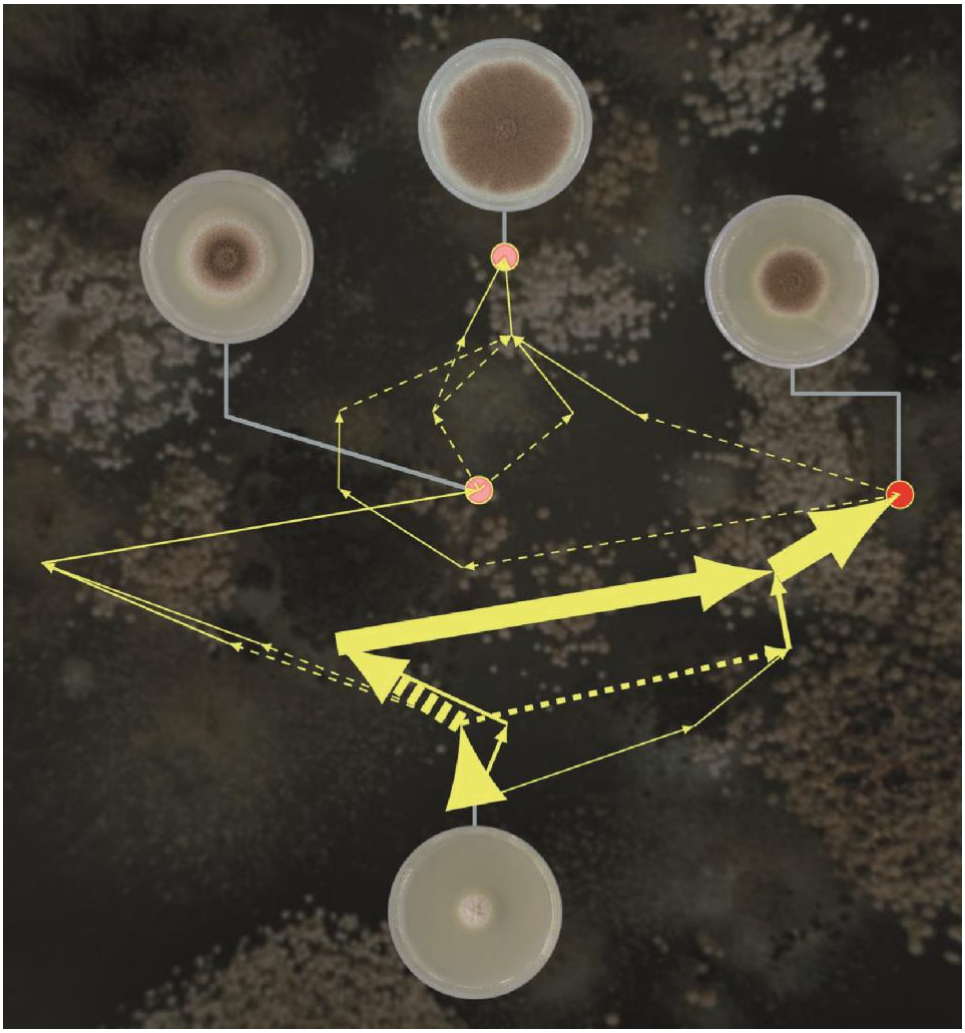
Collaborative Research Centre 680

Molecular Basis of Evolutionary Innovations



Scientific Summary Report
2006 – 2017

University of Cologne

Collaborative Research Centre 680

Molecular Basis of Evolutionary Innovations

Scientific Summary Report
2006 – 2017

University of Cologne

Cover illustration: The diagram shows pathways in genotype space traced out by an ensemble of simulated populations evolving on an empirical fitness landscape of the filamentous fungus Aspergillus niger. Observed evolutionary endpoints are marked by circles. They are connected to images of A. niger mutant colonies with varying degrees of adaptation, as reflected by their different size. The background shows competing colonies of mutants. Image courtesy of Ivan G. Szendro (University of Cologne) and J. Arjan G.M. de Visser (Wageningen University).

Table of contents

1.	Research synopsis	3
2.	Core publications	17
3.	List of scientific projects	21
4.	Scientific events	27
5.	Organization 2014 - 2017	29

1. Research synopsis

The research of CRC 680 took place amidst rapidly increasing data of genome sequences, gene regulation, and metabolic networks in cells. The new molecular data opened unprecedented opportunities to answer core questions of evolution. We endeavored to understand how differences between species build upon variation within species. By comparative sequence and functional analysis, we showed how functional innovations arise from changes of gene interactions in regulatory and metabolic networks. Repeatable network changes explain, for example, the convergent evolution of photosynthesis, the emergence of new flowering cycles and leaf shapes in plants, and the evolution of embryonal developmental patterns in insects. We also found a complementary mode of innovation: novel genes, which sometimes arise from noncoding genome sequence, can generate lineage-specific biological functions.

Of equal importance, CRC 680 developed new, quantitative links between evolutionary experiment, data analysis, and theory. These advances made critical use of time-resolved data from microbial and viral systems, where evolution can be followed in real time. We showed that many such systems have a fast-paced mode of evolution, in which several beneficial mutations compete for success. Which mutations are beneficial is captured by empirical fitness landscapes that we deduced from observed time courses of evolution. But understanding why mutations are beneficial rests on the physical basis of evolution, which was a core theme of the CRC. We studied how mutations affect key biophysical phenotypes in a cell, such as binding affinities, protein stability, and metabolic fluxes, and we quantified how these phenotypes impact organismic functions and fitness. The resulting evolutionary dynamics reveals far-reaching links to statistical physics: equilibrium systems obey a Boltzmann statistics of fitness, while time-dependent fitness drives non-equilibrium adaptive changes. Together, our research lead to a better understanding of fast-evolving systems, including important human pathogens, and paved the way for new applications of evolutionary biology to medicine and public health.

More about our research can be found on the CRC 680 website.

Die Forschung des SFB 680 fand im Kontext rasch wachsender Daten zu Genomsequenzen, Genregulation und metabolischen Netzwerken statt. Die neuen molekularen Daten eröffneten ungeahnte Möglichkeiten zur Beantwortung von Kernfragen der Evolution. Eines unserer wesentlichen Ziele war zu verstehen, wie Differenzen zwischen Spezies auf Variationen innerhalb einer Spezies aufbauen. Durch vergleichende genomische und funktionale Analysen haben wir gezeigt, wie funktionale Innovationen aus Änderungen von Gen-Wechselwirkungen in regulatorischen und metabolischen Netzwerken entstehen. Wiederholbare Veränderungen in Netzwerken erklären zum Beispiel die konvergente Evolution der Photosynthese, die Entstehung neuer Blütezyklen und Blattformen in Pflanzen sowie die Evolution der embryonalen Entwicklung von Insekten. Wir fanden auch einen komplementären Innovationsmodus: neue Gene, die manchmal aus nicht-kodierender Genomsequenz entstehen, können biologische Funktionen bestimmter Abstammungslinien erzeugen.

Von gleicher Wichtigkeit war unsere Entwicklung neuer, quantitativer Verbindungen zwischen Experiment, Datenanalyse und Theorie in der Evolution. Dieser Fortschritt beruhte wesentlich auf zeitaufgelösten Daten mikrobieller und viraler Systeme, in denen die Evolution in Echtzeit verfolgt werden kann. Wir konnten zeigen, dass viele derartige Systeme einen schnellen Evolutionsmodus haben, in dem mehrere vorteilhafte Mutationen im Wettbewerb stehen. Welche Mutationen vorteilhaft sind wird durch empirische Fitnesslandschaften beschrieben, die wir aus beobachteten Zeitreihen evolutionärer Prozesse ableiten konnten. Aber ein Verständnis, warum Mutationen vorteilhaft sind, muss auf den physikalischen Grundlagen der Evolution aufbauen; dies war ein Kernthema des SFBs. Wir haben untersucht, wie Mutationen wichtige biophysikalische Merkmale einer Zelle beeinflussen, darunter Bindungsenergien, die Stabilität von Proteinen und metabolische Flüsse, und wir haben die Auswirkung dieser Phänotypen auf organismische Funktionen und Fitness bestimmt. Die resultierende evolutionäre Dynamik zeigt weitreichende Querverbindungen zur statistischen Physik: Systeme im Gleichgewicht gehorchen einer Boltzmann-Statistik der Fitness, während zeitabhängige Fitness Adaptationsprozesse im Nichtgleichgewicht erzeugt. Insgesamt hat unsere Forschung zu einem besseren Verständnis schnell evolvierender Systeme geführt, darunter wichtiger Krankheitserreger im Menschen, und damit Wege für neue Anwendungen der Evolutionsbiologie in der Medizin eröffnet.

Mehr über unsere Forschung findet sich auf der Webseite des SFB 680.

The research of CRC 680 was organized in three project areas: A – **Evolution between species**, **B – Evolution within species**, and **C – Theory and computational analysis**, with multiple links between projects within and across areas. In the following, we summarize the scientific progress of individual projects, occasionally deviating from their numerical order to display thematic relations.

A number of projects were devoted to the evolution of developmental patterns by changes in gene regulation. Projects A1, A2, and A12 used the red flour beetle *Tribolium castaneum* to explore the evolution of early embryonic patterning in insects (A1, A2) and the consequences of gene duplications for development (A12). The short **project A2** focused on the anteroposterior axis, in particular on the hierarchical gene regulatory network controlling segment formation. In the course of a systematic in situ hybridization screen, new segmentation genes were discovered. A new member of the gap family, *mille-pattes* (*mlpt*) had an unusual gene structure (Savard et al., 2006). It encodes a polycistronic mRNA that gives rise to multiple peptides. The gene *mlpt* turned out to be the prototype of a new eukaryotic gene family (Couso and Patraquim, 2017; Galindo et al., 2007).

Project A1 analysed dorsoventral axis formation with a particular focus on the DV function of the Toll pathway as evolutionary innovation of insects. Outside the insects Toll is only known for its immune function. After initial work with *Tribolium* (Nunes da Fonseca et al., 2008) the approach was extended to the jewel wasp *Nasonia vitripennis* (Ozuak et al., 2014) and the milkweed bug *Oncopeltus fasciatus* (Sachs et al., 2015). The comparative work revealed evolutionary steps by which Toll's DV function gained increasing complexity, culminating in insects like *Drosophila*, where Toll acts as morphogen specifying polarity and pattern along the entire DV axis. The dorsoventral network of *Oncopeltus* might be close to the ancestral situation. Here, Toll lacks all instructive functions and acts only as broadly distributed polarity cue for a dynamic BMP signalling network. A theoretical model of the *Oncopeltus* dorsoventral network, which has been developed in collaboration with project C1, showed that shallow Toll signalling is sufficient to initiate axis formation (Sachs et al., 2015). Broad Toll signalling in insect eggs might originally have had an antimicrobial function from which Toll was recruited for dorsoventral patterning. Thus, our work provides an example for the initial co-option and progressive evolution of a developmental pathway.

Project A12 used the divergence of two hox3/zen paralogues in *Tribolium* to study the emergence of new developmental functions after gene duplication. While *Tc-zen1* has an early patterning role, *Tc-zen2* has a late function during morphogenesis. Evolutionary comparisons suggest that the late function is ancestral while the early function emerged within the Coleoptera (beetles). Functional and expression studies reveal that the late acting paralog (*Tc-zen2*) represses the early acting paralog, and thus restricts its time of action explaining the unusual RNA and protein profiles of *Tc-zen2*. Comparisons with closely related *Tribolium* species suggest positive selection on the coding sequences for the DNA-binding homeodomain that would have promoted paralog functional divergence.

In the first funding period, **project A6** studied the evolution of parthenogenesis in nematodes (Heger et al., 2010). The work involved in-depth analyses of available nematode genomes and lead to the discovery that an important insulator protein, CTCF, was lost during nematode evolution (Heger et al., 2009). The implications of this finding for metazoan evolution led to a collaboration in a new CRC project starting in the second funding period (Heger et al., 2012).

Population genetic processes using the house mouse (*Mus musculus*) were studied in two projects (B1 and B2), which addressed the molecular mechanisms of adaptation in wild populations and the genetics of speciation. **Project B1** used microsatellites to detect selective

sweeps (Teschke et al., 2008) and analysed the divergence of gene expression between two house mouse subspecies. The expression data suggested a neutral model of gene expression change where only a fraction of the change may have been driven by positive selection (Staubach et al., 2010). In the course of this work, the emergence of a new gene in one of the subspecies was observed (Heinen et al., 2009). The results indicated that cryptic signals for transcript regulation and processing exist in intergenic regions and can become the basis for the evolution of new functional genes. Thus, new genes can arise without duplication or recombination of existing genes. This work became the starting point of a new research field devoted to the generation mechanisms of new genes (Neme et al., 2017; Tautz and Domazet-Loso, 2011).

Project B2 addressed the role of gene expression changes for the speciation in the house mouse. By using controlled crosses between subspecies and sampling of individuals from natural hybrid zones loci were identified which contribute to reproductive isolation (Rottscheidt and Harr, 2007; Voolstra et al., 2007). The work resulted in a productive research programme on the molecular mechanisms of speciation (Turner and Harr, 2014).

Project B4 was devoted to detecting loci involved in adaptation to new environments. The research question was related to project B1; however, project B4 focussed on human evolution using dense SNP arrays from four population samples including Polynesians, Melanesians, Southeast Asians and Fijians. Overall 32 genomic regions were identified with clear evidence for deviation from neutrality. A more detailed characterisation of these regions with the aim to identify single loci was not pursued for technical reasons. However, the analyses of this project produced an important contribution to our understanding of the demographic history of Oceania (Wollstein et al., 2010).

Two projects were devoted to interactions between species (B3 and B7), specifically studied host-pathogen interactions and aquatic food web interactions. **Project B3** undertook extensive cell biological studies to identify virulence proteins of the intracellular protozoal parasite, *Toxoplasma gondii*, and the host resistance proteins (IRG proteins) and to characterize the invasive strategy of the parasite at the molecular level (Fleckenstein et al. 2012, Steinfeldt et al. 2010). To understand the evolutionary dynamics of resistance and virulence, the genetic diversity of mouse IRG proteins and *Toxoplasma* virulence factors from infected and uninfected wild mice were analysed (Lilue et al., 2013). An extremely polymorphic gene cluster containing IRG proteins was found to confer resistance. By in vitro reconstruction experiments, the specific IRG protein responsible for differential resistance was identified, and a molecular mechanism of action was proposed and experimentally validated. This study provides the experimental and functional background for analysis of the forces driving reciprocal host-pathogen interactions in the wild.

The short **project B7** studied the planktonic microcrustacean *Daphnia magna*, a key member of aquatic food webs of lakes and ponds. *Daphnia* feeds on algae and cyanobacteria. The cyanobacteria, in turn, produce protease inhibitors, which target the digestive enzymes of *Daphnia*. The project analysed how the seasonal appearance of cyanobacterial protease inhibitors affected the frequency of protease alleles in *Daphnia*. The distribution of chymotrypsin alleles in a Swedish lake provided evidence for local adaptation (Schwarzenberger et al., 2013).

In projects B5 and B10, we used *Arabidopsis thaliana* as a model system to study the genetic basis of phenotypic variation within a species and how this contributes to adaptation to the environment. *Arabidopsis thaliana* is an excellent model system to study adaptation because of its recent expansion into a wide range of environments and the intensive sampling of accessions that has been performed. In **project B5**, we explored the basis of phenotypic variation for life-history traits, particularly genetic variation for the strength of seed dormancy and for flowering time, as these are key check-points in the life cycle of plants. Early work focused on describing

these phenotypic traits in almost two hundred accessions that provided a descriptive basis of variation in life history within the species and enabled subsequent genome-wide association approaches for the identification of underlying polymorphisms (Atwell *et al.*, 2010). We then focused on genetic variation affecting seed germination in natural populations, particularly the *DELAY OF GERMINATION 1* (*DOG1*) locus, which we had identified as a major contributor to variation in seed dormancy in nature (Bentsink *et al.*, 2006). Extensive sequence analysis of *DOG1* alleles present in four different populations led us to propose that selection shaped the geographic variation in *DOG1* sequence and function (Kronholm *et al.*, 2012). As this project developed, we used similar approaches to examine sequence variation at a gene encoding a microRNA, miR824 (de Meaux *et al.*, 2008). In natural populations, two alleles of *MIR824* exist that vary in the efficiency with which miR824 is processed to the mature form and in the effectiveness with which they downregulate expression of their target gene *AGL16*. We showed that AGL16 has an important role in controlling flowering time, and suggested therefore that variation in *MIR824* can contribute to this important adaptive trait (Hu *et al.*, 2014). Thus, we showed how alleles controlling major life history traits are selected for in natural environments.

In **project B10**, we used genetic variation within *Arabidopsis thaliana* to explore trade-offs between induced anti-microbial defences and growth. Intercrossing between *Arabidopsis thaliana* accessions can give rise to hybrid incompatability (HI) in which growth of the hybrids is impaired compared to the parents. These hybrid incompatibilities are often temperature dependent, being less extreme at higher temperatures. Many of these effects have proven to be due to negative epistasis between immune-related genes or allelic forms of these genes giving rise to autoimmunity (Chae *et al.*, 2014), and thus potentially form a basis for trade-offs between growth and disease resistance. We characterized a hybrid incompatibility combination between the North European (NE) *Arabidopsis thaliana* accession (Landsberg, L*er*) and the Central Asian accessions Kashmir (Kas-2) or Kondara (Kond). We identified the parental causal loci underlying immune-related, temperature-conditioned hybrid incompatibility (Alcazar *et al.*, 2010; Alcazar *et al.*, 2014). We visited and collected the natural Polish population of Landsberg and showed that the incompatible allele has persisted in the population along with other alleles at the locus. We also showed that the Kas-2/Kond hybrid incompatibility locus encoded the cell-surface receptor-like kinase gene *Strubbelig Receptor Family3*, *SRF3* (Alcazar *et al.*, 2010). *SRF3* incompatible alleles are widespread in *A. thaliana* Central Asian populations where there is a signature of a recent *SRF3* selective sweep (Alcazar *et al.*, 2010). We established that the L*er SRF3* 'compatible' allele confers lower levels of disease resistance than the Kas-2 'incompatible' *SRF3* allele in near-isogenic backgrounds. Therefore, the spread of incompatible *SRF3* allelic forms in Central Asia might be due to them conferring a local selective advantage, for example under conditions of high pathogen pressure or particular temperature regimes. Our data support the idea that selection for particular resistance gene alleles can create intercrossing barriers between natural populations, shaping genetic variation among plant populations.

Three projects (A9, A10, A13) exploited comparisons between *Arabidopsis thaliana* and its close relatives to study the generation of phenotypic novelties between plant species. These projects elaborated on the deep understanding of developmental processes in *Arabidopsis thaliana* that has emerged over the last twenty years and the more recent generation of genetic and genomic tools across the Brassicaceae family. In **project A13**, we studied the diversity in leaf shape among Brassicaceae species, particularly by using *Cardamine hirsuta* and *Arabidopsis thaliana* as models. In the simple leaf of *Arabidopsis thaliana* only small protrusions called serrations appear on the margin of the leaf, whereas in *Cardamine hirsuta* marginal outgrowths form distinct units that appear as small leaves called leaflets. We identified a Cardamine hirsuta mutant, *reduced complexity* (*rco*), that forms incompletely developed and poorly separated leaflets. We isolated

this gene and showed that it encoded a homeodomain protein (Vlad *et al.*, 2014). Strikingly, transfer of the *Cardamine hirsuta RCO* gene into *Arabidopsis thaliana* was sufficient to cause a lobed leaf form related to that of *Cardamine hirsuta*, indicating that this gene represents a key genetic component for diversification of form (Vlad *et al.*, 2014). We found that *RCO* is not present in the *Arabidopsis thaliana* genome but is a taxonomically restricted gene that underlies elaboration of a species-specific trait. We found that a paralogue of *RCO* called *LMI1* was retained in the *Arabidopsis thaliana* genome and used coding region and promoter swaps between these genes to demonstrate that the species-specific function of *RCO* is attributable to its distinct expression pattern at the leaf base. We constructed a *RCO* gene phylogeny to demonstrate that it represents an evolutionary novelty that arose after a gene duplication that gave rise to the RCO gene within the mustard family (Vlad *et al.*, 2014). *RCO* appears to improve the physiological performance of the plant, measured as increased gas exchange and seed biomass, without pleiotropic effects (Vuolo *et al.*, 2016). Our results therefore provided the first concrete link between causal genetic variation for leaf shape at a macro-evolutionary scale and plant fitness.

Project A10 also utilized phenotypic variation among *Arabidopsis thaliana* relatives to study the evolution of phenotypic novelties. *Arabidopsis thaliana* is a derived semelparous annual species that flowers and reproduces only once during its life time but then produces large numbers of progeny. By contrast, most species in the Brassicaceae are iteroparous perennials that flower and reproduce multiple times over many years, but produce fewer progeny each year (Karl *et al.*, 2012). Thus, in the iteroparous perennial life-cycle there is a trade-off between seed yield and longevity. We developed *Arabis alpina* as a model perennial species and used mutagenesis and inter-species crosses between this perennial and its sister annual species *Arabis montbretiana* to identify loci that control the different patterns of reproduction in these species (Wang *et al.*, 2009; Kiefer *et al.*, 2017). We exploited *Arabis alpina* mutants and genetic variation in natural populations to show that the PERPETUAL FLOWERING 1 (PEP1) MADS box transcription factor plays major roles in restricting flowering temporally and spatially in the perennial (Wang *et al.*, 2009; Albani *et al.*, 2012). Unlike the case of *RCO*, studied in project A13, orthologues of *PEP1* exist in semelparous annuals such as *Arabis montbretiana* and *Arabidopsis thaliana* where they are called *FLOWERING LOCUS C* (*FLC*), but their expression pattern is stably repressed by exposure to winter cold, so that these species flower abundantly and for longer duration. Using inter-species crosses and transgenic complementation we showed that when *FLC* is introduced into the perennial it shows the pattern of expression characteristic of annuals, indicating that the difference in transcriptional pattern between annuals and perennials is due to *cis*-acting variation (Kiefer *et al.*, 2017). Furthermore, the flowering patterns of these *Arabis alpina* plants carrying the annual *FLC* are those shown by the annual progenitor, supporting the critical role of genetic variation at this gene (Kiefer *et al.*, 2017). Also, we showed that only around 15% of the target genes bound by FLC and PEP1 in their respective host species are conserved but that these conserved genes include most of the known targets involved in flowering and reproductive development (Mateos *et al.*, 2017). We proposed that this transcription factor has deeply conserved roles in regulating reproduction in different species, but that during the divergence of the Brassicaceae family it has more recently acquired roles in stress responses that are specific to different lineages within the family (Mateos *et al.*, 2017). This project demonstrated how variation in expression of a single key regulator can contribute to major life-history transitions.

Project A9 also used comparisons between *Arabis alpina* and *Arabidopsis thaliana* to study the evolution of an adaptive process, in this case distribution of hairs on the leaf surface and related epidermal processes. Development and distribution of leaf hairs is controlled by a well-established regulatory network of proteins including several MYB and bHLH transcription factors as well as the WD40 protein TTG1 that collectively form so called MBW complexes (Balkunde *et*

al., 2010). A range of different complexes form due to gene duplications and redundancy between MYB and bHLH transcription factors, and the resulting MBW complexes contribute to five distinct epidermal traits that are collectively important in adaptation to the environment. We examined how redundantly acting genes encoding MBW components conferred different epidermal traits at the micro- and macro-evolutionary scales. At the micro-evolutionary scale, early work used association mapping among *Arabidopsis thaliana* accessions show correlations in trait variation, suggesting that these traits might be controlled by allelic variation at the same genes. Analysis of the closely related species *Arabis alpina* identified mutations with related phenotypes in most MWB components, it also described a patterning mechanism for leaf hairs that is not shown by *Arabidopsis thaliana* and identified a likely gene controlling this process (Chopra *et al.*, 2014). At the macro-evolutionary scale, the ability of MWB components from more distantly related species, cotton, petunia and maize, were tested for their capacity to complement *Arabidopsis thaliana* mutations. These experiments defined differences among MBW proteins involved in leaf hair patterning and pigment formation and detected competitive interactions among components conferring leaf hair patterning (Pesch *et al.*, 2015). These experiments suggest that this competition is an evolutionary protein innovation relevant for patterning. Thus, this project defined evolutionary constraints and innovations in an adaptively important regulatory process.

In **project A5**, we also studied the role of diversification of protein function in generating phenotypic novelties during plant evolution. We focused on the homeobox transcription factor WUSCHEL (WUS) that is required for the specification of the stem cells at the apex of the stem and give rise to all shoot tissues. In higher plants, an extensive family of genes encodes WUS-like proteins (WOX), so that in *Arabidopsis thaliana* these are encoded by 21 genes. WOX proteins have a monophyletic origin in green algae (Deveaux *et al.*, 2008), and we showed by phylogenetic reconstruction that WUS function is a derived trait, which appeared with leptosporangiate ferns, and that the ancestral proteins were related to WOX13 (Nardmann *et al.*, 2009; Nardmann & Werr, 2012). We found that the WUS clade of the family was amplified in higher plants to give rise to WOX2, WOX3, WOX4, WOX5 that serve individual stem cell niches in all seed plants, so for example WOX5 serves the root stem cells. We analysed the protein sequences of 760 WOX proteins by covariance and protein domain analysis and then ascertained the function of selected proteins based on their capacity to complement *wus* mutations of *Arabidopsis thaliana*. The complementation studies showed that despite their very different cellular expression patterns and long independent evolutionary trajectories, most gene products encoded in the *WUS*-clade remained interchangeable (Dolzblasz *et al.*, 2016). Also, the protein analyses demonstrated that WUS proteins differ from the ancestral WOX13 proteins by unique tri-peptide signatures at HD positions 47–49 at the exit from the turn into H3. These changes were adaptive as exchanging the sequences from the ancestral clade with WUS prevented complementation, and these positions probably affect folding of the homeodomain and influence interaction with the DNA minor groove. Furthermore, crystalizing the homeodomain of WUS revealed novel features relative to animal homeodomains. In this project, we have followed the evolutionary history of a family of plant transcription factors that contributed to the evolution of the higher plant body plan, and defined functionally important novelties in their amino acid sequence that enabled their diversification from ancestral proteins.

The early **project A4** was also concerned with deep plant evolution and its relationship to a particular family of transcriptional regulators. In Bryophytes, which represent the most primitive land plants, the gametophytic phase dominates the complete life cycle, whereas the gametophytes of higher plants, the male pollen and the female egg apparatus exhibit a simple morphology and are completely sporophyte-dependent. Thus, extreme reduction of the haploid phase of the life-cycle is a key innovation in seed plant evolution. Project A4 examined whether

this innovation is linked to the evolution of a particular class of MADS box transcription factors, MIKC*-type. However, cloning approaches in Bryophytes showed that MIKC*-type proteins are also present in their gametophytes, so they arose around 450 million years ago in land plants and are not specific to higher plants (Zobell *et al.*, 2010). As the original evolutionary hypothesis was disproven, the project was not continued, but it proved influential in defining a class of gametophyte-specific transcription factors that appeared early in plant evolution.

Project B11 has been concerned with the interplay genetic, antigenic, and morphological changes in colonies of pathogenic bacteria. Bacterial biofilms can generate micro-heterogeneity in terms of surface structures, but little is known about how these structurs affect cell-cell interactions and the architecture of biofilms. Using the type IV pilus of *Neisseria gonorrhoeae*, we showed that rupture forces between pili are fine-tuned by post-translational modification and impact bacterial sorting (*Oldewurtel et al 2015*). Active force generation was necessary for defined morphologies of mixed microcolonies. The observed morphotypes were consistent with a tug-of-war among surface structures of different cells governing cell sorting. We tested the fitness effects of phase and antigenic variation that arise from cell positioning in expanding gonococcal colonies. We found we that loss of pili and subsequent segregation to the front confers a strong selective advantage (*Yüksel et al. 2016*). We sequenced the major pilin gene of spatially segregated sub-populations, and we investigated the spatio-temporal population dynamics. Our findings indicate that pilin phase and antigenic variation generate a standing variation of pilin sequences within the inoculation zone, while variants associated with a non-piliated phenotype segregate to the front of the growing colony (*Zöllner et al. 2017*). We conclude that tuning of attractive forces by phase and antigenic variation is a powerful mechanism governing the dynamics of bacterial colonies.

The associated **project B13** extended controlled experimental evolution of *E. coli* to realistic ecological environments, which is a major step in understanding the impact of such experiments on evolution in the wild (*Lourenço et al 2016*, *Sousa et al. 2017*). Specifically, we found that adaptive evolution of *E. coli* in the mouse gut can be highly repeatable across distinct hosts. This process has strong clonal interference, sometimes together with frequency dependent selection, and some targets of *E. coli* genetic adaptation depend on the individual microbiota composition of each mouse. For the first time, we inferred selection coefficients of *de novo* adaptive mutations within a commensal species of the human microbiota. Recently, using a new colonization model for E. coli in the mouse gut, we found rapid evolution by horizontal gene transfer; this mode precedes and outweighs evolution by point mutations. In collaboration with project C2, we found that these data can consistently be captured by a minimal model of eco-evolutionary change (Frazao A, Sousa A, Lässig M*, Gordo I*, submitted).

Project B12 focused on the the adaptive significance of multiple small-effect mutations that change a molecular quantitative trait. Using high quality protein and RNA data of quantitative trait loci (QTL) in yeast, we developed a simple, yet powerful approach to detect positive selection on molecular traits (*Clement-Ziza et al. 2014*). For the first time, we could show using QTL data that protein levels are under stronger evolutionary selection than transcript levels. In collaboration with the de Meaux lab, we applied QTL inference to transcriptome data from the Arabidopsis genus. We demonstrated that cis-regulatory novelties preferentially contribute to increase basal gene expression levels, whereas increased down-regulation is predominantly provided by trans-acting changes. Furthermore, by contrasting rates of accumulation of *cis*-acting changes in each lineage, we inferred the relative impact of natural selection (*He et al. 2016*).

Fitness effects of quantitative traits were also addressed in the collaborative **project C1**, which focused on the inference of selection on metabolism and gene expression. In early work, we developed evolutionary models of gene and protein interaction networks, including a comparative

method called *graph alignment* aimed at displaying the functional evolution of networks (*Berg and Lässig 2006*). The problem of identifying statistically significant clusters in such networks let to theoretical work on clustering that uncovered deep links to statistical physics (*Luksza et al. 2010*). A combined experimental and theoretical analysis of the lac utilization pathway in *Escherichia coli* produced one of the first quantitative phenotype-fitness maps based on a biophysical model of the lac metabolic pathway (*Perfeito et al. 2011*). This map displays fitness nonlinearities and broad epistasis, which remain important topics of microbial evolution to date. The inference of selection on gene expression triggered a broad research programme in quantitative genetics of adaptive evolution. In particular, we developed evolutionary models for quantitative traits that synthesize directional and stabilizing selection into a joint conceptual framework (*Held et al. 2014*). This theoretical work laid the ground for a recent study of gene expression evolution in flies (*Nourmohammad et al. 2017*). Addressing a long-standing discussion on the role of adaptation in gene expression, this work showed that the majority of macro-evolutionary expression changes throughout the *Drosophila* genome are adaptive.

A parallel development, in **project C2**, was a broad enquiry into the biophysical determinants of molecular evolution. Early work focused on evolutionary models for transcriptional interactions. We showed that fitness landscapes based on binding affinity data quantitatively capture the evolutionary turnover of transcription factor binding sites across neighboring species of yeast (*Mustonen et al. 2008*). To disect adaptive evolution, we developed population genetics models with an explicit dependence of selection on evolutionary time. The basic idea is to associate adaptation with the non-equilibrium part and compensatory changes with the equilibrium part of a stochastic evolutionary process. Using these models, we quantified pace and selective strength of adaptive evolution in the *Drosophila* genome (*Mustonen and Lässig 2007*). We established fitness flux as a new measure of adaptation; this quantity satisfies a fluctuation theorem that provides far-reaching links to modern non-equilibrium statistical physics (*Mustonen and Lässig 2010*). How strong selection pressure plays out under in adaptive processes of fast-evolving microbial and viral populations was a major topic in recent years. We showed that clonal interference leads to emergent neutrality for moderately selected mutations (*Schiffels et al. 2011*), and we established clonal competition between clades in the human influenza virus (*Strelkowa and Lässig 2012*). This has been the first inference of this mode of evolution in a natural population in the wild. It led to a new question: can we predict which of these clades will win the competition, based on past data? We showed that a fitness model for influenza can predict its evolution over periods of one year (*Luksza and Lässig, 2014*), which led to a new method for influenza vaccine strain selection (*Morris et al. 2017*) and integrative fitness models for fast-evolving systems (*Lässig et al. 2017).*

The collaborative **project C3** was focused on empirical fitness landscapes in microbial systems and their implications for adaptive evolution experiments. In the past two funding periods of the SFB680, we have combined experimental and theoretical work to study how the topography of the fitness landscape and population dynamic parameters, such as population size and recombination rate, affect the dynamics and repeatability of evolution. The experimental work involved the characterization of small-scale fitness landscapes of the fungi *Aspergillus niger* and *A. nidulans* and the bacterial antibiotic-resistance enzyme TEM-1 β-lactamase. We also characterized the distribution of beneficial mutation effects and their fixation probability of TEM-1 β-lactamase and showed that it has a heavy tail (*Schenk et al. 2013*). Using the 'rugged' empirical fitness landscapes of *A. niger*, we showed that recombination has mostly deleterious adaptive effects (*de Visser et al. 2009*). We also quantified the accessibility of mutational pathways as a function of the distance to the global optimum (*Franke et al. 2011*), and showed that evolutionary predictability depends non-monotonically on population size on this rugged landscape (*Szendro*

et al. 2013). Using single and double beneficial mutants in *A. nidulans*, we showed that Fisher's geometric model accurately describes the pattern of diminishing-returns epistasis among these mutations (*Schoustra et al. 2016*). Based on the library of beneficial mutations in TEM-1 β-lactamase, we constructed and analyzed several small-scale fitness landscapes, and identified methodological biases in their topography (*Schenk et al. 2013, de Visser and Krug 2014*). We also subjected the enzyme to *in vitro* evolution, showing subtle pleiotropic constraints in the presence of two related antibiotics (*Schenk et al. 2015*), the important role of large-effect driver mutations during selection of mutant libraries (*van Dijk et al. 2017*), and adaptive benefits from small mutation supplies due to strong epistatic constraints (*Salverda et al. 2017*).

Complementary methods for the inference of evolutionary forces micro- and macro-evolutionary scales were the subject of **project C4**. In early work, we formulated and investigated population genetic models to study the imprint of positive selection on patterns of genetic diversity in natural populations. We applied theoretical models to analyze experimental data from natural populations of mouse and fly, resulting in joint publications with experimentalists (*Wiehe et al. 2007, Teschke et al. 2008*). In follow-up studies, key aspects were to contrast and distinguish genetic from demographic causes in shaping the patterns of genetic diversity and to investigate the effect of selection on the haplotype structure of genomes. Starting in the second funding period, we broadened our range of research topics and included mechanisms at macro-evolutionary time scales. We investigated the evolutionary history of so-called chromatin insulator proteins, which are fundamental for defining chromosome architecture and delimiting units of expression (*Heger et al. 2012*). One of these, the DNA binding factor CTCF, is strongly conserved among bilaterian animals and an essential interactor of the developmentally important Hox genes involved in body-plan formation. This insight led us to ask for a more comprehensive catalogue of lineage-specific genes that contribute to the genotypic basis of morphological characters. We have compiled a database of such orthogroup clusters. This list contains functionally very interesting genes, for instance the gene "lefty", which is involved in defining left/right symmetry in bilateria.

Project C7 established an important link between evolution and systems biology. It addressed a central challenge: to explain evolutionary pathways for the formation of complex metabolic innovations that require multiple mutations (*Heckmann et al. 2013*). We proposed that metabolic innovations accessible through the addition of a single reaction serve as stepping stones towards the later establishment of complex metabolic features in another environment, and supported this hypothesis through computer simulations of metabolic network evolution and comparative genomic analyses (*Szappanos et al. 2016*). Our simulations indicate a positive feedback loop between complexity and evolvability: large metabolic networks can integrate new pathways by natural selection more easily, and thereby they promote a further increase of metabolic complexity. We linked observed genotypic and phenotypic evolution on a genome-wide, macro-evolutionary scale by reconstructing the metabolic systems of ancestral *E. coli* strains and identifying 3,323 phenotypic innovations in the history of this clade. Strikingly, every single observed innovation could be traced back to the acquisition of a single DNA segment (*Pang et al. 2018)*. While we found no evidence for the contribution of selectively neutral processes, 10.6% of *E. coli* adaptations to previously unviable environments relied on the support of DNA acquisitions on earlier phylogenetic branches, consistent with the stepwise adaptation to successive environments we observed as a dominant process in our forward simulations.

Eukaryotic genome evolution by repetitive elements was the subject of **project C9**, which entered the SFB in the third funding period. In order to understand how Alu elements re-integrate into the host genome, we analyzed the chromatin structure at all transcribed units of the genome. We developed a hidden Markov model (HMM) that integrates chromatin immunoprecipitation data of general transcription factors and histone modification data, in order to segment the genome into

functionally distinct regions. Our model is publicly available as a Bioconductor software package (STAN, *Zacher et al. 2014*). We included DNA methylation into our later analysis, which was also published on Bioconductor (BEAT, *Akman et al. 2014*). We found that Alu transcripts are transcribed autonomously. They are not splicing side-products of intronic regions, nor do they originate from other gene-related transcription. Moreover, we discovered that Alu synthesis rates are negatively correlated with their age. Evolutionary old Alu families that lost their capacity to re-integrate are generally transcribed at higher rates, while the late Alu-Y families are actively suppressed. The comparison of RNA Polymerase II and RNA Polymerase III ChIP-Seq data surprisingly provided evidence that many Alus could also be Pol II targets. Contrary to common belief, we also found that most Alu transcripts are extraordinarily stable, compared to other non-coding transcripts. A sequence-structure-function association analysis revealed several highly influential loci that when mutated have a strong effect on Alu synthesis rates.

References

Akman K, Haaf T, Gravina S, Vijg J, **Tresch A** (2014). Genome-wide quantitative analysis of DNA methylation from bisulfite sequencing data. Bioinformatics, 30(13).

Albani MC, Castaings L, Wotzel S, Mateos JL, Wunder J, Wang R, **Reymond M**, **Coupland G.** 2012. PEP1 of Arabis alpina is encoded by two overlapping genes that contribute to natural genetic variation in perennial flowering. PLoS Genet 8(12): e1003130.

Alcazar R, Garcia AV, Kronholm I, **de Meaux J**, **Koornneef M**, **Parker JE**, **Reymond M.** (2010). Natural variation at Strubbelig Receptor Kinase 3 drives immune-triggered incompatibilities between Arabidopsis thaliana accessions. Nat Genet 42(12): 1135-1139.

Alcazar R, von Reth M, Bautor J, Chae E, Weigel D, **Koornneef M**, **Parker JE.** (2014). Analysis of a plant complex resistance gene locus underlying immune-related hybrid incompatibility and its occurrence in nature. PLoS Genet 10(12): e1004848.

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, et al. (2010). Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465(7298): 627-631.

Balkunde R, Pesch M, **Hulskamp M.** (2010). Trichome Patterning in Arabidopsis Thaliana: From Genetic to Molecular Models. Plant Development 91: 299-321.

Bentsink L, Jowett J, Hanhart CJ, **Koornneef M.** (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. Proc Natl Acad Sci U S A 103(45): 17042-17047.

Berg J, **Lässig M** 2006. Cross-species analysis of biological networks by Bayesian alignment. PNAS 103:10967-10972.

Chae E, Bomblies K, Kim ST, Karelina D, Zaidem M, Ossowski S, Martin-Pizarro C, Laitinen RA, Rowan BA, Tenenboim H, et al. (2014). Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. Cell 159(6): 1341-1351.

Chopra D, Wolff H, Span J, Schellmann S, **Coupland G**, Albani MC, Schrader A, **Hulskamp M.** (2014). Analysis of TTG1 function in Arabis alpina. Bmc Plant Biology 14.

Clement-Ziza, M., …, **Beyer A** (2014). Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast. Mol. Syst. Biol. 10, 764–764.

Couso, J.P., and Patraquim, P. (2017). Classification and function of small open reading frames. Nat Rev Mol Cell Biol 18, 575-589.

de Meaux J, Hu JY, Tartler U, Goebel U. (2008). Structurally different alleles of the ath-MIR824 microRNA precursor are maintained at high frequency in Arabidopsis thaliana. Proc Natl Acad Sci U S A 105(26): 8994-8999.

Deveaux Y, Toffano-Nioche C, Claisse G, Thareau V, Morin H, Laufs P, Moreau H, Kreis M, Lecharny A. (2008). Genes of the most conserved WOX clade in plants affect root and flower development in Arabidopsis. Bmc Evolutionary Biology 8.

Dolzblasz A, Nardmann J, Clerici E, Causier B, van der Graaff E, Chen JH, Davies B, **Werr W**, Laux T. (2016). Stem Cell Regulation by Arabidopsis WOX Genes. Molecular Plant 9(7): 1028-1039.

Failmezger H, Jaegle B, Schrader A, **Hülskamp M**, **Tresch A.** (2013). Semi-automated 3D Leaf Reconstruction and Analysis of Trichome Patterning from Light Microscopic Images. Plos Computational Biology 9(4).

Fleckenstein, M.C., Reese, M.L., Konen-Waisman, S., Boothroyd, J.C., **Howard, J.C.**, and Steinfeldt, T. (2012). A Toxoplasma gondii pseudokinase inhibits host IRG resistance proteins. PLoS Biol 10, e1001358.

Galindo, M.I., Pueyo, J.I., Fouix, S., Bishop, S.A., and Couso, J.P. (2007). Peptides encoded by short ORFs control development and define a new eukaryotic gene family. PLoS Biol 5, e106.

He F, Arce A, Schmitz G, Rico A., **Beyer A**, **de Meaux J** (2016). The Footprint of Polygenic Adaptation on Stress-Responsive Cis-Regulatory Divergence in the Arabidopsis Genus. Mol Biol Evol. 33(8):2088-101.

Heckmann D, Schulze S, Denton A, Gowik U, Westhoff P, Weber APM, **Lercher MJ** (2013). Predicting C4 Photosynthesis Evolution: Modular, Individually Adaptive Steps on a Mount Fuji Fitness Landscape. Cell 153:1579-1588.

Heger P, Marin B, Bartkuhn M, Schierenberg E and **Wiehe T** (2012). The chromatin insulator CTCF and the emergence of metazoan diversity. PNAS 109:17507.

Heger, P., Kroiher, M., Ndifon, N., **Schierenberg, E**. (2010). Conservation of MAP kinase activity and MSP genes in parthenogenetic nematodes. BMC Dev Biol 10, 51.

Heger, P., Marin, B., Bartkuhn, M., **Schierenberg, E.**, **Wiehe, T.** (2012). The chromatin insulator CTCF and the emergence of metazoan diversity. Proc Natl Acad Sci U S A 109, 17507-17512.

Heger, P., Marin, B., **Schierenberg, E.** (2009). Loss of the insulator protein CTCF during nematode evolution. BMC Mol Biol 10, 84.

Heinen, T.J., Staubach, F., Haming, D., **Tautz, D.** (2009). Emergence of a new gene from an intergenic region. Curr Biol 19, 1527-1531.

Held T, Nourmohammad A, **Lässig M** (2014). Adaptive evolution of molecular phenotypes. J. Stat. Mech., P09029 (39 pages).

Hu JY, Zhou Y, He F, Dong X, Liu LY, **Coupland G**, Turck F, **de Meaux J.** (2014). miR824-Regulated AGAMOUS-LIKE16 Contributes to Flowering Time Repression in Arabidopsis. Plant Cell 26(5): 2024-2037.

Karl R, Kiefer C, Ansell SW, Koch MA. (2012). Systematics and evolution of Arctic-Alpine Arabis alpina (Brassicaceae) and its closest relatives in the eastern Mediterranean. American Journal of Botany 99(4): 778-794.

Kiefer C, Severing E, Karl R, Bergonzi S, Koch M, **Tresch A**, **Coupland G.** (2017). Divergence of annual and perennial species in the Brassicaceae and the contribution of cis-acting variation at FLC orthologues. Mol Ecol.

Kronholm I, Pico FX, Alonso-Blanco C, Goudet J, **de Meaux J.** (2012). Genetic basis of adaptation in Arabidopsis thaliana: local adaptation at the seed dormancy QTL DOG1. Evolution 66(7): 2287-2302.

Lässig M., Mustonen V., Walczak A. (2017). Predicting evolution. Nature Ecol. Evol. 1:0077.

Lilue, J., Muller, U.B., Steinfeldt, T., **Howard, J.C.** (2013). Reciprocal virulence and resistance polymorphism in the relationship between Toxoplasma gondii and the house mouse. Elife 2, e01298.

Lourenço M, Ramiro RS, Güleresi D, Barroso-Batista J, Xavier KB, I **Gordo I**, Sousa A (2016). A mutational hotspot and strong selection contribute to the order of mutations selected for during Escherichia coli adaptation to the gut. PLoS Genetics 12 (11), e1006420.

Luksza M, **Lässig M** (2014). A predictive fitness model for influenza. Nature 507:57-61.

Łuksza M, **Lässig M**, **Berg J** (2010). Significance analysis and statistical mechanics: an application to clustering. Phys. Rev. Lett. 105:220601.

Mateos JL, Tilmes V, Madrigal P, Severing E, Richter R, Rijkenberg CWM, Krajewski P, **Coupland G.** (2017). Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species. Proceedings of the National Academy of Sciences of the United States of America 114(51): E11037-E11046.

Morris et al. 2017 Morris D.H., Gostic K.M., Pompei S., Bedford T., Łuksza, M., Neher, R.A., Grenfell, B.T., **Lässig, M.**, McCauley J.W. (2017). Predictive Modeling of Influenza Shows the Promise of Applied Evolutionary Biology. Trends in Microbiology, 26:101-118.

Mustonen V, Kinney J, Callan CG Jr, **Lässig M** (2008). Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. Proc Natl Acad Sci. 105, 12376-81.

Mustonen V, **Lässig M** (2007). Adaptation to fluctuating selection in *Drosophila*. Proc Natl Acad Sci 104: 2277 - 82.

Mustonen V, **Lässig M** (2010). Fitness flux and ubiquity of adaptive evolution, Proc. Natl. Acad. Sci. 107:4248-53.

Nardmann J, Reisewitz P, **Werr W.** (2009). Discrete Shoot and Root Stem Cell-Promoting WUS/WOX5 Functions Are an Evolutionary Innovation of Angiosperms. Molecular Biology and Evolution 26(8): 1745-1755.

Nardmann J, **Werr W.** (2012). The invention of WUS-like stem cell-promoting functions in plants predates leptosporangiate ferns. Plant Molecular Biology 78(1-2): 123-134.

Neme, R., Amador, C., Yildirim, B., McConnell, E., **Tautz, D.** (2017). Random sequences are an abundant source of bioactive RNAs or peptides. Nat Ecol Evol 1, 0217.

Nourmohammad A, Rambeau J, Held T, Kovacova V, **Berg J**, **Lässig M** (2017). Pervasive adaptation of gene expression in Drosophila, Cell Reports 20 (6) 1385-1395.

Nunes da Fonseca, R., von Levetzow, C., Kalscheuer, P., Basal, A., van der Zee, M., **Roth, S.** (2008). Self-regulatory circuits in dorsoventral axis formation of the short-germ beetle Tribolium castaneum. Dev Cell 14, 605-615.

Oldewurtel, E.R., Kouzel, N., Dewenter, L., Henseler, K., **Maier, B.** (2015). Differential interaction forces govern bacterial sorting in early biofilms, eLife, DOI 10.7554/eLife.1081.

Ozuak, O., Buchta, T., **Roth, S.**, **Lynch, J.A.** (2014). Dorsoventral polarity of the Nasonia embryo primarily relies on a BMP gradient formed without input from Toll. Curr Biol 24, 2393-2398.

Pang TY, **Lercher MJ** (2017). The adaptive acquisition of single DNA segments drives metabolic evolution across E. coli lineages. BioRxiv doi: https://doi.org/10.1101/148460.

Perfeito L, Ghozzi S, **Berg J**, Schnetz K, **Lässig M** (2011). Nonlinear fitness landscape of a molecular pathway. PLoS Genetics 7(7): e1002160.

Pesch M, Schultheiss I, Klopffleisch K, Uhrig JF, Koegl M, Clemen CS, Simon R, Weidtkamp-Peters S, **Hülskamp M** (2015). TRANSPARENT TESTA GLABRA1 and GLABRA1 Compete for Binding to GLABRA3 in Arabidopsis. Plant Physiology 168(2): 584-+.

Rottscheidt, R., **Harr, B.** (2007). Extensive additivity of gene expression differentiates subspecies of the house mouse. Genetics 177, 1553-1567.

Sachs, L., Chen, Y.T., Drechsler, A., Lynch, J.A., **Panfilio, K.A.**, **Lässig, M.**, **Berg, J.**, **Roth, S.** (2015). Dynamic BMP signaling polarized by Toll patterns the dorsoventral axis in a hemimetabolous insect. Elife 4, e05502.

Savard, J., Marques-Souza, H., Aranda, M., **Tautz, D.** (2006). A segmentation gene in tribolium produces a polycistronic mRNA that codes for multiple conserved peptides. Cell 126, 559-569.

Schiffels S, Szöllösi G, Mustonen V, **Lässig M** (2011). Emergent Neutrality in Adaptive Asexual Evolution. Genetics 189:1361 – 75.

Schwarzenberger, A., D´Hondt, S., Vyverman, W., **von Elert, E.** (2013). Seasonal succession of cyanobacterial protease inhibitors and Daphnia magna genotypes in a eutrophic Swedish lake. Aquatic Sciences 75, 433-445.

Sousa A, Frazão N, Ramiro RS, **Gordo I** (2017). Evolution of commensal bacteria in the intestinal tract of mice. Current Opinion in Microbiology 38, 114-121.

Staubach, F., Teschke, M., Voolstra, C.R., Wolf, J.B., **Tautz, D.** (2010). A test of the neutral model of expression change in natural populations of house mouse subspecies. Evolution 64, 549-560.

Steinfeldt, T., Konen-Waisman, S., Tong, L., Pawlowski, N., Lamkemeyer, T., Sibley, L.D., Hunn, J.P., **Howard, J.C.** (2010). Phosphorylation of mouse immunity-related GTPase (IRG) resistance proteins is an evasion strategy for virulent Toxoplasma gondii. PLoS Biol 8, e1000576.

Strelkowa N, **Lässig M** (2012). Clonal interference in the evolution of influenza. Genetics 192:671 - 682.

Szappanos B, Fritzemeier C J, Csörgő B, Lázár V, Lu X, Fekete G, Bálint B, Herczeg R, Nag I, Notebaart R, **Lercher M J**, Pál C, Papp B (2016). Adaptive evolution of complex innovations through step-wise metabolic niche expansion. Nature Communications 7:11607.

Tautz, D., Domazet-Loso, T. (2011). The evolutionary origin of orphan genes. Nat Rev Genet 12, 692-702.

Teschke M., O. Mukabayire, **T. Wiehe**, **D. Tautz** (2008). Identification of selective sweeps in closely related populations of the house mouse based on microsatellite scans. Genetics: 180:1537.

Teschke, M., Mukabayire, O., **Wiehe, T.**, **Tautz, D.** (2008). Identification of selective sweeps in closely related populations of the house mouse based on microsatellite scans. Genetics 180, 1537-1545.

Turner, L.M., **Harr, B.** (2014). Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. Elife 3.

Vlad D, Kierzkowski D, Rast MI, Vuolo F, Dello Ioio R, Galinha C, Gan XC, Hajheidari M, Hay A, Smith RS, et al. (2014). Leaf Shape Evolution Through Duplication, Regulatory Diversification, and Loss of a Homeobox Gene. Science 343(6172): 780-783.

Voolstra, C., **Tautz, D.**, Farbrother, P., Eichinger, L., **Harr, B.** (2007). Contrasting evolution of expression differences in the testis between species and subspecies of the house mouse. Genome Res 17, 42-49.

Vuolo F, Mentink RA, Hajheidari M, Bailey CD, Filatov DA, **Tsiantis M** (2016). Coupled enhancer and coding sequence evolution of a homeobox gene shaped leaf diversity. Genes & Development 30(21): 2370-2375.

Wang R, Farrona S, Vincent C, Joecker A, Schoof H, Turck F, Alonso-Blanco C, **Coupland G**, Albani MC. (2009). PEP1 regulates perennial flowering in Arabis alpina. Nature 459(7245): 423-427.

Wiehe T., V. Nolte, D. Zivkovic and C. Schlötterer (2007). Identification of selective sweeps using a dynamically adjusted number of linked microsatellites. Genetics 175:207.

Wollstein, A., Lao, O., Becker, C., Brauer, S., Trent, R.J., **Nürnberg, P.**, Stoneking, M., **Kayser, M.** (2010). Demographic history of Oceania inferred from genome-wide data. Curr Biol 20, 1983-1992.

Yüksel, M., Power, J.J., Ribbe, J., Volkmann, T., **Maier, B.** (2016) Fitness trade-offs in competence differentiation of *Bacillus subtilis*, Front. Microbiol. 7, 888.

Zacher, B., Lidschreiber, M., Cramer, P., Gagneur, J, **Tresch A.** (2014). Annotation of genomics data using bidirectional hidden Markov models unveils variations in Pol II transcription cycle. Molecular Systems Biology 10, 768, doi:10.15252/msb.20145654.

Zobell O, Faigl W, Saedler H, **Münster T.** (2010). MIKC* MADS-Box Proteins: Conserved Regulators of the Gametophytic Generation of Land Plants. Molecular Biology and Evolution 27(5): 1201-1211.

Zöllner, R.*, Oldewurtel, E.R.*, Kouzel, N., **Maier, B.** (2017) Phase and antigenic variation govern competition dynamics through positioning in bacterial colonies, Sci. Rep. 12151.

2. Core publications

1. **Berg J., Lässig M.** (2006). Cross-species analysis of biological networks by Bayesian alignment. Proc. Natl. Acad. Sci. 103:10967-10972.
2. **Harr B**. (2006). Genomic islands of differentiation between house mouse subspecies. Genome Res 16: 730-737.
3. Savard J., Marques-Souza H., Aranda M., **Tautz D** (2006). A segmentation gene in Tribolium produces a polycistronic mRNA that codes for multiple conserved peptides. Cell 126, 559 - 569.
4. Keurentjes J.J.B., Fu J., de Vos C.H., Lommen A., Hall R.D., Bino R.J., van der Plas L.H., Jansen R.C., Vreugdenhil D., **Koornneef M.** (2006). The genetics of plant metabolism. Nature Genet. 38: 842-849.
5. Park S.C.**, Krug J.** (2007). Clonal interference in large populations. Proc. Natl. Acad. Sci. 104: 18135-18140.
6. Mustonen V., Kinney J., Callan C.G. Jr., **Lässig M.** (2008). Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. Proc Natl Acad Sci. 105, 12376-81.
7. Nunes da Fonseca R., von Levetzow C., Kalscheuer P., Basal A., van der Zee M., **Roth S.** (2008). Self-regulatory circuits in dorsoventral axis formation of the short-germ beetle Tribolium castaneum. Dev Cell 14, 605-615.
8. Heinen T.J., Staubach F., Häming D., **Tautz D.** (2009). Emergence of a new gene from an intergenic region. Curr Biol. 19:1527-31.
9. Mustonen V., **Lässig M.** (2010). Fitness flux and ubiquity of adaptive evolution, Proc. Natl. Acad. Sci. 107:4248-53.
10. Steinfeldt T., Könen-Waisman S., Tong L., Pawlowski N., Lamkemeyer T., Sibley L.D., Hunn J.P., **Howard J.C.** (2010). Phosphorylation of IRG resistance proteins is an evasion strategy for virulent T. gondii strains. PLoS Biology 8: e1000576.
11. Alcázar R., García A.V., Kronholm I., **de Meaux J., Koornneef M., Parker J.E.**, **Reymond M.** (2010). Natural variation at Strubbelig Receptor Kinase 3 drives immune-triggered incompatibilities between A. thaliana accessions. Nature Genet. 42:1135-1139.
12. Franke J., Klözer A., **de Visser J.A.G.M., Krug J.** (2011). Evolutionary accessibility of mutational pathways. PLoS Comp. Biol. 7:e1002134.
13. Perfeito L., Ghozzi S., **Berg J.**, Schnetz K., **Lässig M.** (2011). Nonlinear fitness landscape of a molecular pathway. PLoS Genetics 7(7): e1002160.
14. Alcázar R., **Reymond M.**, Schmitz G., **de Meaux J.** (2011). Genetic and evolutionary perspectives on the interplay between plant immunity and development. Curr. Opin. Plant Biol. 14:378-384.
15. Andrés F., **Coupland G.** (2012). The genetic basis of flowering responses to seasonal cues. Nature Reviews Genetics 13:627.
16. Heger P., Marin B., Bartkuhn M., Schierenberg E., **Wiehe T.** (2012). The chromatin insulator CTCF and the emergence of metazoan diversity. Proc. Natl. Acad. Sci. 109:17507.
17. Lilue J., Mueller U.B., Steinfeldt T., **Howard J.C.** (2013). Toxoplasma gondii and the mouse; reciprocal virulence and resistance polymorphism. eLIFE 10.7554/eLife.01298.

18. Nardmann J., **Werr W.** (2013). Symplesiomorphies in the WUSCHEL clade suggest that the last common ancestor of seed plants contained at least four independent stem cell niches. New Phytol. 199:1081-1092.

19. Heckmann D., Schulze S., Denton A., Gowik U., Westhoff P., Weber A.P.M., **Lercher M.J.** (2013). Predicting C4 Photosynthesis Evolution: Modular, Individually Adaptive Steps on a Mount Fuji Fitness Landscape. Cell 153:1579-1588.

20. Szendro I.G., Franke J., **de Visser J.A.G.M., Krug J.** (2013). Predictability of evolution depends nonmonotonically on population size. Proc. Natl. Acad. Sci. USA 110:571-576.

21. **De Visser J.A.G.M., Krug J.** (2014). Empirical fitness landscapes and the predictability of evolution. Nat. Rev. Genet. 15:480-490.

22. Vlad D., Kierzkowski D., Rast M.I., Vuolo F., Dello Ioio R., Galinha C., Gan X, Hajheidari M., Hay A., Smith R.S., Huijser P., Bailey C.D., **Tsiantis M.** (2014). Leaf shape evolution through duplication, regulatory diversification, and loss of a homeobox gene. Science. 343:780-3.

23. Luksza M., **Lässig M.** (2014). A predictive fitness model for influenza. Nature 507:57-61.

24. Chopra D., Wolff H., Span J., Schellmann S., Coupland G., Albani M.C., Schrader A., **Hülskamp M.** (2014). Analysis of TTG1 function in Arabis alpina. BMC Plant Biol 14, 16.

25. Özuak O., Buchta T., **Roth S., Lynch J.A.** (2014). Dorsoventral polarity of the Nasonia embryo primarily relies on a BMP gradient formed without input from Toll. Curr Biol 24, 2393-2398.

26. Alcazar R., von Reth M., Bautor J., Chae E., Weigel D., **Koornneef M., Parker J.E.** (2014). Analysis of a plant complex resistance gene locus underlying immune-related hybrid incompatibility and its occurrence in nature. PloS Genet. 10, e1004848.

27. Clement-Ziza M., …, **Beyer A.** (2014). Natural genetic variation impacts expression levels of coding, non-coding, and antisense transcripts in fission yeast. Mol. Syst. Biol. 10, 764–764.

28. Oldewurtel E.R., Kouzel N., Dewenter L., Henseler K., **Maier B.** (2015). Differential interaction forces govern bacterial sorting in early biofilms, eLife, DOI 10.7554/eLife.10811.

29. Sachs L., Chen Y.T., Drechsler A., **Lynch J.A., Panfilio K.A., Lässig M., Berg J., Roth S.** (2015). Dynamic BMP signalling polarized by Toll patterns the dorsoventral axis in a hemimetabolous insect. Elife 4, e05502.

30. Szappanos B., Fritzemeier C.J., Csörgő B., Lázár V., Lu X., Fekete G., Bálint B., Herczeg R., Nag I., Notebaart R.**, Lercher M.J.**, Pál C., Papp B. (2016). Adaptive evolution of complex innovations through step-wise metabolic niche expansion. Nature Communications 7:11607.

31. Müller U.B., **Howard J.C.** (2016). The impact of Toxoplasma gondii on the mammalian genome. Current Opinion in Microbiology. 32:19-25.

32. Vuolo F., Mentink R.A., Hajheidari M., Bailey C.D., Filatov D.A., **Tsiantis M.** (2016). Coupled enhancer and coding sequence evolution of a homeobox gene shaped leaf diversity. Genes Dev. 1;30:2370-2375.

33. Liu Y., **Beyer A.,** Aebersold R. (2016) On the dependency of cellular protein levels on mRNA abundance. (Review) Cell 165:535-50.

34. Yüksel M., Power J.J., Ribbe J., Volkmann T., **Maier B.** (2016) Fitness trade-offs in competence differentiation of Bacillus subtilis, Front. Microbiol. 7, 888.

35. **Lässig M.**, Mustonen V., Walczak A. (2017). Predicting evolution. Nature Ecol. Evol. 1:0077.

36. Nourmohammad A., Rambeau J., Held T., Kovacova V., **Berg J.**, **Lässig M.** (2017). Pervasive adaptation of gene expression in Drosophila, Cell Reports 20 (6) 1385-1395.

37. Mateos J.L., Tilmes V., Madrigal P., Severing E., Richter R., Rijkenberg C.W.M., Krajewski P. & **Coupland G.** (2017). Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species. Proc. Natl. Acad. Sci. USA **114**:E11037-E11046.

38. Sousa A., Frazão N., Ramiro R.S., **Gordo I.** (2017). Evolution of commensal bacteria in the intestinal tract of mice. Current Opinion in Microbiology 38, 114-121.

39. Kiefer C., Severing E., Karl R., Bergonzi S., Koch M., **Tresch A., Coupland G.** (2017). Divergence of annual and perennial species in the Brassicaceae and the contribution of cis-acting variation at FLC orthologues. Molecular Ecology 26:3437-3457. doi:10.1111/mec.14084.

40. **Panfilio K.A.**, Angelini D.R. (2018). By land, air, and sea: hemipteran diversity through the genomic lens. Current Opinion in Insect Science 25:106-115 (invited review).

3. List of scientific projects

Project	Titel	Research area	Principal investigator(s) affiliation	Funding period
Project area A: Evolution between species				
A1	The evolution of signalling pathways and regulatory networks involved in dorsoventral patterning in insects	Evolution of development, dorsoventral patterning	**Siegfried Roth,** Institut für Entwicklungsbiologie, Universität zu Köln	2006-2009
A1	The evolution of gene regulatory networks for dorsoventral patterning in insects	Evolution of development, dorsoventral patterning	**Siegfried Roth** & **Jeremy Lynch,** Institut für Entwicklungsbiologie, Universität zu Köln	2010-2013
A1	The emergence of Toll signalling as a major component of the dorsoventral patterning network during insect evolution	Evolution of developmental mechanisms in insects	**Siegfried Roth,** Institut für Entwicklungsbiologie, Universität zu Köln	2014-2017
A2	Evolution of the regulatory interactions of the segmentation gene network in insects	Evolution of development, segmentation	**Diethard Tautz,** Institut für Genetik, Universität zu Köln	2006-2008
A4	Analysis of a new type of MADS-box gene that is involved in the evolution of land plant gametophytes	Evolutionary genetics, functional evolution of transcription factors	**Thomas Münster,** MPI für Züchtungsforschung, Köln	2006-2009
A5	Stem cell fate control in grasses: modulations during domestication of maize and in the orientation of the primary embryonic axis	Evolution of development, plant embryos	**Wolfgang Werr,** Institut für Entwicklungsbiologie, Universität zu Köln	2006-2009
A5	The evolution of the root and shoot stem cell niches in land plants: contribution of the WOX gene family	Evolution of development, plant embryos	**Wolfgang Werr,** Institut für Entwicklungsbiologie, Universität zu Köln	2010-2013
A5	The evolution of seed plant stem cell niches	Evolution of developmental mechanisms in plants	**Wolfgang Werr,** Institut für Entwicklungsbiologie, Universität zu Köln	2014-2017
A6	Evolution of the reproductive mode in nematodes	Evolution of development, reproductive processes	**Einhard Schierenberg,** Zoologisches Institut, Universität zu Köln	2006-2009

A9	Micro-evolution of a regulatory network in Brassicaceae	Plant Biology, Genetics	**Martin Hülskamp,** Institut für Botanik, Universität zu Köln	2010-2013
A9	Evolutionary comparison of a gene regulatory network between Arabidopsis thaliana and Arabis alpina	Evolution of developmental pathways in plants	**Martin Hülskamp,** Institut für Botanik, Universität zu Köln	2014-2017
A10	Evolution of annual and perennial life histories in the *Brassicaceae*	Evolution of adaptive traits in plants, evolution of transcriptional regulation	**George Coupland,** MPI für Planzen-züchtungsforschung, Köln	2010-2013
A10	Evolution of transcriptional regulation during the divergence of annual and perennial plant life histories	Evolution of adaptive traits in plants, evolution of transcriptional regulation	**George Coupland,** MPI für Planzenzüchtungs-forschung, Köln	2014-2017
A12	The changing roles of *Hox3* genes in insect evolution	Evolution of transcriptional control in insects	**Kristen Panfilio,** Institut für Entwicklungsbiologie, Universität zu Köln	2014-2017
A13	An interdisciplinary approach for understanding diversification of leaf form	Evolution of development in plants, regulatory networks, morphogenesis	**Miltos Tsiantis,** MPI für Planzenzüchtungs-forschung, Köln	2014-2017
Project area B: Evolution within species				
B1	Characterization of potential adaptive trait loci in house mouse populations	Population genetics, adaptive mutations	**Diethard Tautz,** Institut für Genetik, Universität zu Köln	2006-2008
B2	The role of regulatory changes in speciation processes of the house mouse	Population genetics, evolution of transcriptional regulation	**Bettina Harr,** Institut für Genetik, Universität zu Köln	2006-2008
B3	The origin and evolutionary dynamics of the p47 GTPases, a vertebrate pathogen resistance mechanism	Evolutionary genetics, immunity	**Jonathan Howard,** Institut für Genetik, Universität zu Köln	2006-2009
B3	Reciprocal polymorphism and co-adaptation in the *Toxoplasma*-mouse parasite-host relationship	Host-pathogen coevolution at the species level, origin and evolution of a complex polymorphism	**Jonathan Howard,** Institut für Genetik, Universität zu Köln	2010-2013
B3	Evolution of virulence polymorphism in a co-evolving system	Population genetics, host-pathogen co-evolution	**Jonathan Howard,** Institut für Genetik, Universität zu Köln	2014-2017
B5	The genetic basis of adaptative responses to local demands on seed dormancy in two *Arabidopsis* species	Population genetics, evolution of transcriptional regulation	**Juliette de Meaux** & **Maarten Koornneef,** MPI Planzen-züchtungsforschung, Köln	2006-2009

B5	The adaptive value of pleiotropic variation: miR824 and the MADS-Box Gene Network	Plant biology, Genetics, Natural Variation, Evolutionary biology, pleiotropy, microRNA	**Juliette de Meaux** & **Maarten Koornneef,** MPI für Planzen-züchtungsforschung, Köln	2010-2013
B7	Evolutionary adaptation of *Daphnia* to protease inhibitors in cyanobacteria	Local adaptation, food quality, microevolution, herbivore	**Eric von Elert,** Zoologisches Institut, Universität zu Köln	2010-2013
B10	Molecular analysis and evolutionary history of an epistatic interaction leading to incompatibility	Plant biology, genetics, natural variation, evolutionary biology	**Jane E. Parker** & **Matthieu Reymond,** MPI für Planzenzüchtungs-forschung, Köln	2010-2013
B10	Temperature modulation of plant immune responsiveness and fitness	Evolution of plant defence networks, stress hormone pathway homeostasis	**Jane E. Parker,** MPI für Planzenzüchtungs-forschung, Köln	2014-2017
B11	Cost and benefit of bacterial transformation	Gene transfer, experimental evolution, microbial genetics	**Berenike Maier,** Institut für Theoretische Physik, Universität zu Köln und Biozentrum	2011-2013
B11	Cost and benefit of bacterial transformation	Adaptive evolution of bacteria	**Berenike Maier,** Institut für Biologische Physik, Universität zu Köln und Biozentrum	2014-2017
B12	The adaptive significance of small-effect mutations in yeast and plants	Population genetics, evolution of regulation	**Andreas Beyer** Institut für Genetik Universität zu Köln	2014-2017
B13 associated	Adaptation of Escherichia coli to a complex ecosystem	Adaptive evolution of bacteria, gut microbiota	**Isabel Gordo,** Instituto Gulbenkian de Ciência	2014-2017
Project area C: Theory and computational analysis				
C1	Evolutionary Analysis of Molecular Networks	Biological systems analysis, network evolution	**Johannes Berg & Michael Lässig** Institut für Theoretische Physik, Universität zu Köln	2006-2009
C1	Comparative analysis of gene expression	Gene expression, population genetics	**Johannes Berg & Michael Lässig** Institut für Theoretische Physik, Universität zu Köln	2010-2013
C1	Evolution of gene expression	Gene expression, population genetics	**Johannes Berg & Michael Lässig** Institut für Biologische Physik, Universität zu Köln	2014-2017
C2	Evolution of regulatory DNA	Evolutionary biology, bioinformatics	**Michael Lässig,** Institut für Theoretische Physik, Universität zu Köln, **Nikolaus Rajewsky,** NYU (affiliated)	2006-2009

C2	Time-dependent selection and adaptation	Population genetics, adaptive evolution, evolution of regulation	**Michael Lässig,** Institut für Theoretische Physik, Universität zu Köln	2010-2013
C2	Adaptative evolution of microbial and viral populations	Evolutionary theory, adaptation	**Michael Lässig,** Institut für Biologische Physik, Universität zu Köln	2014-2017
C3	Sequence space models of adaptive evolution	Evolutionary theory, adaptation	**Joachim Krug,** Institut für Theoretische Physik, Universität zu Köln	2006-2009
C3	Epistasis, recombination and predictability in adaptive evolution	Theoretical population genetics, experimental evolution	**Joachim Krug,** Institut für Theoretische Physik, Universität zu Köln **Arjan de Visser,** Wageningen University	2010-2013
C3	Epistasis, recombination and predictability in adaptive evolution	Theoretical population genetics, experimental evolution of bacterial and fungal systems	**Joachim Krug,** Institut für Biologische Physik, Universität zu Köln **Arjan de Visser,** Wageningen University	2014-2017
C4	Multi-locus selection models in population genetics and application to population variability data from mammals, insects and plants	Population genetics, modeling and statistical analysis	**Thomas Wiehe,** Institut für Genetik, Universität zu Köln	2006-2009
C4	Theory of shortest unique substrings in population genetics and comparative genomics	Population genetics and bioinformatics	**Thomas Wiehe,** Institut für Genetik, Universität zu Köln	2006-2009
C4	Evolutionary innovations – a case study on chromatin insulators	Theoretical population genetics, regulatory evolution in Drosophila, bioinformatics	**Thomas Wiehe,** Institut für Genetik, Universität zu Köln	2010-2013
C6	Control and evolution of somite pattern formation during early developement	Evolution of development, segmentation	**Ulrich Gerland** Institut für Theoretische Physik, Universität zu Köln	2008-2009
C7	Bacterial genome dynamics: gene acquisitions and losses in response to life style changes	Evolution of metabolic and regulatory networks in bacteria, horizontal gene transfer	**Martin Lercher,** Institut für Informatik, Heinrich-Heine-Universität Düsseldorf	2010-2013
C7	The role of horizontal gene transfer in life style changes and metabolic adaptations	Adaptive evolution, microbial evolution, lateral gene transfer	**Martin Lercher,** Institut für Informatik, Heinrich-Heine-Universität Düsseldorf	2014-2017

C9	The evolution of repetitive DNA and its methylation	Protein evolution and function	**Achim Tresch,** Institut für Botanik, Universität zu Köln und MPI für Planzenzüchtungs-forschung, Köln	2014-2017
Project area Z: Central facilities				
Z2	Computational biology service unit	Sequence analysis, bioinformatics	**Andreas Beyer** Institut für Genetik **Michael Lässig** Institut für Biologische Physik Universität zu Köln	2014-2017

4. Scientific events

CRC 680 hosted a broad spectrum of events, ranging from international CRC Conferences to topical Seminar Days. These events greatly contributed to scientific exchange, reflecting and advancing new research directions. Here we list the major events, a full list can be found on the CRC website.

1. **Cologne Spring Meeting: The variable genome**
   18.3. – 20.3. 2009

   Invited speakers: **Stephan Beck** (University College London), **Sebastian Bonhoeffer** (Swiss Federal Institute of Technology, Zürich), **Antony Dean** (University of Minnesota), **Manolis Dermitzakis** (Wellcome Trust Sanger Institute), **Anna Di Rienzo** (University of Chicago), **Richard Durbin** (Wellcome Trust Sanger Institute), **Xavier Estivill** (Centre for Genomic Regulation, Barcelona), **Adam Eyre-Walker** (University of Sussex), **Timothy Frayling** (University of Exeter), **Kelly Frazer** (The Scripps Research Institute), **Laurence Hurst** (University of Bath), **Steve Jones** (University College London), **Andrew Leigh Brown** (University of Edinburgh), **John Mattick** (University of Queensland), **Rasmus Nielsen** (University of California, Berkeley), **Howard Ochman** (University of Arizona), **Redmond O'Hanlon** (Church Hanborough, UK), **Svante Pääbo** (Max Planck Institute for Evolutionary Anthropology, Leipzig), **Nikolaus Rajewsky** (Max Delbrück Center for Molecular Medicine, Berlin), **Stephan C. Schuster** (Pennsylvania State University), **Kári Stefánsson** (deCODE genetics, Reykjavik), **Shamil Sunyaev** (Harvard Medical School, Boston), **Amalio Telenti** (University of Lausanne), **Harmen van de Werken** (Hubrecht Institute, Utrecht), **Joris Veltman** (Radboud University Nijmegen).

2. **Molecular Basis of Evolutionary Innovations**
   01.07. – 03.07. 2010

   Invited speakers: **Ralf Bundschuh** (Ohio State University), **Curtis Callan** (Princeton University), **Andrew Clark** (Cornell University), **John Colbourne** (Indiana University), **Erez Dekel** (Weizmann Institute), **Laurence Hurst** (University of Bath, **Santiago Elena** (Universidad Polytecnica de Valencia), **Eugene Koonin** (National Center for Biotechnology Information9, **Martin Kreitman** (University of Chicago), **Stanislas Leibler** (Rockefeller University, New York and Institute for Advanced Study, Princeton), **Thomas Lengauer** (Max Planck Institute for Informatics), **Sebastian Maerkl** (Ecole Polytechnique de Lausanne), **Christopher Marx** (Harvard University), **Leonid Mirny** (Massachusetts Intitute of Technology), **Ville Mustonen** (Sanger Institute), **C. Pal** (Biological Research Center, Szeged), **Dmitry Petrov** (Stanford University), **Boris Shraiman** (Kavli Institute for Theoretical Physics, Santa Barbara), **Shamil Sunyaev** (Harvard University), **Diethard Tautz** (Max-Planck-Institute for Evolutionary Biology).

3. **Evolutionary Genomics: New Data, New Challenges**
   SFB 680 / FOR 1078 Joint Conference, 29.09. – 01.10. 2011

   Invited speakers: **Rubén Alcázar** (MPI for Plant Breeding Research, Cologne), **John Baines** (Christian-Albrechts-University, Kiel), **Thomas Bataillon** (University of Aarhus), **Frank Chan** (MPI for Evolutionary Biology, Plön), **George Coupland** (MPI for Plant Breeding Research, Cologne), **Susanne Foitzik** (Johannes Gutenberg-University, Mainz), **Isabel Gordo** (Instituto Gulbenkian, Lisbon), **Oskar Hallatschek** (MPI for Dynamics and Self-Organization, Göttingen), **Jonathan Howard** (University of Cologne), **JinYong Hu** (MPI for Plant Breeding Research, Cologne), **Jeffrey Jensen** (University of Massachusetts, Medical School, Worchester), **Michael Lässig** (University of Cologne), **Dirk Metzler** (Ludwig-Maximilians-University, Munich), **Ville Mustonen** (Welcome Trust Sanger Institute), **John Parsch** (Ludwig-Maximilians-University, Munich), **Frank Rosenzweig** (University of Montana, Missoula) **Christian Schlötterer** (University of Veterinary Medicine, Vienna), **Shamil Sunyaev** (Brigham & Women's Hospital and Harvard Medical School), **Karl Schmid** (University of Hohenheim), **Ana Sousa** (Instituto Gulbenkian, Lisbon) **Diethard Tautz** (MPI for Evolutionary Biology, Plön), **Xavier Vekemans** (University of Lille).

4. **Cologne Spring Meeting: Molecular Ecology and Evolution**
   22.02. – 24.02.2012

   Invited Speakers: **Ian Thomas Baldwin** (MPI Jena), **Nitin Baliga** (ISB Seattle), **Andrew Beckermann** (University of Sheffield), **Joy Bergelson** (University of Chicago), **Michael Boots** (University of Sheffield), **John Colbourne** (Indiana University), **Santiago Elena** (IBMCP Valencia), **Duncan Greig** (MPI Plön), **Bryan Grenfell** (Princeton University), **Eddie Holmes** (Pennsylvania State Universita), **Britt Koskella** (Oxford University), **Thomas Mitchell-Olds** (Duke University), **Hélène Morlon** (Ecole Polytechnique Paris), **Wayne Potts** (University of Utah), **Michael. Purugganan** (New York University), **Andrew Rambaut** (University of Edinburgh**) Walter Salzburger** (University of Basel), **Johanna Schmitt** (Brown University), **Diethardt. Tautz** (MPI Ploen), **Miltos Tsiantis** (University of Oxford) **Daniel Weinreich** (Brown)

5. **Viral Evolution: Linking genetics to Epidemics**
   Satellite meeting of Cologne Spring Meeting 2012, 24.02 – 25.02.2012

   Invited Speakes: **Trevor Bedford** (University of Edinburgh), **Nico Beerenwinkel** (ETH Zürich), **Julia Gog** (University of Cambridge), Bryan Grenfell (Princeton University), **Alice McHardy** (University of Düsseldorf), **Sergei Kryazhimskiy** (Harvard University), **Ville Mustonen** (Wellcome Trust Sanger Institute), **Richard Neher** (MPI Tübingen), **Nico Pfeifer** (MPI Saarbrücken), Oliver Pybus (University of Oxford), Annemie Vandamme (KU Leuven)

6. **Evolution of Development**
   27.08.2014

   Invited Speakers: **Cassandra Extavour** (Harvard University), **Nicolas Gompel** (LMU Munich), **Angela Hay,** (MPIPZ Cologne), **Felicity Jones** (FML Tübingen) **Kristen Panfilio** (University of Cologne)

7. **Perspectives in Biophysics**
   08.10. – 10.10.2014

   Invited Speakers: **Timo Betz** (Institute Curie, Paris), **Tobias Bollenbach** (IST Austria), **Konstantin Doubrovinski** (Princeton University), **Damien Faivre** (MPIKG, Potsdam), **Stefano Pagliara** (University of Cambridge), **Aleksandre Persat** (Princeton University), **Florian Rehfeld** (Georg-August-Universität, Göttingen), **Steffen Sahl** (Stanford University), **Kurt Schmoller** (Stanford University), **Ingmar Schön** (ETH Zürich) , **Max Ulbrich** (Uni Freiburg)

8. **Forecasting Evolution?**
   08.07. – 11.07.2015

   Invited Speakers: Dan Andersson (Uppsala University), **Trevor Bedford** (Fred Hutchinson Cancer Research Center), **Jesse Bloom** (Fred Hutchinson Cancer Research Center), **Arup Chakraborty** (Massachusetts Institute of Technology), **Sarah Cobey** (University of Chicago), **Michael Desai** (Harvard University), **Michael Doebeli** (University of British Columbia), **Daniel Fisher** (Stanford University), **Marco Gerlinger** (Institute of Cancer Research London), **Michael Hochberg** (Centre National de la Recherche Scienti que Montpellier), **Christopher Illingworth** (Cambridge University), **Roy Kishony** (Harvard University), **Natalia Komarova** (University of California Irvine), **Richard Lenski** (Michigan State University), **Stanislas Leibler** (Rockefeller University), **Marta Łuksza** (Institute for Advanced Study Princeton), **Luke Mahler** (University of Toronto), **Berenike Maier** (University of Cologne), **Leonid Mirny** (Massachusetts Institute of Technology), **Richard Neher** (Max Planck Institute Tübingen), **Armita Nourmohammad** (Princeton University), **Colin Russell** (University of Cambridge), **Sohrab Shah** (University of British Columbia), **Boris Shraiman** (University of California Santa Barbara) **Olivier Tenaillon** (Inserm Paris), **Aleksandra Walczak** (Ecole Normale Supérieure)

The Seminar Days and the Cologne Evolution Colloquium had a total of 128 external speakers, as detailed on the CRC website. In addition, there have regular long-term visits of external scientists, ranging from project work of a few weeks to sabbatical stays.

5. Organization 2014 - 2017

Scientific Coordinator

Michael Lässig

CRC Board

George Coupland
Martin Lercher
Michael Lässig
Berenike Maier
Siegfried Roth
Thomas Wiehe

Administrative coordinator

Christa Stitz

Participating institutions

Universität zu Köln
Institut für Botanik
Institut für Genetik
Institut für Entwicklungsbiologie
Institut für Theoretische Physik
Institut für Zoologie

Heinrich-Heine-Universität Düsseldorf
Institut für Bioinformatik

Wageningen University
Laboratory of Genetics

Max-Planck-Institut (MPI) für Pflanzenzüchtungsforschung, Köln