douglas Silva Domingues[9], Alexandre Paschoal[10], Gerrit Kuhn[11], Jonas Korlach[11], Jason Chin[11], David Sankoff[12], Chunfang Zheng[12], Victor Albert[13]

1: Nestlé R&D Tours France, 2: IRD, Montpellier France, 3: NIHS, Lausanne Switzerland, 4: BTI Cornell USA, 5: EMBRAPA Brazil, 6: CIRAD France, 7: ENEA Roma Itlaie, 8: University of Illinois Urbana-Champaign USA, 9: IAPAR Londrina Brazil, 10: University of Londrina Brazil, 11: Pacific Biosciences USA, 12: University of Ottawa Canada, 13: University of Buffalo USA

Coffea arabica which accounts for 70% of world coffee production is an allotetraploid with a genome size of approximately 1.3 Gb and is derived from the hybridization of *C. canephora* (710 Mb) and *C. eugenioides* (670 Mb). To elucidate the evolutionary history of *C. arabica*, and generate critical information for breeding programs, a sequencing project is underway to finalize a reference genome using a dihaploid line and a set of 30 *C. arabica* accessions. For the reference genome, we have generated two assemblies, one from Illumina data (>150x coverage) and a second from PacBio sequences (>50x coverage). The present assemblies cover 1,031 and 1,042 Mb, respectively. After further refinement, using optical mapping, the genome assemblies will be annotated using IsoSeq. Resequencing of *C. eugenioides* and *C. canephora* using PacBio technology is also underway and will used to better assess homeologs within the sub-genomes. Furthermore, 30 *C. arabica* accessions, representing wild and cultivated genotypes, are being resequenced (20x coverage) using Illumina. A *C. arabica* genetic map, currently including over 650 SSR markers, that differentiate between the two sub-genomes, is used to anchor the assemblies. Newly identified SNP markers are being added to the map.

The final goals of the project are to produce a high quality reference genome, assess an eventual neo-diversification occurring in the cultivated varieties, have a better understanding of the species formation and evolution, and develop tools that will make the finished genome accessible and useful to breeders and researchers.

A reference genome for Sugarcane / Une séquence de référence pour le génome de la canne à sucre
(**Olivier GARSMEUR - CIRAD Montpellier**)

TOWARD A REFERENCE SEQUENCE OF THE GENE RICH PART OF THE SUGARCANE GENOME

Olivier Garsmeur[1], Bernard Potier[2], Karen Aitken[3], Paul Berkman[3], Gaetan Droc[1], Carine Charron[1], Guillaume Martin[1], Bandie Harrison[4], Edwin van der Vossen[5], Robert Henry[6] and Angélique D'Hont[1].

[1] CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), UMR AGAP, F-34398 Montpellier, France.
[2] SASRI (South African Sugarcane Research Institute), Mount Edgecombe, 4300, South Africa.
[3] CSIRO (Commonwealth Scientific and Industrial Research Organisation), Agriculture, St Lucia, QLD 4067, Australia.
[4] Amplicon Express Inc., 2345 NE Hopkins Ct., Pullman, WA 99163, USA.
[5] Keygene N.V., Agro Business Park 90, 6708 PW, Wageningen, the Netherlands.
[6] QAAFI (Queensland Alliance for Agriculture and Food Innovation), University of Queensland, St Lucia, QLD 4072, Australia.

Keywords: Sequencing, Genome, Sugarcane, Genes.

The sugarcane genome poses challenges that have not been addressed in any prior genome sequencing project. The main difficulties reside in its high polyploidy (2n ~ 12x ~ 120), and its high level of heterozygosity which makes an assembly of the genome very challenging through classical whole genome shotgun sequencing approaches.

Previous studies demonstrated that sugarcane hom(e)ologous chromosomes share a very high level of micro-colinearity among themselves and show good micro-colinearity with sorghum. These findings suggested that sorghum could provide a good template to select and sequence a minimum tiling path (MTP) of sugarcane BACs representing the gene-rich part of a monoploid genome composed of 10 basic chromosomes. We exploited the Whole Genome Profiling (WGP™) technology of Keygene to analyze a set of 20,736 BACs from cultivar R570, representing an approximate 2 fold coverage of the monoploid genome of sugarcane. The WGP technology generates short sequence tags from the terminal ends of restriction fragments from pooled BACs. An average of 37.2 sequence tags per BAC was generated that allowed anchoring more than 11,000 of the profiled R570 BACs on the sorghum sequence. A minimum tiling path (MTP) of 5,000 BACs has been selected and is currently being sequenced, through international collaboration, to generate a reference sequence of the gene-rich part of the sugarcane genome. So far, around 2,500 BACs corresponding to 5 out of the 10 sugarcane basic chromosomes have been sequenced. We are currently testing the possibility to further anchor the sugarcane sequence to R570 chromosomes through genetic mapping. We have tested the genotyping by sequencing (GBS) on a sugarcane mapping population, developed bioinformatics tools to identify simplex SNP markers and built a R570 genetic map comprising more than 7,000 SNP markers. We acknowledge members of the International Consortium for Sugarcane Biotechnology (ICSB) for their support.

SESSION 2 - STRUCTURAL VARIATION

Structural variations of the hexaploid wheat chromosome 3B / Variations structurales sur le chromosome 3B du blé hexaploïde (**Etienne PAUX - INRA Clermont-Ferrand**)

Variations structurales du chromosome 3B de blé tendre

Emeric Dynomant, Hélène Rimbert, Frédéric Choulet, Etienne Paux*

Genetics, Diversity & Ecophysiology of Cereals
INRA Auvergne-Rhône-Alpes 5 chemin de Beaulieu
63039 Clermont-Ferrand cedex 2FRANCE
Contact : etienne.paux@clermont.inra.fr

Sous le terme variations structurales (SV) sont rassemblées un ensemble de réarrangements tels que des délétions, duplications, inversions et translocations impactant la structure du génome d'une espèce et conduisant à des différences entre individus. Souvent associées aux gènes, où elles sont définies en tant que variations du nombre de copies (CNV) ou variations de type présence / absence (PAV), les SV peuvent également concerner les éléments transposables (TE). Chez le blé hexaploïde, les SV peuvent apparaître à trois niveaux : (1) entre génomes homéologues ; (2) entre variétés différentes ; (3) à différents niveaux de ploïdie. Toutefois, sans séquence de référence disponible chez cette espèce à génome complexe, les connaissances sur ce sujet restent très limitées. Notre équipe a produit récemment la première séquence de référence d'un chromosome de blé, le 3B, qui constitue une ressource clé pour la découverte et la caractérisation des variations structurales chez cette espèce d'intérêt agronomique majeur. Pour cela, nous avons également reséquencé les chromosomes 3B de 44 lignées de blé hexaploïdes et tétraploïdes. Une analyse du paysage transpositionnel et des variations du contenu en gènes de ces accessions a été réalisée. Les résultats de cette étude seront présentés et mis en regard des caractéristiques structurales et fonctionnelles particulières déjà mises en évidence sur le chromosome 3B.

Structural evolutionary dynamics in synthetic oilseed rape / Dynamique structurale et évolutive chez des colzas synthétiques (**Mathieu ROUSSEAU-GUEUTIN - INRA Rennes**)

Structural evolutionary dynamics in synthetic oilseed rape

M. Rousseau-Gueutin[1], J. Morice[1], S. Negre[1], G. Trotoux[1], C. Falentin[1], O. Coriton[1], F. Kerbrat[1], G. Deniot[1], M. Gilet[1], M. Lode[1], A. Pele[1], F. Eber[1], H. Berges[2] et AM. Chevre[1].

[1] INRA, UMR 1349, Institut de Génétique, Environnement et Protection des Plantes, 35653 Le Rheu Cedex, France.

[2] INRA – Centre National de Ressources Génomiques Végétales, 31326 Castanet Tolosan Cedex, France.

Allopolyploidy is at the origin of an evolutionary dynamic that is critical for the adaptive success of newly formed lineages. Presently, one of the best biological system to study this phenomenon is *Brassica napus* (oilseed rape) since its whole genome sequence and those of its parental diploid species (*B. rapa* and *B. oleracea*) have been sequenced recently. These data, in addition with the availability of synthetic *B. napus* individuals (produced in our laboratory) allows to decipher, at a level never attained before, the evolutionary dynamics of *B. napus* subgenomes at the structural and functional levels.

A Illumina SNP60k array was used to investigate the structural modifications (deletion, Non Reciprocal Translocation and genic conversion) that may occur after the formation of a allopolyploid *B. napus* (3rd generation of outcrossing). By comparing the genome of 15 synthetic *B. napus* (in 3rd generation) with its diploid progenitors, we were able to determine that some chromosomes were more prone to rearrangements and that a subgenome was supra-dominant (i.e. pivotal). The various genetic modifications observed were then verified experimentally using BAC-FISH (deletion and NRT) or by direct sequencing (genic conversion). From this study, we notably observed that more than 5% percent of *B. napus* genome may be shuffled and that many thousands of one homoeologous gene copies may be lost in only three generations.

Do monkeys play with banana genomes? / Des singes ont-ils joué avec les génomes de bananiers ?
(**Guillaume MARTIN - CIRAD Montpellier)**

Do monkeys play with banana genomes?

Guillaume Martin[1], Franc-Christophe Baurens[1], Paco Derouault[1], Gaëtan Droc[1], Mathieu Rouard[2], Angélique D'Hont[1].

1. CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), UMR AGAP, F-34398 Montpellier, France

2. Bioversity International, Parc Scientifique Agropolis II, 34397 Montpellier Cedex 5, France.

Banana cultivars are derived from hybridization between *Musa acuminata* subspecies (A genome) and for some with the species *M. balbisiana* (B genome). These hybrids have reduced fertility, disturbed meiosis and strong segregation distortions. These characteristics, attributed to chromosomal rearrangements between species and subspecies, complicate genetic analysis and breeding programs. In this context it is important to characterize these chromosomal structural variations affecting bananas. We have developed and tested new approaches, based on the recent availability of a reference sequence of the banana and high - throughput sequencing technologies to characterize these chromosomal structural variations. The tools we developed were tested on a few banana accessions and showed promising results.

SESSION 3 - PLANT GENOME EVOLUTION

Design of high-throughput genotyping arrays to analyze structural variants in maize: first results / Premiers résultats du génotypage haut-débit de variants structuraux chez le mais (**Stéphane NICOLAS - INRA Le Moulon**)

A new high-throughput approach to characterize Presence/Absence Variant in Maize based on Affymetrix technology

Stéphane Nicolas[1], Aude Darracq[1], Ali Pirani[3], Hélène Rimbert[2], Clément Mabire[1], Delphine Madur[1], Nathalie Rivière[2], Jean-Philippe Pichon[2], Johann Joets[1], Jorge Duarte[2]

1 Génétique Quantitative et Evolution - Le Moulon, INRA - Université Paris-Sud - CNRS - AgroParisTech, Ferme du Moulon, F-91190, Gif-sur-Yvette, France

2 Biogemma - Upstream Genomics Group, route d'Ennezat, CS 90126, Chappes 63720, France

3 Affymetrix - 3420 Central Expy, Santa Clara, CA, 95051, USA

Structural variations are pervasive in plants, notably in maize with large insertions/deletions (>1kbp) harbouring hundreds of expressed genes. Although these polymorphisms, also called Presence/Absence Variants (PAVs), could play a key role in agronomic trait variations, heterosis and plant evolution, their distribution and frequency along the genome within the maize germplasm remains poorly characterized. On the contrary to SNPs, discovery and high-throughput genotyping of PAVs remains challenging and has not been achieved yet. We developed an original high-throughput approach based on Affymetrix Axiom technology to genotype 108,240 PAVs ranging from 115pb to 130kbp with a combination of probes designed on (i) PAV breakpoints (ii) PAV internal sequences and (iii) SNPs within PAV sequences. We used a dedicated bioinformatics pipeline to discover these insertions/deletions regarding the B73 reference genome from resequencing data of 3 key founders of breeding material: an european flint line F2, an american iodent line PH207 and a lancaster dent line C103. Twenty-five million probes were designed using Affymetrix algorithms, from which we selected 662,772 probes (108K PAVs), according to their type and distribution along the PAV sequences. 480 maize inbred lines representing worldwide diversity were genotyped using this new array. Affymetrix developed dedicated pipelines corresponding to the 3 probes types to call PAVs and classify probes. Preliminary results show that of the interrogated probes, 64% (breakpoints probes), 81% (monomorphic internal PAV probes) and 61% (polymorphic internal PAV probes) could be called. We will present our results and discuss our ability to genotype PAVs regarding their origin, the probe type and Affymetrix scoring.

Keywords: Copy Number Variant, Genotyping, Maize, Present/Absent Variant, Breakpoint, Diversity, Axiom Affymetrix

Comparative analysis of five chloroplast genomes in the legume genus *Lupinus*

Keller J[1], Rousseau-Gueutin M[13], Martin GE[2], Cabello-Hurtado F[1], Morice J[3], Boutte J[1], Coissac E[4], Aïnouche M[1], Salmon A[1], Aïnouche A[1]

[1]UMR CNRS 6553 Ecobio, OSUR (Observatoire des Sciences de l'Univers de Rennes), Université de Rennes 1; [2]CIRAD (Centre de coopération Internationale en Recherche Agronomique pour le Développement), UMR AGAP, F-34398 Montpellier, France ; [3]Institut national de la recherche agronomique, Le Rheu, 35653 Le Rheu CEDEX, France ; Laboratoire d'Ecologie Alpine, CNRS - Université Grenoble 1 - Université de Savoie, France.

State of the art and aim: The chloroplast genome plays an essential role in plants and is involved in major processes such as photosynthesis, and other gene expression and metabolic machineries. Due to the slow evolutionary rate of the plastome and to its maternal inheritance, chloroplast sequence data are generally useful for phylogeny, detection of reticulate evolution, phylogeography, molecular dating or DNA barcoding studies. With the recent progress of Next Generation Sequencing technologies (NGS), the number of sequenced plastomes has rapidly expanded in the last decade. Until recently, only few chloroplast genomes were available for the legume family, with all of them only belonging to one Papilionoid lineage (the NPAAA clade including species such as *Medicago*, *Glycine*…), leaving a large gap for the other lineages (Cardoso *et al*. 2012; Magee *et al*. 2010).

Results and discussion: Last year the first sequenced chloroplast genome of Genistoid legume revealed new specific rearrangement shared by lupines and Genistoids. This consists in a 36-kb inversion included in an already known inversion of 50-kb (Martin, Rousseau-Gueutin *et al.*, 2014). Moreover this study shows a loss of the *rpl22* gene, which has been successfully transferred to the nucleus.

Following this work, chloroplast genomes of four lupines have been sequenced and annotated (151,808 to 152,272 bp, 111 genes). Comparison of the 5 known lupine chloroplast genomes revealed a high overall stability (98 % identity). Analysis of gene content, evolutionary rate of coding and non-coding sequences, as well as analyses of repeats, resulted in the identification of both highly conserved and stable regions on one side, and remarkably variable regions representing hotspots on the other side. For example, two of the five lupines studied here have more or less completely lost the *rps16* gene. Additionally, these regions exhibit numerous variable characters which represent a novel source of potentially informative characters to be used in molecular evolutionary studies at various taxonomic levels.

Key words: Lupinus, legumes, chloroplast genome evolution, lineage-specific 36-kb inversion, plastome hotspots, variable sequences, phylogenetic utility

References:

Cardoso D *et al*. (2012) *American Journal of Botany*

Magee AM *et al*. (2010) *Genome Research*, 20, 1700

Martin GE *et al.* (2014) *Annals of Botany*,

Fusion sites in wheat chromosomes / Etude des sites de fusion de chromosomes chez le blé (**Arnaud BELLEC - INRA Toulouse**)

Exploring structure of Chromosome Fusion sites in hexaploid wheat

Arnaud BELLEC – CNRGV -INRA Toulous

The explosion of genomes sequencing in the past decade has highlighted whole genome duplications (WGDs) as a major evolutionary driving force shared by multiples branches of the tree of life. Thus WGDs have shaped ancestral vertebrates genomes and have occurred frequently in plants. However, the genome sizes and chromosome numbers have not increased exponentially over time because WGDs have been balanced by chromosome number reduction (CNR) and gene losses. These two mechanisms have been found through genomes sequences comparisons and ancestral karyotype reconstruction in plants (Salse, Curr. Plant Biol., 2012). Models have been proposed explaining CNR by chromosomes fusions where two Chromosomes can merge by their extremities or by centromeric/telomeric interaction leading to nested chromosome fusions (NCFs). However, molecular mechanisms driving ancestral chromosome fusions that have led to the present day diversity of plant are still largely unknown.

Following this evolutionary mechanisms modern Triticeae (wheat, barley) with a basic number of 7 chromosomal groups have been shaped from a monocot ancestral genome structured in 12 protochromosomes following 5 NCFs (Murat et al. Genome Biol. Evol., 2014). Hexaploid bread wheat has been used to clone one of the 5 NFCs based on BAC libraries screening and sequencing of the targeted BACs contigs using long reads sequencing technologies. Ultimately the annotation and the comparison of such sequences will be valuable to unveil mechanisms of DNA interaction that drove plant chromosomes repatterning during evolution through NFCs.

Peculiar features of common bean subtelomeres

Valérie Geffroy - IPS2, Bat630, UPS, 91 405 Orsay cedex

In common bean (*Phaseolus vulgaris*; *Pv*) genome, most of the large disease resistance (*R*) loci are localized at the end of linkage groups (LG). This is the case of the *Co-2* and *B4 R* gene clusters, localized at one end of LGB11 and LGB4, respectively. Many specific *R* genes and QTL conferring resistance to a diverse selection of pathogens including fungi (*Colletotrichum lindemuthianum*, *Uromyces appendiculatus*) and bacteria (*Pseudomonas syringae*) have been mapped to these two *R* clusters. We sequenced ≈650kb and ≈1Mb spanning the *B4* and *Co-2* clusters, respectively, and annotated more than 100 genes for each cluster with ≈1/3 of them corresponding to Coiled-Coil-Nucleotide-Binding-Site-Leucine-Rich-Repeat (CNL) genes. Concerning the *Co-2* cluster, synteny analyses revealed that it corresponds to the complex *Rpg1b* CNL cluster in soybean and that *Co-2*-related CNL sequences are also present in the corresponding region of *Medicago truncatula* (*Mt*) in chromosome *Mt8*. A completely different situation was observed for the *B4 R* gene cluster. Indeed, although conserved microsynteny was observed for non-CNL sequences between the *Pv B4* locus and corresponding regions of *Mt* and *Lotus japonicus* (*Lj*), in chromosomes *Mt6* and *Lj2,* respectively, the CNL sequences were completely absent in these regions. The origin of the *B4* CNL sequences was investigated through phylogenetic analysis, which revealed that, in the *Pv* genome, paralogous CNL genes are shared among nonhomologous chromosomes (4 and 11). Together, our results suggest that *B4* CNL derived from *Co-2* CNL, through an ectopic recombination event. Integration of the soybean genome data enabled us to date more precisely this event and to infer that a single CNL moved from the *Co-2* to the *B4* cluster. FISH analyses revealed that both *R* clusters are localized in a peculiar genomic environment, adjacent to knobs and associated with a 528-bp subtelomeric satellite repeat, referred to as *khipu*. For both *R* gene clusters, *khipu* is tightly interspaced between CNL sequences and is also present in the adjacent knobs. *Khipu* is specific to the *Phaseolus* genus and present on most chromosomal termini, indicating the existence of frequent ectopic recombination events in *Pv* subtelomeric regions. Our results highlight the importance of ectopic recombination in the evolution of *R* gene clusters and their distribution around the genome. These findings also point out the shortcomings of using data from model species to map *R* genes in crop specie

SESSION 4- MEIOTIC RECOMBINATION

A meiotic transcriptome overview in a polyploïd species (*Brassica napus*) / Le transcriptome méiotique d'une espèce polyploïde, le colza (**Aurélien BLARY – INRA Versailles**)

More information at https://colloque.inra.fr/worshop-series-dynagev

A meiotic transcriptome overview in a polyploïd species (*Brassica napus*)

Aurélien Blary, Andrew Lloyd, Guillem Rigaill Joseph Tran, Catherine Charpentier, Sandrine Balzergue, Etienne Delannoy, Delphine Charif, Eric Jenczewski

INRA, UMR1318, Institut Jean-Pierre Bourgin, RD10, F-78000 Versailles, France

C'est au cours de la méiose que se créent les nouvelles combinaisons alléliques grâce au brassage génétique produit par les Crossing Over (CO), l'un des produits de la recombinaison méiotique. Les CO, échanges réciproques de matériel génétique entre chromosomes, sont indispensables pour assurer une bonne ségrégation des chromosomes homologues lors de la première division de la méiose. Chez les espèces allopolyploïdes, la présence de plus d'un partenaire possible lors de la recombinaison requiert un niveau supplémentaire de contrôle ; il est nécessaire d'empêcher la formation de CO entre les chromosomes dits homéologues (hérités des génomes parentaux) pour assurer le bon équilibre des gamètes et le maintien de la fertilité de l'espèce. Chez le blé hexaploïde, ce contrôle semble résulter de la régulation fine du niveau d'expression d'un réseau de gènes méiotique.

Dans notre équipe, nous nous intéressons au colza (Brassica napus), jeune allotétraploïde (AACC, 2n=38) résultant de croisements multiples entre le choux (B. oleracea ; 2n=18 ; CC) et la navette (B. rapa ; 2n=20 ; AA). Plusieurs locus impliqués dans le contrôle de la formation des CO entre homéologues ont été cartographiés chez des plantes allohaploïdes (AC, n=19) (Jenczewski et al., 2003 ; Liu et al., 2006). Une analyse de type RNA-seq a été mise en œuvre pour vérifier dans quelle mesure le transcriptome méiotique varie entre les 2 variétés de colza, Darmor et Yudal, ayant permis de cartographier ces QTLs.

Pour cette analyse, les ARNs ont été extraits à partir de méiocytes isolés à partir de plantes euploïdes (AACC) et haploïdes (AC) de Darmor et de Yudal, avec pour chaque génotype 3 réplicas biologiques et 2 réplicas techniques.

La différence de niveau de transcription entre copies homéologues (portées par le génome A ou C) est la principale source de variation observée dans notre jeu de données sans pour autant refléter une dominance d'un génome sur l'autre. L'effet variété constitue la seconde source de variation : une majorité des gènes transcrits en méiose (40 à 50% des 101040 gènes décrits dans la version publiée du génome) sont différentiellement exprimés (DE) entre Darmor et Yudal. Vient enfin le niveau de ploïdie, nous avons identifié parmi les 28000 gènes DE un sous-ensemble commun de gènes aux 2 variétés en réponse à l'effet ploïdie que nous sommes en train de caractériser.

Nos analyses ont par ailleurs permis de mettre en évidence l'existence de translocations (échange intergénomique produit par un CO entre homéologues) ; certaines récentes, ségrégant entre nos réplicas biologiques, et d'autres plus anciennes, fixées dans le génome des 2 variétés étudiées. Nous sommes en train de vérifier dans quelles mesures ces translocations peuvent avoir un impact sur les analyses différentielles. Un certain nombre d'éléments laissent à penser qu'au moins une translocation pourrait co-localiser avec l'un des QTLs impliqués dans le contrôle de la formation des CO entre homéologues.

Characterization of recombination hot spots in bread wheat: Example of chromosome 3B.

Benoit DARRIER, Alain LOUSSERT, Hélène RIMBERT, Marie-Claire DEBOTE, Frédéric CHOULET, Etienne PAUX, Aurèlie EVRARD, François BALFOURIER, Pierre SOURDILLE. (UMR GDEC - INRA Clermont Ferrand)

INRA UMR 1095 GDEC Domaine de Crouël, 63100, Clermont-Ferrand France.

In a context of a changing world, we need to adapt our cultivated plants for food production to the new environmental constraints in order to achieve a more sustainable agriculture. For this purpose, genetic diversity can be used to introduce new original alleles for genes of agronomical interest in elite lines through introgression using recombination as tool for genetic shuffling.

However, in bread wheat like in many species with large genome, 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. At the present time, chromosome 3B of bread wheat is the first wheat chromosome sequenced and organized into an anchored and annotated pseudo-molecule which represents roughly 800Mb. This pseudo-molecule could allow us to investigate relationships between sequence features and recombination hot spots and linkage disequilibrium.

To understand the mechanisms that drive recombination at sequence level we precisely mapped recombination hot spots using 1271 Chinese Spring X Renan (CsRe) progeny (F6-SSD). Then we used two collections of 90 lines corresponding to Asian and European genetic pools including Chinese Spring (Asian) and Renan (European) to study the variation of linkage disequilibrium at fine scale in the regions covering our previously detected hot spots.

At the present time, 30 recombination hot spots were detected in regions of less than 10kb. These hot spots will be compared to sequence features (Epigenetic landmarks, genes, transposable elements, specific motifs) available for chromosome 3B. Concerning linkage disequilibrium, our first results showed a rough decrease of LD in our two populations in a region bearing a hot spot with some differences of linkage between surrounding markers.

SESSION 5- GENE AND GENOME DUPLICATION

Tracking duplicate copies in high-redundant *Spartina* genomes / A la recherche des copies dupliquées dans le génome hautement polyploïde des Spartines (**Julien BOUTTE - Univ-Rennes**)

Tracking duplicate copies in high-redundant *Spartina* genomes

Boutte Julien., Ainouche Malika., Salmon Armel.

UMR CNRS 6553 Ecobio, OSUR, University of Rennes 1. Bât 14A Campus Scientifique de Beaulieu, 35 042 Rennes Cedex (France)

Next generation sequencing (NGS) technologies offer new opportunities to explore polyploid genomes and their corresponding transcriptomes; however, identification of duplicated (homoeologous) gene copies remains challenging, most particularly when the actual (diploid) parental genomes are not available, and in the context of recurrent whole genome duplication events.

The *Spartina* genus (Poaceae, Chloridoideae) is particularly affected by recent hybridization and polyploidy events. Five polyploid species including the hexaploid parents *S. maritima* (2n=6x=60) and *S. alterniflora* (2n=6x=62), their homoploid hybrids (*S. x townsendii* and S*. x neyrautii*) and the recently formed allododecaploid *S. anglica* that has invaded saltmarshes on several continents, present a good system to study the short term consequences of hybridization and polyploidization events. However, identification of the duplicated genes is particularly challenging in these species, which requires the development of adapted tools.

We consequently developed a bioinformatics approach to detect homoeologues from NGS datsets. Roche-454 and Illumina data were assembled and co-assembled, obtaining a total of 110 423 annotated contigs for the 5 species (37 867 non-redundant contigs). Sequences were co-aligned to detect SNPs and different putative duplicates were detected within 286 937 contigs. This method allowed identification of homoeoSNPs (which discriminate homoelogous sequences) in the hexaploid parents and their derived hybrid and allopolyploid; it also provided the opportunity to evaluate the divergence between duplicated genes. Using synonymous substitutions (K_s), it was possible to identify and confirm the recent duplication events in the *Spartina* genus. These results now allow us to identify duplicated copies at any ploidy level in *Spartina* and to explore the fate of duplicate genes in these complex genomes as well as in any other polyploid species.

Evolution of duplicated pathways in allopolyploid cotton / Evolution de voies métaboliques dupliquées chez le coton allotétraploïde (**Joseph GALLAGHER - Iowa State University**)

Evolution of duplicated pathways in polyploid cotton

Joseph P. Gallagher, Corrinne E. Grover, Jonathan F. Wendel

The phenomenon of polyploidy, or whole genome duplication, has occurred many times throughout the evolutionary history of the land plants, particularly in the angiosperms. While the evolutionary trajectory of duplicated genes following polyploidy has been considered frequently, pathway and network features have only just begun to be incorporated into these studies. Here, we examine how the pathway characteristics of the genes of the anthocyanin biosynthetic pathway affect their evolutionary trajectory in polyploid cotton. By integrating the evolutionary effects of polyploidy and pathways, we are able to better understand the outcomes of gene duplication and polyploidy.

Importance of transposon-based genomic changes in natural and synthetic *Nicotiana* allopolyploids / Rôle des changements génomiques associés aux éléments transposables chez des

formes allotétraploïdes naturelles et synthétiques de Nicotiana (**Corinne MHIRI – INRA Versailles)**

IMPORTANCE OF TRANSPOSON-BASED GENOMIC CHANGES IN NATURAL AND SYNTHETIC NICOTIANA ALLOPOLYPLOIDS

Corinne MHIRI[1], Natacha SENERCHIA[2], Christian PARISOD[2], Julien DANIEL[1], François DORLHAC de BORNE[3], Ales KOVARIK[4], Andrew LEITCH[5] and Marie-Angèle GRANDBASTIEN[1]

[1]Institut J-P Bourgin - UMR1318, INRA, Route de Saint Cyr, 78026 Versailles Cedex, France

[2] Laboratory of Evolutionary Botany, University of Neuchâtel, 2000 Neuchâtel, Switzerland

[3] Tobacco Institute, Imperial Tobacco Group, 24100 Bergerac, France

[4] Institute of Biophysics, Academy of Sciences of the Czech Republic, CZ-61265 Brno, Czech Republic

[5] School of Biological and Chemical Sciences, Queen Mary, University of London, London E1 4NS, UK

The allopolyploidy process, by merging and doubling in a single nucleus chromosome sets originated from different species, can be considered as a “genomic shock” as predicted by Barbara McClintock. Merged genomes may undergo a wide range of structural, epigenetic and functional changes. As transposable elements (TEs) are major components of plant genomes, they may play a key role in the genetic and functional modifications produced by the allopolyploidy process.

We are currently assessing the extent of genomic changes associated with TEs in recent allopolyploid species of the *Nicotiana* genus that probably formed less than 200,000 years ago from different diploid progenitors species. We focused our study on *N. rustica*, *N. arentsii* and *N. tabacum* (tobacco), for which it is possible to synthetize *de novo* synthetic hybrids from the current actual parental species. To unravel the extent of TE-associated structural changes, we performed comparative analysis of SSAP (Sequence-Specific Amplification Polymorphism) profiles obtained for six different endogenous TE populations in both natural and synthetic accessions as well as their diploid progenitors. We estimated for each species or hybrid the importance of the genomic changes (non-additive SSAP bands : new bands and lost bands) associated with each TE. The analysis of the young *Nicotiana* allopolyploids shows that band loss is the main event, and that each TE family displays a specific evolutionary trajectory. TE divergence between progenitors is also strongly correlated with the TE-associated restructuring level in the corresponding polyploid, in agreement with the genome shock model. In *Nicotiana* synthetic hybrids, we observed mainly additive profile as expected but also some TE-related genomic restructuring in F1 hybrids, indicative of TE-associated genome reorganization at early hybridization steps.

Persistence, dispersal and genetic evolution of recently formed *Spartina* homoploid hybrids and allopolyploids in Southern England

Authors: Dalibor Huska[1*], Ilia J Leitch[2*] Julie Ferreira de Carvalho[3], Andrew R. Leitch[4], Armel Salmon[3], Malika Ainouche[3] and **Ales Kovarik**[1**]

Address: [1]Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i, Královopolská 135, Brno Czech, CZ-612 65, Czech Republic; [2]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB UK, [3]UMR 6553 ECobio Université de Rennes 1, Bât. 14A Campus de Beaulieu, France; [4]Queen Mary University of London, School of Biological and Chemical Sciences, Mile End Road, London E1 4NS, UK

Abstract

In Southampton Water, UK, the recent (c. 150 years ago) interspecific hybridisation between *Spartina alterniflora* (2n=6x=62; A-genome) and *S. maritima* (2n=6x=60; M-genome) gave rise to the homoploid hybrid (*S.* × *townsendii*, 2n=6x=62), and subsequently to the invasive allododecaploid species *S. anglica* (2n=12x=120-124) that has since spread worldwide. To address the question of dynamics of mixed ploidy populations involving these plants, we analysed three populations (fifty one individuals) in Southern England, UK., one of which was the presumed place of origin of the homoploid hybrid and the derived allopolyploid (Hythe). Using a combination of flow cytometry, ribosomal DNA (rDNA) genotyping and amplified fragment length polymorphisms (AFLPs) we were able to identify the genomic composition and ploidy level of each individual analysed. The data show that the homoploid hybrid still dominates the population at Hythe (82% of individuals analysed). We also identified *S.* × *townsendii* for the first time on Hayling Island (66% of analysed plants), indicating dispersal beyond its likely origin. The fertile allododecaploid *S. anglica* was mainly found in populations outside the initial hybridisation site (Hayling Island and Eling Marchwood). Quantification of the rDNA contributions from each parental genome showed that the ratios were mostly balanced in *S.* × *townsendii.* However, two (3%) *S. anglica* individuals have lost nearly all M-genome homeologs, indicating extensive repeat loss or homogenisation. Such variation indicates that despite the presumed single allopolyploid origin of *S. anglica* and genetic uniformity at other loci, it has undergone substantial changes at the rDNA loci following genome duplication.

Brassica allohexaploids: problems, processes and potential / Potentialités et problèmes associés à la formation d'allohexaploïdes de Brassica (**Annaliese MASON - Justus Liebig University Giessen**)

Brassica allohexaploids: problems, processes and potential

Annaliese S Mason[1,2], Junko Takahira[3], Wallace A Cowling[3], Gustavo Moreira Alves[2,4], Arkaprava Chaudhuri[2,5], Chhaya Atri[6], Alice Hayward[2], Ning Chen[2], Mohana E Ragu[2], Jessica Dalton-Morgan[2], Birgit Samans[1], Olivier Coriton[7], Virginie Huteau[7], Frédérique Eber[7], Anne-Marie Chèvre[7], Jacqueline Batley[2,3], Matthew N Nelson[3]

Abstract

Brassica crops *B. napus* (oilseed rape, 2n = AACC), *B. carinata* (Ethiopian mustard, 2n = BBCC) and *B. juncea* (Indian mustard, 2n = AABB) are allopolyploid, with two genomes each from diploid species *B. rapa* (Chinese cabbage, turnip; 2n =AA), *B. nigra* (black mustard; 2n = BB) and *B. oleracea* (cabbage, cauliflower, broccoli; 2n = CC). No naturally occurring allohexaploid *Brassica* (2n = AABBCC) exists. However, combining the three genomes may allow production of a new crop benefiting from increased hybrid vigour through allelic heterosis, as well as shedding light on polyploid speciation processes in *Brassica*.

Genome stability and meiotic behaviour was investigated in different allohexaploid *Brassica* populations. High-throughput molecular genotyping and cytogenetics approaches were used to track inheritance of the A, B and C genome alleles, flow cytometry was used to estimate DNA content and fertility data were collected.

The majority of interactions between homologous chromosome pairs in the first generation were normal, although homoeologous chromosome interactions occurred at low frequencies for most chromosomes (0 - 14%). Somewhat surprisingly, strong selective pressure for 2n = AABBCC karyotypes was not observed, although balanced translocations (duplication/deletions between homoeologues) were more common than unbalanced duplication or deletion events. Most chromosomes present in only one copy (univalents) were just as likely to be lost as to be passed on to the next generation, and increasingly aneuploid chromosome complements were observed in some second-generation progeny. However, fertility in some allohexaploid lines has remained high for several generations.

Analysis of this material may shed light on genetic and genomic regions associated with meiotic stability in *Brassica*, and facilitate genomic selection for meiotically stable allohexaploid genotypes.

[1]Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany

LISTES DES PARTICIPANTS

AINOUCHE	Malika	Univ Rennes I
AINOUCHE	Kader	Univ Rennes I
AKAKPO	Roland	IRD-Montpellier
BARRIERE	Yves	CNRGV-Toulouse
BAURENS	Franc-Christophe	CIRAD Montpellier
BELLEC	Arnaud	CNRGV INRA TOULOUSE
BERGES	Hélène	CNRGV-Toulouse
BILLY	Béatrice	GEVES Beaucouze
BLARY	Aurélien	INRA VERSAILLES
BONNET	Béatrice	GEVES
BOUTTE	Julien	Univ-Rennes I
CABELLO	Francisco	Univ-Rennes I
CHALHOUB	Boulos	URGV-EVRY
CHEVRE	Anne-Marie	UMR APBV - RENNES
CORITON	Olivier	UMR APBV - RENNES
COURTIAL	Aurdrey	CNRGV-Toulouse
CURK	Franck	INRA CORSE
DARRIER	Benoit	INRA-Clermont
DE KOCHKO	Alexandre	IRD-DIADE
DENIOT	Gwenaelle	UMR IGEPP Rennes
D'HONT	Angélique	CIRAD_AGAP
FALENTIN	Cyril	UMR IGEPP Rennes
GALLAGNER	Joseph	US
GARCIA	Sonia	Universitat de Barcelona
GARNATJE ROCA	Teresa	Institut Botànic de Barcelona
GARSMEUR	Olivier	CIRAD-AGAP
GEFFROY	Valerie	IPS2
GONZALO	Adrian	INRA VERSAILLES
GRAS MAS	Airy	Universitat de Barcelona
GUYOT	Romain	IRD Montpellier
HAMON	Perla	IRD Montpellier
JENCZEWSKI	Eric	INRA-Versaille
JOETS	Johann	INRA-Le Moulon
JOUSSEAUME	Thibault	INRA-Rennes
KELLER	Jean	Univ-Rennes I
KERBRAT	Florian	INRA-Rennes
KOVARIK	Ales	Institute of Biophysics, Czech Republic
LASHERMES	Philippe	IRD-RPB
LODE	Maryse	UMR APBV - RENNES
MARANDE	William	CNRGV-Toulouse
MARTIN	Guillaume	CIRAD Montpellier
MASON	Annaliese	Australie
NEGRE	Sylvie	INRA LE RHEU
NICOLAS	Stéphane	UMR GV
MHIRI	Corinne	INRA Versailles
MICHARD	Robin	INRA Clermont
PAUX	Etienne	INRA Clermont
PICAULT	Nathalie	Univ-Perpignan
PRAT	Elise	CNRGV-Toulouse
REMAY	Arnaud	GEVES
ROUSSEAU-GUEUTIN	Mathieu	Univ-Rennes I
ROUSSEAU	Helène	Univ-Rennes I
SALMON	Armel	Univ-Rennes I
TROTOUX	Gwenn	UMR IGEPP Rennes