Promotion IUF 2016
Rapport d'activité (2016-2021)

NOM : Schacherer
PRÉNOM : Joseph
DATE DE NAISSANCE : 15 mars 1976
GRADE : Professeur des Universités (PR1)
DISCIPLINE PRINCIPALE : Génétique
CNU : section 65
UNIVERSITÉ OU ÉTABLISSEMENT D'APPARTENANCE : Université de Strasbourg
UNITÉ DE RECHERCHE D'APPARTENANCE : Génétique Moléculaire, Génomique et Microbiologie (GMGM) – UMR7156

CATÉGORIE : JUNIOR

THÉMATIQUE DE RECHERCHE :

VARIATION INTRA-SPÉCIFIQUE ET ÉVOLUTION DES GÉNOMES

RÉSUME SCIENTIFIQUE À PROPOS DE LA RÉALISATION DU PROJET DE RECHERCHE IUF:

Over the past 5 years, the team focused on two main topics: (i) the study of the genetic and phenotypic diversity at a population level, *i.e.* population genomic surveys (ii) the characterization of the genetic basis and architecture of complex traits using yeast as a model organism. We have made important progress in our field and I only detail the significant results below.

Understanding how genetic variation among individuals leads to the phenotypic diversity observed within a species is a long-standing and fundamental goal in biology. Advances in whole genome sequencing have allowed genome-wide association studies (GWAS) to correlate genetic variants with various traits, including many human diseases. However, while thousands of causal variants have been identified, they often only explain a fraction of the observed phenotypic variance. While there are a number of possible explanations for this "missing heritability", rare variants, genetic interactions and structural variants are likely to be a significant culprit. During this reporting period, we took advantage of lessons learned from systematic functional genomic analyses in *S. cerevisiae* and we use it to explore the origin of the missing heritability from population data. In parallel we built new tools that allow us to perform large scale functional genomics in diverse strain backgrounds.

Species-wide exploration of the genetic and phenotypic landscapes

To lay the foundation of the exploration of the genetic complexity of traits, we first completely sequenced and phenotyped a large collection of 1,011 yeast isolates, coming from diverse ecological niches across the world on plates containing different stressors. The generated dataset revealed an undescribed evolutionary history as well as the driving forces of genome evolution, and has provided insights into the genotype–phenotype relationship.

To dissect the genetic basis of complex traits, we developed an efficient, standardized and quantitative high-throughput screening strategy based on an automated robotic plateform (Singer RoToR HDA robot) to replicate the set of strains onto various media (more than 50 growth conditions). In total, we determined and analyzed 34,956 phenotypic measurements covering a large number of traits providing a comprehensive analysis of their inheritance patterns. The combination of these two datasets allowed to perform genome-wide association studies. We performed mixed-model association and detected 35 variants significantly associated with 14 conditions, with an enrichment and high association scores for Copy Number Variants (22 CNVs *vs.* 13 SNPs). In addition, some of the detected variants are linked to variable ORFs, which are not present in the reference genome. Phenotypic variance explained was estimated by running a new association with a similarity matrix containing the significantly associated markers. For five of the tested traits, the phenotypic variation explained is surprisingly greater than 25%. In fact, CNVs explained larger proportions of trait variance compared to SNPs, with a median of 36.8% and 4.49% of the variance explained, respectively. Our genome-wide association analyses, including an exhaustive catalog of genome content and CNVs present in the 1,011 genomes, highlighted the overall importance of these genetic variants on the phenotypic diversity.

Interestingly, the difference between the estimated genome-wide heritability and explained phenotypic variance gives an overview of the extent of missing heritability. Many SNPs are present at low frequencies, which echoes observations previously made in human GWAS and raised the question of whether rare SNPs have an important role in modulating the phenotypic landscape.

Rare variants contribute significantly to quantitative trait variation

Based on the genomic and phenotypic data from the collection of 1,011 *S. cerevisiae* isolates, a total of 55 isolates that are diploid, homozygous, genetically diverse and present unbiased population structure have been selected. In total, we created 3,025 hybrids, representing 2,970 heterozygous hybrids with a unique parental combination and 55 homozygous hybrids. We screened the entire set of isolates, and hybrids for high-resolution quantification of mitotic growth ability across 53 conditions, using the automated, standardized robotic platform. The conditions included different carbon sources and chemical compounds impacting various physiological and cellular responses such as membrane and protein stability, signal transduction, sterol biosynthesis, transcription, translation as well as osmotic and oxidative stress. This phenotyping step led to the characterization of more than 160,00 hybrid/trait combinations.

Our dataset gave us the opportunity to have a species-wide view of the relative contributions of additive and non-additive components to overall phenotypic variation in our large sample of 3,025 hybrids. Separation of variance components of phenotypic variation is a key point in the understanding of genetic architecture of traits and a potential source of the missing heritability. In addition, because we tested many traits across a large population, we highlighted interesting variations between conditions as well as isolates, suggesting the presence of rare functional variants and potential expressivity cases.

Because all the 55 isolates used in the diallel cross were completely sequenced, we examined the genetic basis of traits across a large panel of traits using GWAS. The main point about using a diallel cross design was that we can use the redundancy of the haplotype which is intrinsic of the pairwise crosses to our advantage. Indeed, with only 34 base genomes, we could generate *in silico* the 595 genomes corresponding to a half pairwise matrix of 561 hybrids with 34 homozygous. Each parental genome is present 34 times hence creating haplotype mixing across the matrix. This high level of haplotype shuffling and repetition gives the advantage of offering allele overrepresentation compared to the use of a population with the same number of independent individuals. Minor allele frequency (MAF) will be fundamentally changed in a diallel compared to the species level because of the smaller number of parents involved. In our diallel panel, out of the total 31,632 SNPs retained, 3.5% (1,128) which had a MAF < 5% in the 1,011 *S. cerevisiae* genomes happen to

surpass this threshold in the diallel panel and thus are now detectable, going up to a MAF of 32%. Surprisingly, 12.1% of the significantly associated SNPs also surpassed this threshold meaning that they could not have been detected by a classical GWAS approach. Altogether, our results have major implications for our understanding of the genetic architecture of traits in the context of unexplained heritability. They clearly highlight the extensive role of low-frequency and rare variants on the phenotypic variation at the population level.

Global picture of inheritance patterns and the genetic complexity of traits

We then took advantage of the diallel panel to assess the overall genetic complexity of traits and the prevalence of phenotypic expressity at a population-scale. In this context, we first selected a subset of the large diallel hybrid panel in order to have 190 unique hybrids coming from 20 natural isolates representative of the *S. cerevisiae* genetic diversity. For each of these hybrids, a large progeny of 160 individuals (corresponding to 40 full tetrads) was obtained, leading to a total of 30,400 offspring individuals. Their mitotic growth has been assessed on 40 growth conditions inducing various cellular stress.

The main objective of this work has been to infer complexity level of traits at a population scale and assess its dynamic across multiple genetic backgrounds. To do so, we conducted a large-scale phenotyping of the whole panel of 30,400 haploid progeny coming from 190 hybrids. Overall, more than three million phenotypic measurements were performed and grouped for each cross and condition (trait) to obtain 7,600 phenotypic distributions of haploid progenies *i.e.* one distribution for each cross/trait combination. Manually inferring the complexity level for each of the cross/trait combination would be tedious and error prone. To help us in this task, we based our analysis on a constructed decision tree to classify distributions into different inheritance categories based on their underlying genetic complexity. Yet, the first step of this process was the determination of unimodality vs. bimodality of the distribution. This distinction is far from trivial and required to be assessed in a very specific manner. To do so, we used a machine learning algorithm, more precisely, we build a random forest classifier.

We classified the distributions in one of three complexity level: monogenic, oligogenic and complex. Overall, 80.3% of the considered distributions displayed inheritance patterns corresponding to a complex inheritance pattern. In the meantime, 11.2% appear as monogenic and only 4% as oligogenic. The remaining 4.5% failed to be sorted into one of the previous categories for various reasons, either the parents could not be confidently attributed to one cluster or the tetrad segregation phenotype could not result in a confident classification. These results confirm the fact that inheritance patterns are mainly complex but also that in a non-negligible number of cases, one gene is actually responsible for most of the observed genetic variance. However, this overview can be completed by the fact that this repartition of the complexity is highly dependent on the condition considered. Indeed, extensive variation in the complexity repartition can be observed in the 40 conditions explored here. When further dissecting the 153 distributions corresponding to an oligogenic inheritance, we highlighted several types of digenic interaction. We detected 87 cases of recessive epistasis. In 66 cross/trait combinations, modifier gene suggesting the presence of a suppressor have been identified.

Overall, we were able to assess the complexity level of traits at a species-wide level. We also highlighted the prevalence of expressivity with most of the followed variants displaying departure from monogenic inheritance patterns. Finally, this works lays the ground for a more complete and in detail exploration of variants displaying different levels of expressivity by dissecting the genetic basis of the observed cases. This dissection is the next step and will allow to have a better insight into the phenotypic expressivity landscape.

PRODUCTION SCIENTIFIQUE DE LA PÉRIODE 2016-2021 :

1. Eberlein C, Abou Saada O, Friedrich A, Albertin W, Schacherer J. Different trajectories of polyploidization shape the genomic landscape of the *Brettanomyces bruxellensis* yeast species. Genome Res. 2021 Nov 23;31(12):2316. doi:10.1101/gr.275380.121.

2. Papaioannou IA, Dutreux F, Peltier FA, Maekawa H, Delhomme N, Bardhan A, Friedrich A, Schacherer J, Knop M. Sex without crossing over in the yeast *Saccharomycodes ludwigii*. Genome Biol. 2021 Nov 3;22(1):303. doi:10.1186/s13059-021-02521-w.

3. Peltier E, Bibi-Triki S, Dutreux F, Caradec C, Friedrich A, Llorente B, Schacherer J. Dissection of quantitative trait loci in the *Lachancea waltii* yeast species highlights major hotspots. G3 (Bethesda). 2021 Sep 6;11(9):jkab242. doi: 10.1093/g3journal/jkab242.

4. Dutta A, Dutreux F, Schacherer J. Loss of heterozygosity results in rapid but variable genome homogenization across yeast genetic backgrounds. Elife. 2021 Jun 23;10:e70339. doi: 10.7554/eLife.70339.

5. Bleykasten-Grosshans C, Fabrizio R, Friedrich A, Schacherer J. Species-wide transposable element repertoires retrace the evolutionary history of the *Saccharomyces cerevisiae* host. Mol Biol Evol. 2021 Sep 27;38(10):4334-4345. doi:10.1093/molbev/msab171.

6. Abou Saada O, Tsouris A, Eberlein C, Friedrich A, Schacherer J. nPhase: an accurate and contiguous phasing method for polyploids. Genome Biol. 2021 Apr 29;22(1):126. doi:10.1186/s13059-021-02342-x.

7. D'Angiolo M, De Chiara M, Yue JX, Irizar A, Stenberg S, Persson K, Llored A, Barré B, Schacherer J, Marangoni R, Gilson E, Warringer J, Liti G. A yeast living ancestor reveals the origin of genomic introgressions. Nature. 2020 Nov;587(7834):420-425. doi:10.1038/s41586-020-2889-1.

8. Stoneman HR, Wrobel RL, Place M, Graham M, Krause DJ, De Chiara M, Liti G, Schacherer J, Landick R, Gasch AP, Sato TK, Hittinger CT. CRISpy-Pop: A Web Tool for Designing CRISPR/Cas9-Driven Genetic Modifications in Diverse Populations. G3 (Bethesda). 2020 Nov 5;10(11):4287-4294. doi:10.1534/g3.120.401498.

9. De Chiara M, Friedrich A, Barré B, Breitenbach M, Schacherer J, Liti G. Discordant evolution of mitochondrial and nuclear yeast genomes at population level. BMC Biol. 2020 May 11;18(1):49. doi:10.1186/s12915-020-00786-4.

10. Brion C, Caradec C, Pflieger D, Friedrich A, Schacherer J. Pervasive phenotypic impact of a large non-recombining introgressed Region in Yeast. Mol Biol Evol. 2020 Sep 1;37(9):2520-2530. doi:10.1093/molbev/msaa101.

11. Gounot JS, Neuvéglise C, Freel KC, Devillers H, Piskur J, Friedrich A, Schacherer J. High complexity and degree of genetic variation in *Brettanomyces bruxellensis* population. Genome Biol Evol. 2020 Jun 1;12(6):795-807. doi:10.1093/gbe/evaa077.

12. Fournier T, Abou Saada O, Hou J, Peter J, Caudal E, Schacherer J. Extensive impact of low-frequency variants on the phenotypic landscape at population-scale. Elife. 2019 Oct 24;8:e49258. doi:10.7554/eLife.49258.

13. Defenouillère Q, Verraes A, Laussel C, Friedrich A, Schacherer J, Léon S. The induction of HAD-like phosphatases by multiple signaling pathways confers resistance to the metabolic inhibitor 2-deoxyglucose. Sci Signal. 2019 Sep 3;12(597):eaaw8000. doi: 10.1126/scisignal.aaw8000.

14. Fleiss A, O'Donnell S, Fournier T, Lu W, Agier N, Delmas S, Schacherer J, Fischer G. Reshuffling yeast chromosomes with CRISPR/Cas9. PLoS Genet. 2019 Aug 29;15(8):e1008332. doi:10.1371/journal.pgen.1008332.

15. Peltier E, Friedrich A, Schacherer J, Marullo P. Quantitative trait nucleotides impacting the technological performances of industrial *Saccharomyces cerevisiae* strains. Front Genet. 2019 Jul 23;10:683. doi:10.3389/fgene.2019.00683. PMID: 31396264; PMCID: PMC6664092.

16. Ishchuk OP, Ahmad KM, Koruza K, Bojanovi*ƒ*ç K, Sprenger M, Kasper L, Brunke S, Hube B, ST, Hellmark T, Gullstrand B, Brion C, Freel K, Schacherer J, Regenberg B, Knecht W, Piskur J. RNAi as a Tool to study virulence in the pathogenic yeast *Candida glabrata*. Front Microbiol. 2019 Jul 24;10:1679.
doi: 10.3389/fmicb.2019.01679.

17. Fairhead C, Fischer G, Liti G, Neuvéglise C, Schacherer J. André Goffeau's imprinting on second generation yeast "genomologists". Yeast. 2019 Apr;36(4):167-175. doi: 10.1002/yea.3377.

18. Raghavan V, Bui DT, Al-Sweel N, Friedrich A, Schacherer J, Aquadro CF, Alani E. Incompatibilities in mismatch repair genes *MLH1-PMS1* contribute to a wide range of mutation rates in Human isolates of baker's yeast. Genetics. 2018 Dec;210(4):1253-1266. doi: 10.1534/genetics.118.301550.

19. Peter J, De Chiara M, Friedrich A, Yue JX, Pflieger D, Bergstrom A, Sigwalt A, Barre B, Freel K, Llored A, Cruaud C, Labadie K, Aury JM, Istace B, Lebrigand K, Barbry P, Engelen S, Lemainque A, Wincker P, Liti G, Schacherer J. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature. 2018 Apr;556(7701):339-344. doi: 10.1038/s41586-018-0030-5.

20. Avramova M, Cibrario A, Peltier E, Coton M, Coton E, Schacherer J, Spano G, Capozzi V, Blaiotta G, Salin F, Dols-Lafargue M, Grbin P, Curtin C, Albertin W, Masneuf-Pomarede I. *Brettanomyces bruxellensis* population survey reveals a diploid-triploid complex structured according to substrate of isolation and
geographical distribution. Sci Rep. 2018 Mar 7;8(1):4136. doi:10.1038/s41598-018-22580-7.

21. Fournier T, Gounot JS, Freel K, Cruaud C, Lemainque A, Aury JM, Wincker P, Schacherer J, Friedrich A. High-quality *de novo* genome assembly of the *Dekkera bruxellensis* yeast using Nanopore MinION Sequencing. G3 (Bethesda). 2017 Oct 5;7(10):3243-3250. doi: 10.1534/g3.117.300128.

22. Fournier T, Schacherer J. Genetic backgrounds and hidden trait complexity in natural populations. Curr Opin Genet Dev. 2017 Dec;47:48-53. doi:10.1016/j.gde.2017.08.009.

23. Zhou N, Bottagisi S, Katz M, Schacherer J, Friedrich A, Gojkovic Z, Swamy KBS, Knecht W, Compagno C, Piskur J. Yeast-bacteria competition induced new metabolic traits through large-scale genomic rearrangements in *Lachancea kluyveri*. FEMS Yeast Res. 2017 Sep 1;17(6). doi:10.1093/femsyr/fox060.

24. Smukowski Heil C, Burton JN, Liachko I, Friedrich A, Hanson NA, Morris CL, Schacherer J, Shendure J, Thomas JH, Dunham MJ. Identification of a novel interspecific hybrid yeast from a metagenomic spontaneously inoculated beer sample using Hi-C. Yeast. 2018 Jan;35(1):71-84. doi:10.1002/yea.3280.

25. Brion C, Legrand S, Peter J, Caradec C, Pflieger D, Hou J, Friedrich A, Llorente B, Schacherer J. Variation of the meiotic recombination landscape and properties over a broad

evolutionary distance in yeasts. PLoS Genet. 2017 Aug 1;13(8):e1006917. doi:10.1371/journal.pgen.1006917.

26. Ohnuki S, Okada H, Friedrich A, Kanno Y, Goshima T, Hasuda H, Inahashi M, Okazaki N, Tamura H, Nakamura R, Hirata D, Fukuda H, Shimoi H, Kitamoto K, Watanabe D, Schacherer J, Akao T, Ohya Y. Phenotypic diagnosis of lineage and differentiation during sake yeast breeding. G3 (Bethesda). 2017 Aug 7;7(8):2807-2820. doi:10.1534/g3.117.044099.

27. Istace B, Friedrich A, d'Agata L, Faye S, Payen E, Beluche O, Caradec C, Davidas S, Cruaud C, Liti G, Lemainque A, Engelen S, Wincker P, Schacherer J, Aury JM. *de novo* assembly and population genomic survey of natural yeast isolates with the Oxford Nanopore MinION sequencer. Gigascience. 2017 Feb 1;6(2):1-13. doi: 10.1093/gigascience/giw018.

28. Bui DT, Friedrich A, Al-Sweel N, Liti G, Schacherer J, Aquadro CF, Alani E. Mismatch repair incompatibilities in diverse yeast populations. Genetics. 2017 Apr;205(4):1459-1471. doi:10.1534/genetics.

29. Hou J, Schacherer J. Fitness trade-offs lead to suppressor tolerance in yeast. Mol Biol Evol. 2017 Jan;34(1):110-118. doi: 10.1093/molbev/msw225.

30. Hou J, Sigwalt A, Fournier T, Pflieger D, Peter J, de Montigny J, Dunham MJ, Schacherer J. The hidden complexity of Mendelian traits across natural yeast populations. Cell Rep. 2016 Jul 26;16(4):1106-1114. doi:10.1016/j.celrep.2016.06.048.

31. Sigwalt A, Caradec C, Brion C, Hou J, de Montigny J, Jung P, Fischer G, Llorente B, Friedrich A, Schacherer J. Dissection of quantitative traits by bulk segregant mapping in a protoploid yeast species. FEMS Yeast Res. 2016 Aug;16(5):fow056. doi:10.1093/femsyr/fow056.

32. Schacherer J. Beyond the simplicity of Mendelian inheritance. C R Biol. 2016 Jul-Aug;339(7-8):284-8. doi:10.1016/j.crvi.2016.04.006.

33. Hou J, Fournier T, Schacherer J. Species-wide survey reveals the various flavors of intraspecific reproductive isolation in yeast. FEMS Yeast Res. 2016 Aug;16(5):fow048. doi:10.1093/femsyr/fow048.

34. Vakirlis N, Sarilar V, Drillon G, Fleiss A, Agier N, Meyniel JP, Blanpain L, Carbone A, Devillers H, Dubois K, Gillet-Markowska A, Graziani S, Huu-Vang N, Poirel M, Reisser C, Schott J, Schacherer J, Lafontaine I, Llorente B, Neuvéglise C, Fischer G. Reconstruction of ancestral chromosome architecture and gene repertoire reveals principles of genome evolution in a model yeast genus. Genome Res. 2016 Jul;26(7):918-32. doi:10.1101/gr.204420.116.

35. Brion C, Pflieger D, Souali-Crespo S, Friedrich A, Schacherer J. Differences in environmental stress response among yeasts is consistent with species-specific lifestyles. Mol Biol Cell. 2016 May 15;27(10):1694-705. doi:10.1091/mbc.E15-12-0816.

36. Jung PP, Sigwalt A, Ohnuki S, de Montigny J, Ohya Y, Schacherer J. Large-scale survey of intraspecific fitness and cell morphology variation in a protoploid yeast species. G3 (Bethesda). 2016 Apr 7;6(4):1063-71. doi:10.1534/g3.115.026682.

37. Freel KC, Friedrich A, Sarilar V, Devillers H, Neuvéglise C, Schacherer J. Whole-genome sequencing and intraspecific analysis of the yeast species *Lachancea quebecensis*. Genome Biol Evol. 2016 Jan 5;8(3):733-41. doi:10.1093/gbe/evv262.

38. Peter J, Schacherer J. Population genomics of yeasts: towards a comprehensive view across a broad evolutionary scale. Yeast. 2016 Mar;33(3):73-81. doi:10.1002/yea.3142.

39. Hou J, Schacherer J. On the mapping of epistatic genetic interactions in natural isolates: combining classical genetics and genomics. Methods Mol Biol. 2016;1361:345-60. doi:10.1007/978-1-4939-3079-1_19.

Since 2016, I had the chance to be invited to more than 30 international conferences and research institutes to talk about the research enabled by IUF.

ENCADREMENT DOCTORAL (Direction de thèses) :

Jing Hou, 2012 - 2016, Encadrement 100%, CDI publique (CRCN CNRS)

Anastasie Sigwalt, 2012 - 2016, Co-encadrement 70% avec Jacky de Montigny, CDI publique (professeur de l'enseignement publique)

Peter Jackson, 2013 - 2017, Co-encadrement 50% avec Anne Friedrich, CDD publique (postdoctorat)

Jean-Sébastien Gounod, 2014 - 2018, Co-encadrement 50% avec Anne Friedrich, CDD publique (postdoctorat)

Téo Fournier, 2015 - 2019, Co-encadrement 80% avec Jacky de Montigny, CDI privé (Responsable d'exploitation, GenoScreen)

Elodie Caudal, 2017 - 2021, Encadrement 100%, CDI (YCE Partners)

Omar Abou Saada, 2017 - 2021, Co-encadrement 50% avec Anne Friedrich, CDD publique

AUTRES AVANCÉES SIGNIFICATIVES AU COURS DE LA PÉRIODE : /

PRIX ET DISTINCTIONS SCIENTIFIQUES OBTENUS AU COURS DE LA PÉRIODE

This project and the preliminary results we obtained allowed to access to 3 new fundings:

2019-2023 ANR - Projet BrettAdapt, Partner, 270 k€
Populational and multi-dimensional survey of the evolution, impacts and consequences of polyploidization in a yeast model

2019-2023 ANR - Projet RecombFun, Partner, 240 k€
Partner Evolution of the recombinational landscape and functional patterns across yeast species.

2018-2023 ERC (European Research Council) Consolidator, Coordinator, 1999 k€
Inheritance, expressivity and epistasis hidden behind the phenotypic landscape of natural populations

AUTRES OBSERVATIONS : /

Acceptez-vous la mise en ligne de ce document sur le site internet de l'IUF : OUI