

Contents

October 2015 • VOLUME 5 • ISSUE 10 • www.g3journal.org

MEETING REPORT

- 1945–1947 **Meeting Report on the Challenge of Inference from Genome to Phenome**
Bevan Emma Huang, Antonio Reverter, Ian Purvis, and Scott Chapman

INVESTIGATIONS

- 1949–1960 **Endogenous Small RNA Mediates Meiotic Silencing of a Novel DNA Transposon**
Yizhou Wang, Kristina M. Smith, John W. Taylor, Michael Freitag, and Jason E. Stajich
 The *Neurospora crassa* genome is a graveyard of transposon relics due to its efficient genome defense. “Meiotic silencing by unpaired DNA” (sometimes called MSUD) is one post-transcriptional defense mechanism that silences unpaired genomic segments during karyogamy. The authors discovered a novel DNA transposon family (called *Sly*) with multiple copies in several lab strains that are absent from most wild isolates. This is the first demonstration of an intact and likely active DNA transposon in the genus *Neurospora*. Here, small RNAs are required for meiotic silencing in order to suppress spreading of transposons.
- 1961–1971 **Genome-Wide Association Study of Down Syndrome-Associated Atrioventricular Septal Defects**
Dhanya Ramachandran, Zhen Zeng, Adam E. Locke, Jennifer G. Mulle, Lora J.H. Bean, Tracie C. Rosser, Kenneth J. Dooley, Clifford L. Cua, George T. Capone, Roger H. Reeves, Cheryl L. Maslen, David J. Cutler, Eleanor Feingold, Stephanie L. Sherman, and Michael E. Zwick
 Children with Down syndrome (DS) have 2000-fold elevated risk of developing atrioventricular septal defects (AVSD), a severe congenital heart abnormality. To uncover the genetic architecture behind this elevated risk, Ramachandran *et al.* conducted a genome-wide association study on 210 cases with DS and complete AVSD and 242 controls with DS and documented structurally normal hearts, making this the largest heart study conducted to date on a trisomic background. No common variant of large effect size (odds ratio > 2.0) was identified, thus supporting a complex and multifactorial etiology for AVSD even in the sensitized DS population.
- 1973–1981 **The Genetics of Resistance to *Morinda* Fruit Toxin During the Postembryonic Stages in *Drosophila sechellia***
Yan Huang and Deniz Erezyilmaz
 The fruit fly *Drosophila sechellia* is narrowly adapted to the fruit of *Morinda citrifolia*, which is toxic to its sibling species, *Drosophila simulans*. The authors combined a high-throughput, high-resolution genomic tool with QTL mapping to detect evolved genomic regions that confer resistance to *M. citrifolia* in *D. simulans*-*D. sechellia* hybrids. Their analysis shows that *M. citrifolia* resistance is conferred by two large effect loci that act with multiple smaller-effect regions. The discovery of novel, major effect loci is an important first step towards identifying the molecular basis of adaptation in a multicellular organism.

- 1983–1990** **Development of a Tightly Controlled Off Switch for *Saccharomyces cerevisiae* Regulated by Camphor, a Low-Cost Natural Product**
Shigehito Ikushima, Yu Zhao, and Jef D. Boeke
- This article presents a novel transcription switch named the Camphor-Off switch in *Saccharomyces cerevisiae*. Expression of reporters (*gfp* and *ADE2*) under control of the switch was activated in the absence of camphor, and turned off in the presence of micromolar concentration of the compound. Importantly, the switch showed orthogonality with the Tet-Off system that controls expression of a gene with doxycycline. Thus, the Camphor-Off switch provides new options for regulating gene expression in *S. cerevisiae*.
- 1991–1998** **Comparison of Genomic Selection Models to Predict Flowering Time and Spike Grain Number in Two Hexaploid Wheat Doubled Haploid Populations**
Saravanan Thavamanikumar, Rudy Dolferus, and Bala R. Thumma
- Thavamanikumar *et al.* compared the ability of different genomic selection models to predict time to young microspore, spike grain number under control conditions, and spike grain number under osmotic stress conditions in two wheat biparental doubled haploid populations with unrelated parents. Prediction accuracies were compared using BayesB, Bayesian LASSO (BL), ridge regression best linear unbiased prediction (RR-BLUP), partial least square regression (PLS), and sparse partial least square regression (SPLS) models. Results from this study indicate that BayesB and SPLS capture the linkage disequilibrium between markers and traits, effectively leading to higher accuracies. Excluding markers from QTL studies reduces prediction accuracies.
- 1999–2006** **Fingerprinting Soybean Germplasm and Its Utility in Genomic Research**
Qijian Song, David L. Hyten, Gaofeng Jia, Charles V. Quigley, Edward W. Fickus, Randall L. Nelson, and Perry B. Cregan
- The USDA Soybean Germplasm Collection, including 18,480 domesticated and 1168 wild soybeans, were genotyped with the SoySNP50K BeadChip, containing >50K SNPs. Redundant accessions were identified, and distinct genetic backgrounds of soybean from different geographic origins were observed that could be a resource for soybean genetic improvement. Song *et al.* detected a dramatic reduction of genetic diversity based on linkage disequilibrium and haplotype structure analyses of the wild, landrace, and North American cultivar populations, and identified regions associated with domestication and selection imposed by breeding. The authors constructed the first soybean haplotype block maps, which are crucial for association mapping to identify genes.
- 2007–2019** **Second-Generation Linkage Maps for the Pacific Oyster *Crassostrea gigas* Reveal Errors in Assembly of Genome Scaffolds**
Dennis Hedgecock, Grace Shin, Andrew Y. Gracey, David Van Den Berg, and Manoj P. Samanta
- Hedgecock *et al.* present a high-density, second-generation linkage map for the Pacific oyster, based on more than 1100 coding single-nucleotide polymorphisms (SNPs), with average marker-spacing of 1.0 cM and genome coverage of 86%. The authors find that nearly 40% of 260 genome scaffolds with two or more SNPs map to two or more different linkage groups, suggesting substantial errors in the published scaffold assemblies. Variation in marker orders and inter-marker distances among families are attributed largely to widespread distortions of segregation ratios caused by early selective mortality and to genotyping error. A consensus high-density linkage map containing 656 framework markers is presented.
- 2021–2026** **High-Density Genotypes of Inbred Mouse Strains: Improved Power and Precision of Association Mapping**
Christoph D. Rau, Brian Parks, Yibin Wang, Eleazar Eskin, Petr Simecek, Gary A. Churchill, and Aldons J. Lusis
- Human genome-wide association studies (GWAS) have identified thousands of loci associated with disease phenotypes. GWAS studies in rodent models require one or more orders of magnitude fewer strains than human studies. To enable well-powered studies in mice, Rau *et al.* generated high-density genotypes for ~175 inbred strains using the Mouse Diversity Array. The new genotypes have increased marker density, fewer missing values, and provide better identification of heterozygous regions compared to earlier efforts. The authors identified new loci from previously reported association studies using the new genotypes. The data are freely available for download. Web-based tools provide access for association mapping and viewing.

- 2027–2036** **GC-Content of Synonymous Codons Profoundly Influences Amino Acid Usage**
Jing Li, Jun Zhou, Ying Wu, Sihai Yang, and Dacheng Tian
- Centering on the combined GC-proportion of all the synonymous codons for a particular amino acid (GC_{syn}), the authors revealed a number of novel genomic patterns addressing the inner connections between codon/amino acid usage, GC-content, and GC_{syn} . They discovered a unique spectrum of amino acid usage; both the amino acid and codon usage frequencies are nearly constant in the regions that have similar GC-content. This new discovery will greatly help to better understand the mysterious yet intimate relationships between codon, amino acid, and GC-content.
- 2037–2041** **Genome-Wide Linkage Analysis Identifies Loci for Physical Appearance Traits in Chickens**
Yanfa Sun, Ranran Liu, Guiping Zhao, Maiqing Zheng, Yan Sun, Xiaoqiong Yu, Peng Li, and Jie Wen
- Physical appearance traits such as feather-crested head, comb size and type, beard, wattles size, and feathered feet are used to distinguish between breeds of chicken, and may also be associated with economic traits. In this study, Sun *et al.* used a genome-wide linkage analysis to identify candidate regions and genes. Their findings provide important genetic information about physical appearance traits in chickens.
- 2043–2049** **A Tool Set for the Genome-Wide Analysis of *Neurospora crassa* by RT-PCR**
Jennifer M. Hurley, Arko Dasgupta, Peter Andrews, Alexander M. Crowell, Carol Ringelberg, Jennifer J. Loros, and Jay C. Dunlap
- Real Time-PCR (RT-PCR) is a common tool used to track mRNA levels in the model organism *Neurospora crassa*. Two key elements of RT-PCR include efficiently and optimally designed primers as well as invariantly expressed reference genes to serve as internal controls. To provide resources for these specific issues, Hurley *et al.* identified optimal reference genes from several experimental conditions using a Prediction Interval Ranking Score or PIRS. The authors also developed a genome-wide data set of five optimal RT-PCR primers for each gene in the *Neurospora* genome and validated a selection of these primers.
- 2051–2059** **Splicing Machinery Facilitates Post-Transcriptional Regulation by FBFs and Other RNA-Binding Proteins in *Caenorhabditis elegans* Germline**
Preston Novak, Xiaobo Wang, Mary Ellenbecker, Sara Feilzer, and Ekaterina Voronina
- Genetic enhancer screens are a powerful tool to probe complex regulatory networks in development. Novak *et al.* used this approach to identify cofactors of FBF-1 and FBF-2, PUF-family translational repressors important for the maintenance of germline stem cells. They identified several splicing factors contributing to FBF function in stem cell maintenance. This was surprising, since splicing factors were thought to promote differentiation rather than germline stem cell renewal. The authors propose that splicing factors are broadly required for post-transcriptional regulation, thus splicing factor knockdowns produce opposite synthetic phenotypes in different sensitized backgrounds.
- 2061–2071** **Analysis of Circadian Rhythms in the Basal Filamentous Ascomycete *Pyronema confluens***
Stefanie Traeger and Minou Nowrousian
- Adaptation to daily changes in the environment is a central feature for most organisms. Circadian clocks aid with these adaptations by enabling organisms to anticipate instead of simply responding to changes. In *Neurospora crassa*, a model organism for studies of circadian clocks, the frequency (*frq*) gene is an important component of the molecular oscillator. Traeger and Nowrousian analyzed circadian rhythms in the basal filamentous ascomycete *Pyronema confluens*, and show that a rhythmically expressed *frq* is present in this fungus. This suggests that *frq* was already associated with the circadian clock in the last common ancestor of filamentous ascomycetes.
- 2073–2084** **Ensemble Learning of QTL Models Improves Prediction of Complex Traits**
Yang Bian and James B. Holland
- Bian and Holland developed a new method to improve the prediction ability of quantitative trait locus (QTL) mapping models. The method works by taking stratified subsets of linkage map markers (“thinning”), mapping QTL using the thinned map, predicting phenotype values for untested lines, and then averaging predictions over thinned maps. This method avoids some statistical problems that arise when linkage maps are very dense and improves prediction ability.

- 2085–2089 **Using Targeted Resequencing for Identification of Candidate Genes and SNPs for a QTL Affecting the pH Value of Chicken Meat**
Xidan Li, Xiaodong Liu, Javad Nadaf, Elisabeth Le Bihan-Duval, Cécile Berri, Ian Dunn, Richard Talbot, and Dirk-Jan De Koning
- Finding the functional mutation underlying a QTL is a major challenge, especially in non-model species. Previously, Li *et al.* used targeted gene expression to narrow down a QTL on chromosome 1 affecting pH in chicken muscle. In this study, the authors applied next generation sequencing of the candidate region in 10 chickens, five of each QTL genotype. They identify a large number of candidate SNPs, which they prioritize based on location and potential effect. The authors now have a manageable number of candidate SNPs for further testing.
- 2091–2104 **Endoplasmic Reticulum Stress–Related Genes in Yellow Catfish *Pelteobagrus fulvidraco*: Molecular Characterization, Tissue Expression, and Expression Responses to Dietary Copper Deficiency and Excess**
Yu-Feng Song, Zhi Luo, Chao Huang, Qi-Liang Chen, Ya-Xiong Pan, and Yi-Huan Xu
- The full-length cDNA sequences of two ER molecular chaperones (GRP78 and CRT) and three ER stress sensors (PERK, IRE-1 α , and ATF-6 α) cDNAs were first characterized from yellow catfish and their tissue-specific expressions were determined. Meanwhile, the effect of dietary Cu deficiency and excess on their mRNA levels was explored. The results will be beneficial for elucidating the molecular basis of ER stress and for further evaluation of the effects of dietary Cu levels in fish at a molecular level based on the upstream pathway of lipid metabolism. This will greatly increase understanding of the nutrition of Cu in fish.
- 2105–2112 **A Flippase-Mediated GAL80/GAL4 Intersectional Resource for Dissecting Appendage Development in *Drosophila***
Brittany N. Smith, Arash M. Ghazanfari, Rudolf A. Bohm, William P. Welch, Bing Zhang, and John P. Masly
- The ability to target specific tissues or subsets of cells within tissues for genetic manipulation using the GAL4-UAS binary system revolutionized *Drosophila* genetics. However, many of the existing GAL4 expression patterns are rather broad, and increasing the cell and tissue specificity of genetic manipulation would allow researchers to investigate biological phenomena at finer spatial resolutions. In this study, Smith *et al.* characterize a large collection of enhancer-trap Flippase tools that enable such fine-scale dissection of existing GAL4 expression patterns in *Drosophila* imaginal discs. The authors illustrate the potential of these tools for studying developmental, cellular, and molecular processes at the genetic level.
- 2113–2126 **Genomic-Enabled Prediction of Ordinal Data with Bayesian Logistic Ordinal Regression**
Osvaal A. Montesinos-López, Abelardo Montesinos-López, José Crossa, Juan Burgueño, and Kent Eskridge
- The ordinal logistic regression model is often preferred over the ordinal probit model, as it provides regression coefficients that are more interpretable due to their connection to odds ratios and because the logit model is more robust for modeling misspecifications. However, due to the easy implementation of the Bayesian probit ordinal regression (BPOR) model, logistic regression is rarely implemented in genomic prediction. The authors propose a Bayesian logistic ordinal regression (BLOR) model with similar full conditional distributions of the BPOR model. Results indicate that BLOR model is a good alternative for analyzing ordinal data in the context of genomic prediction.
- 2127–2135 **Developmental Transcriptome for a Facultatively Eusocial Bee, *Megalopta genalis***
Beryl M. Jones, William T. Wcislo, and Gene E. Robinson
- Social insects exhibit striking examples of developmental phenotypic plasticity, yet the evolutionary origins of this plasticity are not well understood. Using *Megalopta genalis* bees, Jones *et al.* sequenced and assembled the first developmental transcriptome for a species with facultative eusociality, a trait that offers insights into evolutionary transitions in sociality. The authors document life stage-specific gene expression and shifts in dominant functional processes, including transcriptional regulation and metabolic processes, during development. A number of other distinct biological processes also show characteristic changes during each life stage, facilitating the use of *M. genalis* as a new model to investigate the origins of developmental plasticity.

- 2137–2154 **SUMO-Enriched Proteome for *Drosophila* Innate Immune Response**
Mithila Handu, Bhagyashree Kaduskar, Ramya Ravindranathan, Amarendranath Soory, Ritika Giri, Vijay Barathi Elango, Harsha Gowda, and Girish S. Ratnaparkhi
 Handu *et al.* measured the changes to the SUMO proteome in *Drosophila* cells in response to an immune challenge and find global changes, with protein domains, cellular pathways, and protein complexes responding to immune stress. A confident set of 710 proteins is identified as the immune-induced SUMO proteome that contains *bona fide* SUMO targets such as Caspar, Jra, Kay, cdc42, 14-3-3 ϵ , p38b, and also cellular proteins with diverse functions, such as pros β 4, Rps10b, SmD3, Tango7, and Arginyl tRNA synthetase. This study will help increase understanding of the roles for SUMO modification in the host response to pathogens.
- 2155–2164 **A Genomic Selection Index Applied to Simulated and Real Data**
J. Jesus Ceron-Rojas, José Crossa, Vivi N. Arief, Kaye Basford, Jessica Rutkoski, Diego Jarquín, Gregorio Alvarado, Yoseph Beyene, Kassa Semagn, and Ian DeLacy
 Genomic selection is now a standard method to predict breeding values and to select individuals as parents to the next generation. A genomic selection index (GSI) is a linear combination of genomic estimated breeding values of individuals used to select parents in an unobserved testing population, for which there is only marker information. In this article, the authors propose a GSI and compare its efficiency with that of the Smith phenotypic selection index (PSI). They used simulated and real data sets to compare the efficiency of the GSI and the PSI.
- 2165–2176 **Similar Efficacies of Selection Shape Mitochondrial and Nuclear Genes in Both *Drosophila melanogaster* and *Homo sapiens***
Brandon S. Cooper, Chad R. Burrus, Chao Ji, Matthew W. Hahn, and Kristi L. Montooth
 Nonrecombining genomes, such as the mtDNA and heteromorphic Y/W chromosomes, are predicted to accumulate deleterious mutations. Prior studies comparing samples of mitochondrial and nuclear genes have largely supported the idea that selection is less effective at removing deleterious variants from mitochondrial relative to nuclear genes. Using large samples of mitochondrial and nuclear whole-genome data from fruit flies and humans, Cooper *et al.* find little evidence that mitochondrial genes, on average, harbor more nonsynonymous polymorphism than do nuclear genes. Their findings challenge the current understanding of animal mitochondrial evolution and will lead to investigations of how selection maintains protein sequence in nonrecombining genomes.
- 2177–2186 **Identification and Correction of Sample Mix-Ups in Expression Genetic Data: A Case Study**
Karl W. Broman, Mark P. Keller, Aimee Teo Broman, Christina Kendziorski, Brian S. Yandell, Śaunak Sen, and Alan D. Attie
 In a large mouse eQTL study with 500 intercross mice and genome-wide gene expression data on six tissues, the authors identified a high proportion (18%) of sample mix-ups in the genotype data, as well as a smaller number of mix-ups in the expression data for each tissue. They describe a simple and practical approach for identifying and correcting such problems. The approach has been implemented in the R package R/lineup, which is available at <https://github.com/kbroman/lineup>.
- 2187–2197 **Genetic Networks Required to Coordinate Chromosome Replication by DNA Polymerases α , δ , and ϵ in *Saccharomyces cerevisiae***
Marion Dubarry, Conor Lawless, A. Peter Banks, Simon Cockell, and David Lydall
 Genome replication is an important, complex, and difficult task. In eukaryotes, DNA replication depends principally on three DNA polymerase activities. Dubarry *et al.* used systematic genetic approaches to investigate DNA replication when each polymerase is defective. They identify hundreds of genetic interactions affecting replication. Comparison of fitness profiles between defective DNA polymerase strains allows them to distinguish factors that affect different aspects of replication. New online tools facilitate access and interaction with these data, which should improve the understanding how eukaryotic DNA replication is coordinated.

2199–2207 **Selection Against Maternal microRNA Target Sites in Maternal Transcripts**

Antonio Marco

The first steps of development in any animal are mediated by maternal products, including microRNAs. Maternal microRNAs are found in many species, but their function seems to be suppressed. Whether maternal microRNAs are functional in *Drosophila* is still unknown. In this article Marco characterizes a set of maternal microRNAs in *Drosophila*. Through the study of the evolutionary profile of their targeted transcripts, a functional impact became evident: many maternal transcripts avoid being targeted by maternal microRNAs. The study also reveals that maternal microRNAs are often the product of transcriptional hitchhiking. Marco concludes that maternal microRNAs have an impact in maternal gene evolution.

MUTANT SCREEN REPORT

2209–2215 **A Forward Genetic Screen for Suppressors of Somatic P Granules in *Caenorhabditis elegans***

Ashley L. Kelly, Michael J. Senter-Zapata, Anne C. Campbell, Hannah E. Lust, Monique E. Theriault, Karolina M. Andralojc, and Dustin L. Updike

Germline programs are actively repressed in somatic tissue. While components of the synMuv B chromatin remodeling complex are known contributors to this repression, the full scope of pathways that suppress germline expression in the soma is unclear. Here the authors describe a forward genetic screen to identify suppressors of somatic P-granule expression in *Caenorhabditis elegans*. Of the eight suppressors recovered from the screen, all are components of the synMuv B chromatin remodeling pathway. This suggests that suppressors outside of this pathway are either rare or required for viability.

2217 **CORRIGENDUM**