but enhanced by GTP-γS and AlF_4^- (Fig. 2, C and D), indicating that the cytoplasmic domains of β_3 and β_1 can directly interact with $G\alpha_{13}$ and that GTP enhances the interaction. The $G\alpha_{13}$-β_3 interaction was enhanced in platelets adherent to fibrinogen, and by thrombin, which stimulates GTP binding to $G\alpha_{13}$ via GPCR (Fig. 2E). Hence, the interaction is regulated by both integrin occupancy and GPCR signaling.

To map the β_3 binding site in $G\alpha_{13}$, we incubated cell lysates containing Flag-tagged wild type or truncation mutants of $G\alpha_{13}$ (fig. S5) with GST-β_3CD beads. GST-β_3CD associated with wild-type $G\alpha_{13}$ and the $G\alpha_{13}$ 1 to 212 fragment containing α-helical region and switch region I (SRI), but not with the $G\alpha_{13}$ fragment containing residues 1 to 196 lacking SRI (Fig. 2F). Thus, SRI appears to be critical for β_3 binding. To further determine the importance of SRI, $G\alpha_{13}$-β_3 binding was assessed in the presence of a myristoylated synthetic peptide, Myr-LLARRPTKGIHEY (mSRI), corresponding to the SRI sequence of $G\alpha_{13}$ (197 to 209) (*21*, *22*). The mSRI peptide, but not a myristoylated scrambled peptide, inhibited $G\alpha_{13}$ binding to β_3 (Fig. 2G), indicating that mSRI is an effective inhibitor of β_3-$G\alpha_{13}$ interaction. Therefore, we further examined whether mSRI might inhibit integrin signaling. Treatment of platelets with mSRI inhibited integrin-dependent phosphorylation of c-Src Tyr^{416} and accelerated RhoA activation (Fig. 3A). The effect of mSRI is unlikely to result from its inhibitory effect on the binding of RhoGEFs to $G\alpha_{13}$ SRI because $G\alpha_{13}$ binding to RhoGEFs stimulates RhoA activation, which should be inhibited rather than promoted by mSRI (*22*). Thus, these data suggest that β_3-$G\alpha_{13}$ interaction mediates activation of c-Src and inhibition of RhoA. Furthermore, mSRI inhibited integrin-mediated platelet spreading (Fig. 3B), and this inhibitory effect was reversed by C3 toxin (which catalyzes the ADP ribosylation of RhoA) or Y27632, confirming the importance of $G\alpha_{13}$-dependent inhibition of RhoA in platelet spreading. Thrombin promotes platelet spreading, which requires cdc42/Rac pathways (*23*). However, thrombin-promoted platelet spreading was also abolished by mSRI (Fig. 3B), indicating the importance of $G\alpha_{13}$-β_3 interaction. Thus, $G\alpha_{13}$-integrin interaction appears to be a mechanism that mediates integrin signaling to c-Src and RhoA, thus regulating cell spreading.

To further determine whether $G\alpha_{13}$ mediates inhibition of integrin-induced RhoA-dependent contractile signaling, we investigated the effects of mSRI and depletion of $G\alpha_{13}$ on platelet-dependent clot retraction (shrinking and consolidation of a blood clot requires integrin-dependent retraction of platelets from within) (*7*, *8*). Clot retraction was accelerated by mSRI and depletion of $G\alpha_{13}$ (Fig. 4, A and B, and fig. S6), indicating that $G\alpha_{13}$ negatively regulates RhoA-dependent platelet retraction and coordinates cell spreading and retraction. The coordinated cell spreading-retraction process is also important in wound healing, cell migration, and proliferation (*24*).

The function of $G\alpha_{13}$ in mediating the integrin-dependent inhibition of RhoA contrasts with the traditional role of $G\alpha_{13}$, which is to mediate GPCR-induced activation of RhoA. However, GPCR-mediated activation of RhoA is transient, peaking at 1 min after exposure of platelets to thrombin, indicating the presence of a negative regulatory signal (Fig. 4, D and F). Furthermore, thrombin-stimulated activation of RhoA occurs during platelet shape change before substantial ligand binding to integrins (Fig. 4, C, D, and F). In contrast, after thrombin stimulation, β_3 binding to $G\alpha_{13}$ was diminished at 1 min when $G\alpha_{13}$-dependent activation of RhoA occurs, but increased after the occurrence of integrin-dependent platelet aggregation (Fig. 4, E and F). Thrombin-stimulated binding of $G\alpha_{13}$ to $\alpha_{IIb}\beta_3$ and simultaneous RhoA inhibition both require ligand occupancy of $\alpha_{IIb}\beta_3$ and are inhibited by the integrin inhibitor Arg-Gly-Asp-Ser (RGDS) (Fig. 4, D to F). Thus, our study demonstrates not only a function of integrin $\alpha_{IIb}\beta_3$ as a noncanonical $G\alpha_{13}$-coupled receptor but also a new concept of $G\alpha_{13}$-dependent dynamic regulation of RhoA, in which $G\alpha_{13}$ mediates initial GPCR-induced RhoA activation and subsequent integrin-dependent RhoA inhibition (Fig. 4G). These findings are important for our understanding of how cells spread, retract, migrate, and proliferate, which is fundamental to development, cancer, immunity, wound healing, hemostasis, and thrombosis.

References and Notes

1. R. O. Hynes, *Cell* **110**, 673 (2002).
2. M. H. Ginsberg, A. Partridge, S. J. Shattil, *Curr. Opin. Cell Biol.* **17**, 509 (2005).
3. Y. Q. Ma, J. Qin, E. F. Plow, *J. Thromb. Haemost.* **5**, 1345 (2007).
4. S. J. Shattil, *Trends Cell Biol.* **15**, 399 (2005).
5. A. Obergfell *et al.*, *J. Cell Biol.* **157**, 265 (2002).
6. E. G. Arias-Salgado *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13298 (2003).
7. P. Flevaris *et al.*, *J. Cell Biol.* **179**, 553 (2007).
8. P. Flevaris *et al.*, *Blood* **113**, 893 (2009).
9. N. A. Riobo, D. R. Manning, *Trends Pharmacol. Sci.* **26**, 146 (2005).
10. L. F. Brass, D. R. Manning, S. J. Shattil, *Prog. Hemost. Thromb.* **10**, 127 (1991).
11. A. Moers *et al.*, *Nat. Med.* **9**, 1418 (2003).
12. T. Kozasa *et al.*, *Science* **280**, 2109 (1998).
13. M. J. Hart *et al.*, *Science* **280**, 2112 (1998).
14. B. Klages, U. Brandt, M. I. Simon, G. Schultz, S. Offermanns, *J. Cell Biol.* **144**, 745 (1999).
15. V. Senyuk *et al.*, *Cancer Res.* **69**, 262 (2009).
16. B. S. Coller, *Blood* **55**, 169 (1980).
17. Z. Li, G. Zhang, R. Feil, J. Han, X. Du, *Blood* **107**, 965 (2006).
18. M. Gu, X. Xi, G. D. Englund, M. C. Berndt, X. Du, *J. Cell Biol.* **147**, 1085 (1999).
19. W. T. Arthur, L. A. Petch, K. Burridge, *Curr. Biol.* **10**, 719 (2000).
20. S. Tanabe, B. Kreutz, N. Suzuki, T. Kozasa, *Methods Enzymol.* **390**, 285 (2004).
21. Single-letter abbreviations for amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
22. J. S. Huang, L. Dong, T. Kozasa, G. C. Le Breton, *J. Biol. Chem.* **282**, 10210 (2007).
23. C. Vidal, B. Geny, J. Melle, M. Jandrot-Perrus, M. Fontenay-Roupie, *Blood* **100**, 4462 (2002).
24. K. Moissoglu, M. A. Schwartz, *Biol. Cell* **98**, 547 (2006).
25. X. P. Du *et al.*, *Cell* **65**, 409 (1991).
26. X. D. Ren, M. A. Schwartz, *Methods Enzymol.* **325**, 264 (2000).

27. This work was supported by grants HL080264, HL062350, and HL068819 from the National Heart, Lung, and Blood Institute (X.D.) and GM061454 and GM074001 from the National Institute of General Medical Sciences (T.K.). We thank G. Nucifora for help with bone marrow transplantation and K. O'Brien and M. K. Delaney for proofreading.

Supporting Online Material

www.sciencemag.org/cgi/content/full/327/5963/340/DC1
Materials and Methods
Figs. S1 to S6
References

9 April 2009; accepted 4 December 2009
10.1126/science.1174779

Functional and Evolutionary Insights from the Genomes of Three Parasitoid *Nasonia* Species

The *Nasonia* Genome Working Group*†

We report here genome sequences and comparative analyses of three closely related parasitoid wasps: *Nasonia vitripennis*, *N. giraulti*, and *N. longicornis*. Parasitoids are important regulators of arthropod populations, including major agricultural pests and disease vectors, and *Nasonia* is an emerging genetic model, particularly for evolutionary and developmental genetics. Key findings include the identification of a functional DNA methylation tool kit; hymenopteran-specific genes including diverse venoms; lateral gene transfers among Pox viruses, *Wolbachia*, and *Nasonia*; and the rapid evolution of genes involved in nuclear-mitochondrial interactions that are implicated in speciation. Newly developed genome resources advance *Nasonia* for genetic research, accelerate mapping and cloning of quantitative trait loci, and will ultimately provide tools and knowledge for further increasing the utility of parasitoids as pest insect-control agents.

Parasitoid wasps are insects whose larvae parasitize various life stages of other arthropods (for example, insects, ticks, and mites). Female wasps sting, inject venom, and lay eggs on or in the host, where the developing offspring consume and eventually kill it. Parasitoids

Downloaded from www.sciencemag.org on January 15, 2010

are widely used in the biological control of insect pests, and they are very diverse, with estimates of over 600,000 species (*1*, *2*). *Nasonia* is the second genus of Hymenoptera to have whole-genome sequencing, after *Apis mellifera* (Fig. 1), and *Nasonia* comprises four closely related parasitoid species: *N. vitripennis*, *N. giraulti*, *N. longicornis*, and *N. oneida* (*3*, *4*). *Nasonia* are genetically tractable organisms with short generation time (~2 weeks), large family size, ease of laboratory rearing, and cross-fertile species. Like other hymenopterans, haploid males develop from unfertilized eggs, and diploid females develop from fertilized eggs. Cross-fertile species facilitate the mapping and cloning of genes that are involved in species differences. Haploid genetics assist efficient genotyping, mutational screening (*5*), and evaluation of gene interactions (epistasis) without the added complexity of genetic dominance. As a result, *Nasonia* are now emerging as genetic model organisms, particularly for complex trait analysis, developmental genetics, and evolutionary genetics (*4*).

We sequenced, assembled, annotated, and analyzed the genome of *N. vitripennis* from sixfold Sanger sequence genome coverage by using a highly inbred line of *N. vitripennis* (*6*). The draft genome assembly comprises 26,605 contigs [total length of 239.8 Mb, with half of the bases residing in contigs larger than 18.5 kb (N_{50}), 40.6% guanine plus cytosine content (GC)]. Contigs were placed with mate-pair information into 6,181 scaffolds (total size 295 Mb, N_{50} =709 kb). We assessed the *N. vitripennis* assembly for completeness and accuracy by comparing it with 19 finished bacterial artificial chromosome (BAC) sequences and 18,000 expressed sequence tags (ESTs). The genome assembly contained 98% of the BAC and 97% of the EST sequences, with an error rate of 5.9×10^{-4}. Thus the assembly is a high-quality representation of both genomic and transcribed *N. vitripennis* sequences.

Highly inbred lines of the two sibling species *N. giraulti* and *N. longicornis* (Fig. 1B) were sequenced with onefold Sanger and 12-fold, 45–base pair (bp) Illumina genome coverage. Assembled by alignment to the *N. vitripennis* reference using stringent criteria (*6*), these reads cover 62% and 62.6% of the *N. vitripennis* assembly, and 84.7% and 86.3% of protein coding regions, respectively. These were used for genome comparisons and provided resources [for example, single nucleotide polymorphisms (SNPs) and microsatellites] for scaffold, gene, and quantitative trait loci (QTL) mapping. Sequence error rates for the *N. giraulti* alignment are estimated to be 3.8×10^{-3} for the entire alignment and 1.47×10^{-4} for coding sequences on the basis of comparison to three finished *N. giraulti* BACs (*6*). Sequences of 25 coding genes in both species perfectly matched their respective aligned sequences.

Normally, the intracellular bacteria *Wolbachia* prevent the formation of interspecies hybrids; however, antibiotically cured strains are cross-fertile (*7*). Hybrid crosses (Fig. 1C) (*6*) were used to map scaffolds and visible mutations onto the five chromosomes of *Nasonia* (Fig. 2). Several interspecies QTL have already been mapped using genetic/genomic resources, including wing size (*8*, *9*), host preference (*10*), female mate preference (*11*), and in this study, sex-ratio control and male courtship (*6*). Linkage analysis has revealed that the genome-wide recombination rate in *Nasonia* is 1.4 to 1.5 centimorgans (cM)/Mb, which is lower than that of honeybees (*12*, *13*), and shows a 100-fold difference in rate between high- and low-recombination regions of the genome (Fig. 2) (*6*).

An official gene set (OGS v1.1) was generated from comparisons to *A. mellifera*, *Tribolium castaneum*, *Drosophila melanogaster*, *Pediculus humanus*, *Daphnia pulex*, and *Homo sapiens* [details are given in (*6*)]. Overall, *Nasonia* encodes a typical insect gene repertoire (Fig. 3) (*6*), of which 60% of genes have a human ortholog, 18% are arthropod-specific, and 2.4% appear to be hymenoptera-specific, showing high conservation between *Nasonia* and *Apis* and low conservation or absence in other taxa. An additional 12% are either *Nasonia*-specific or without clear orthology. Many (63%) single-copy orthologs shared between *Nasonia* and *Apis* occur in microsynteny blocks, which is similar to the amount of microsynteny blocks in *Aedes aegypti*/*Anopheles gambiae* and *H. sapiens*/*Gallus gallus* (*14*). Four hundred and forty-five orthologs between *Nasonia* and humans lack a candidate homolog in *D. melanogaster* (table S1), including the human transcription factors E2F7 and E2F8, which are involved in cell-cycle regulation. Further refinement of the gene set resulted in OGS v1.2 (*15*), which totals 17,279 genes, of which 74% have tiling microarray or EST support (*6*).

Nasonia is abundant in transposable elements (TEs) and other repetitive DNA (table S2 and fig. S1). This contrasts with a paucity of TEs in *A. mellifera* (*16*). TE diversity in *Nasonia* is 30% higher (2.9 TE types/Mb) than the next most diverse insect (*Bombyx mori*, 2.1 TE types/Mb), and is 10-fold higher than the average dipteran (*6*, *17*). *Nasonia* also contains an unusual abundance of nuclear-mitochondrial insertions and a higher density of microsatellites (10.9 kb/Mb) than most other arthropod species (*18*, *19*), suggesting that the accumulation of repetitive DNA is a feature of these insects.

The *Nasonia* genome encodes a full DNA methylation tool kit, including all three DNA cytosine-5-methyltransferase (Dnmt) types (Fig. 1A).

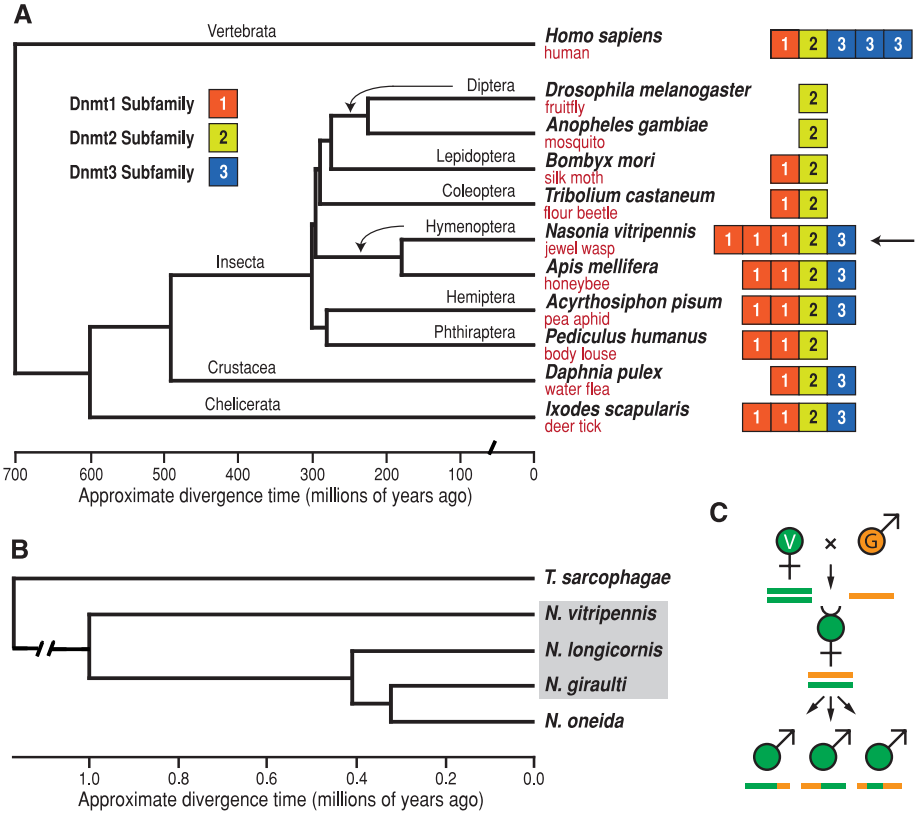

Fig. 1. Phylogenetic relationships of *Nasonia* and the DNA methylation tool kit. (**A**) *Nasonia* relationships to other sequenced genomes (*6*). Right: DNA methyltransferase subfamilies (Dnmt1, Dnmt2, Dnmt3) in these taxa. (**B**) Relationships among the three sequenced *Nasonia* genomes. (**C**) Crossing scheme used for mapping scaffolds on the *Nasonia* chromosomes and for studies of nuclear-cytoplasmic incompatibility.

*All authors with their affiliations appear at the end of this paper.
†To whom correspondence should be addressed. E-mail: werr@mail.rochester.edu (J.H.W.); stephenr@bcm.tmc.edu (S.R.)

Downloaded from www.sciencemag.org on January 15, 2010

In vertebrates, Dnmt3 establishes DNA methylation patterns, Dnmt1 maintains these patterns, and Dnmt2 is involved in tRNA methylation (*20*). The *Nasonia* genome encodes three Dnmt1 genes, one Dnmt2, and one Dnmt3, in contrast with *D. melanogaster*, which has only Dnmt2. The presence of all three subfamilies in both *Nasonia* and *Apis* (Fig. 1) raises the question of whether methylation has similar regulatory functions in Hymenoptera as it does in vertebrates. DNA methylation is important in *Apis* caste development (*21*) and is suggested for *Nasonia* sex determination (*22*). Coding exons of both *Nasonia* and *Apis* show bimodal distributions in observed/expected CpG (fig. S2) (*6*, *23*), which is consistent with mutational biases due to DNA methylation of hyper- and hypomethylated genes. We confirmed methylated CpG dinucleotides in five examined *N. vitripennis* genes by bisulfite sequencing (fig. S3). These results suggest that epigenetic modifications by DNA methylation may be important in Hymenoptera. *Nasonia* also has the largest number of ankyrin (ANK) repeat–containing proteins (over 200) so far found in any insect (table S3) (*6*), suggesting a regulatory importance through protein-protein interactions (*24*).

Systemic RNA interference (RNAi) in *Nasonia* allows for gene expression knockdowns (*4*, *25*). The *Nasonia* genome encodes homologs for the majority of genes implicated in small RNA processes (table S4). However, as in *Tribolium* and *Apis*, *Nasonia* lacks an RNA-dependent RNA polymerase (RdRp) ortholog, indicating a different systemic RNAi mechanism than in *Caenorhabditis*. Using various computational approaches (*6*), we identified 52 putative micro RNAs (miRNAs) with homologies to known miRNAs (*26*), nine that were previously unknown, and 11 additional Hymenoptera-specific miRNAs (table S5). Small-RNA library sequencing confirmed 39 predicted and identified 59 additional miRNAs (table S6).

Nasonia shares a long germ-band mode of embryonic development with *Drosophila*, but exhibits significant differences in the genetic mechanisms involved (*5*, *27*, *28*) (see fig. S4). All major components of the dorso-ventral patterning system are present, with many *Nasonia*-specific gene duplications in the Toll pathway. Orthologs of vertebrate genes absent from *Drosophila* include the transforming growth factor–β (TGFβ) ligands *ADMP* and *myostatin*, and the bone morphogenesis protein (BMP) inhibitors *BAMBI* and *DAN*, but their functions in *Nasonia* are not yet known. *A. mellifera* shows an expansion of the *yellow/major royal jelly* (*yellow/MRJP*) genes that are linked to caste formation and sociality (*29*). *Nasonia* has the largest number of *yellow/MRJP* genes so far found in any insect, including an independent amplification of MRJP-like proteins (fig. S5) (*6*, *29*). Although their function in *Nasonia* is unknown, these genes are expressed broadly in different tissues and life stages (table S7). The insect sex peptide/receptor system, which causes female re-mating refractoriness (*30*), is highly conserved in insects but is absent in *Nasonia* and *Apis* (table S8) (*6*). Instead, *Nasonia* males inhibit female re-mating behaviorally with a special "post-copulatory display" (*31*). Additional features analyzed (*6*) include those related to sex determination (fig. S6), pathogens and immunity (fig. S7), neuropeptides (tables S9 and S10), cuticular proteins (table S11), xenobiotics (fig. S8), and diapause (table S12).

We investigated genome microevolution, including rapidly evolving genes that are potentially involved in species differences and speciation, by using the genomes of the three closely related *Nasonia* species. Synonymous divergence between *N. vitripennis* and its sibling species *N. giraulti* and *N. longicornis* is 0.031 ± 0.0002 SE and 0.030 ± 0.0002 SE, respectively, and between *N. giraulti* and *N. longicornis* is 0.014 ± 0.0001 SE (*6*), which is comparable to those among *Drosophila* sibling species (*32*). We compared the ratio of synonymous-to-nonsynonymous substitutions (*dN*/*dS*) between *Nasonia* species pairs with respect to gene ontology (GO) term categories, using genes with high-quality alignments and 1:1 orthologs between *Nasonia* and *Drosophila*. Nuclear genes that interact with mitochondria revealed significantly elevated *dN*/*dS* [by comparison of *dN*/*dS* distributions for each GO term to resampled distributions, see (*6*) and table S13], specifically those encoding mitochondrial ribosomes ($P < 0.003$ for all species pairs) and oxidative phosphorylation complex I ($P < 0.03$ for *N. vitripennis*/*N. giraulti* and *N. vitripennis*/*N. longicornis*) and complex V ($P < 0.04$ for all species pairs). This finding is consistent with the rapid evolutionary rate of *Nasonia* mitochondria (*33*) and studies implicating nuclear-mitochondrial incompatibilities in F2 hybrid breakdown (*7*, *31*). For example, reciprocal crosses between *N. giraulti* × *N. vitripennis* have identical F1 nuclear genotypes, but their mitochondrial haplotypes differ. Yet, microarray hybridization (Fig. 2) (*6*) of DNA from pooled surviving adult F2 haploid males shows distortion in the recovery of particular regions of the genome, which is dependent upon their mitochondrial haplotype (*giraulti* versus *vitripennis*). Because hybrid mortality is post-embryonic (*7*) and embryo ratios are Mendelian (*33*), these distortions reflect larval to adult mortality. In particular, F2 males with *N. vitripennis* alleles on the left arm of chro-

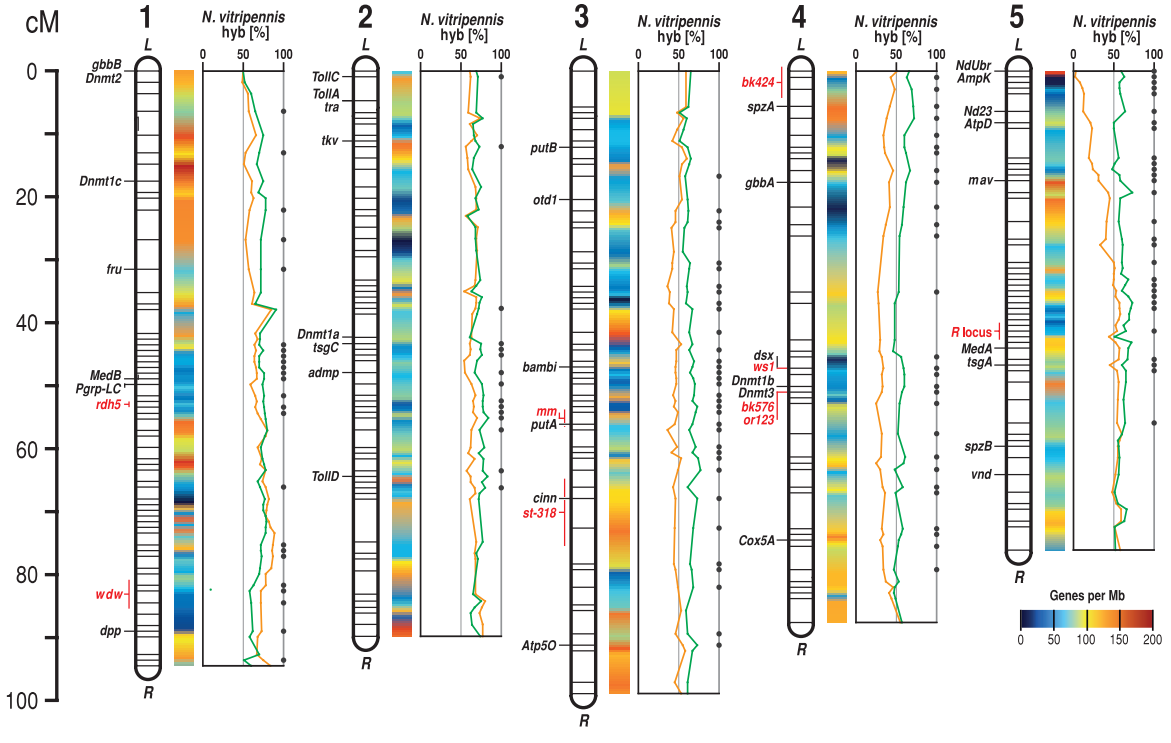

Fig. 2. A high-resolution recombination map of the five *Nasonia* chromosomes is shown (*6*), with estimated gene density and locations of visible markers, landmark genes, and QTL. The hybridization percentage to *N. vitripennis* alleles is shown among surviving adult *N. vitripennis* × *N. giraulti* F2 hybrid males with either *N. vitripennis* (green curve) or *N. giraulti* (orange curve) mitochondria. Dots specify genome regions with significant differences in the hybridization ratio between the reciprocal crosses ($P < 0.01$).

Downloaded from www.sciencemag.org on January 15, 2010

mosome 5 and *N. giraulti* mitochondria suffer nearly 100% mortality (Fig. 2). This region contains three genes encoding mitochondrial interacting proteins, *atpD*, *ampK*, and *nadh-ubiquinone oxireductase* (Fig. 2). Coevolution of nuclear and mitochondrial genomes can accelerate evolution (*34*, *35*), and these findings indicate that such interactions contribute to reproductive incompatibility and speciation in *Nasonia*.

Sequences of 25 gene regions from multiple strains for the three *Nasonia* species (*6*) show low levels of intraspecific variation (table S14) with synonymous site variation ranging from 0.0005 in *N. giraulti* to 0.0026 in *N. vitripennis*, which are much lower than in *Drosophila* species and more akin to levels observed in humans (*36*). This low nuclear variation could be explained by founder events, purging of deleterious mutations in haploid males, or inbreeding.

Recent lateral gene transfers from the bacterial endosymbiont *Wolbachia* into the genomes of *Nasonia* and other arthropods have been identified (*37*). Detecting ancient lateral transfers is more problematic. By examining protein domain arrangements in *Nasonia* relative to other organisms, we uncovered an ancient lateral gene transfer involving Pox viruses, *Wolbachia*, and *Nasonia*. Thirteen ANK repeat–bearing proteins encoded in the *N. vitripennis* genome also contain C-terminal PRANC (Pox proteins repeats of ankyrin–C terminal) domains. This domain was previously only described in Pox viruses, where it is associated with ANK repeats and inhibits the nuclear factor κB (NF-κB) pathway in mammalian hosts (*38*). A computational screen revealed ANK-PRANC–bearing genes in some *Wolbachia* and a related Rickettsiales (Fig. 4). Screening additional *Wolbachia* confirmed the presence of ANK-PRANC genes in diverse *Wolbachia*. The *Nasonia* PRANC genes are clearly integrated in the genome (*6*) and are expressed in different life stages (table S15). Phylogenetic analysis of the PRANC-domain sequences suggests that the *Nasonia* lineage acquired one or more of these proteins from *Wolbachia*, with subsequent amplification and divergence (Fig. 4). Such lateral gene transfers between bacteria and animals could be an important source of evolutionary innovation (*37*).

Nasonia is a carnivore, feeding on an amino acid–rich diet both as larva and adult (*4*). Mapping of *Nasonia* genes onto metabolic pathways (*39*) revealed loss or rearrangement in some amino acid metabolic pathways, including tryptophan and aminosugar metabolism (fig. S9) (*6*). The changes may reflect its specialized carnivorous diet and can inform efforts to produce artificial diets for more economical parasitoid rearing.

The venom of parasitoids, injected into a host before oviposition, serves to condition the host for successful development of wasp progeny (*1*, *2*). Unlike the defensive *Apis* venom that inflicts pain and damage, parasitoid venoms have diverse physiological effects on hosts, including developmental arrest; alteration in growth and physiology; suppression of immune responses; induction of paralysis, oncosis, or apoptosis; and alteration of host behavior (*40*). The identification of *Nasonia*

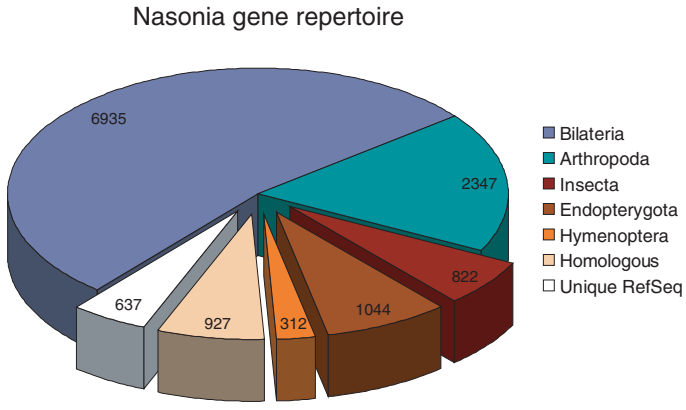

Fig. 3. Distribution of recognizable *Nasonia* orthologs and *Nasonia*-specific genes among gene models with expression sequencing support (*6*).

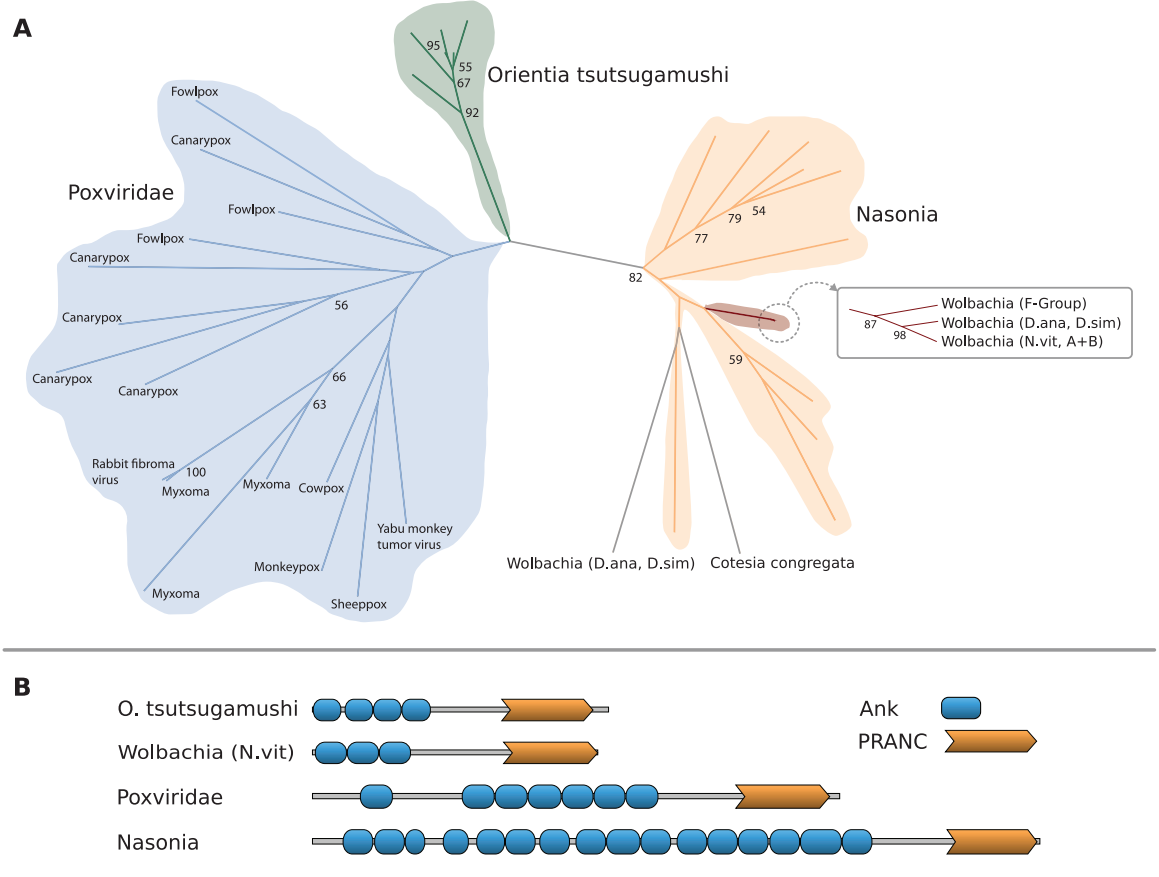

Fig. 4. PRANC domain proteins in *Nasonia*, Pox viruses, and *Wolbachia*. (**A**) Maximum-likelihood tree of PRANC-domain sequences found in Pox viruses, rickettsia (*Wolbachia* and *Orientia*), and parasitoids (*N. vitripennis* and *Cotesia congregata*). The tree was estimated using RaxML with 1000 bootstrap replicates and model settings estimated by ProtTest [see (*6*); alignment deposited in Treebase with ID SN4709]. Bootstrap values above 50% are shown by the corresponding nodes. The phylogenetic relationships suggest lateral transfer from *Wolbachia* to the *Nasonia* lineage. (**B**) Representative domain arrangements for ANK-PRANC proteins.

Downloaded from www.sciencemag.org on January 15, 2010

genes with venom features and proteomic analyses of venom reservoir tissues have uncovered a rich assemblage of 79 candidate venom proteins (table S16) (*41*). Some *Nasonia* venom reservoir proteins belong to previously known insect venom families such as serine proteases; however, nearly half were not related to any known insect venoms. As expected, many of these venom candidates show highly elevated expression in the female reproductive tract, which includes the venom glands and reservoirs. Venom genes also showed significantly higher *dN*/*dS* ratios between *N. vitripennis* and *N. giraulti* than nonvenom genes did (Mann-Whitney U test, $P < 2 \times 10^{-6}$), suggesting that changes in host use between the species may be accompanied by rapid evolution of venom proteins. The large venom protein set found in *Nasonia* with diverse physiological effects (*40*) and abundance of parasitoid species (*1*, *2*) suggests that parasitoids may contain a rich venom pharmacopeia of potential new drugs.

N. vitripennis is a generalist parasitoid with a wide host utilization of many fly species, whereas the other *Nasonia* species are specialists (*4*, *10*). Using genomic tools, a major host preference locus has been mapped to a region of ~2 cM (*10*). Other genes in the *Nasonia* genome that are potentially involved in host finding include odorant binding proteins (table S17) and chemoreceptors (*42*), which show expansions, contractions, and pseudogenization, indicative of rapid turnover.

A suite of genetic tools and resources is available or under development for the *Nasonia* system (*4*, *6*, *11*, *28*), and the genome resources presented here can be used for fine-scale mapping (*6*, *9-11*) and positional cloning (*8*) of QTLs. By combining haploid genetics, ease of rearing, short generation time, systemic RNAi, interfertile species, and new genome resources for three species, *Nasonia* shows promise as a genetic model system for evolutionary and developmental genetics. Genome resources described here and our resulting enhanced understanding of parasitoid biology will also open avenues for improving parasitoid utility in biological control of pests of agricultural and medical importance.

References and Notes

1. D. L. J. Quicke, *Parasitic Wasps* (Chapman & Hall, London, 1997).
2. J. Heraty, in *Insect Biodiversity: Science and Society*, R. Tootti and P. Alder, Eds. (Wiley-Blackwell, Hoboken, NJ, 2009), pp. 445–462.
3. R. Raychoudhury *et al., Heredity* 10.1038/hdy.2009.147 (2010).
4. J. H. Werren, D. Loehlin, *Cold Spring Harb. Protocols* 10.1101/pdb.emo134 (2009).
5. M. A. Pultz *et al.*, *Genetics* **154**, 1213 (2000).
6. Materials and methods and supplementary text are available as supporting material on *Science* Online.
7. J. A. J. Breeuwer, J. H. Werren, *Evolution* **49**, 705 (1995).
8. D. W. Loehlin *et al., PLoS Genet.* 10.1371/journal.pgen.1000821 (2010).
9. D. W. Loehlin, L. S. Enders, J. H. Werren, *Heredity* 10.1038/hdy.2009.146 (2010).
10. C. A. Desjardins, F. Perfecti, J. D. Bartos, L. S. Enders, J. H. Werren, *Heredity* 10.1038/hdy.2009.145 (2010).
11. B. J. Velthuis, W. Yang, T. van Opijnen, J. H. Werren, *Anim. Behav.* **69**, 1107 (2005).
12. O. Niehuis *et al.*, *PLoS ONE* 10.1371/journal.pone.0008597 (2010).
13. L. Wilfert, J. Gadau, P. Schmid-Hempel, *Heredity* **98**, 189 (2007).
14. E. M. Zdobnov, P. Bork, *Trends Genet.* **23**, 16 (2007).
15. The official gene set OGS v1.2 is available at http://nasoniabase.org/nasonia_genome_consortium/datasets.html.
16. Honeybee Genome Sequencing Consortium, *Nature* **443**, 931 (2006).
17. C. D. Smith *et al.*, *Gene* **389**, 1 (2007).
18. L. Viljakainen, D. C. S. G. Oliveira, J. H. Werren, S. K. Behura, *Insect Mol. Biol.* **19**, 27 (2010).
19. B. A. Pannebakker, O. Niehuis, A. Hedley, J. Gadau, D. M. Shuker, *Insect Mol. Biol.* **19**, 91 (2010).
20. T. P. Jurkowski *et al.*, *RNA* **14**, 1663 (2008).
21. R. Kucharski, J. Maleszka, S. Foret, R. Maleszka, *Science* **319**, 1827 (2008).
22. L. W. Beukeboom, A. Kamping, L. van de Zande, *Semin. Cell Dev. Biol.* **18**, 371 (2007).
23. N. Elango, B. G. Hunt, M. A. D. Goodisman, S. V. Yi, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11206 (2009).
24. L. K. Mosavi, T. J. Cammett, D. C. Desrosiers, Z. Y. Peng, *Protein Sci.* **13**, 1435 (2004).
25. J. A. Lynch, C. Desplan, *Nat. Protoc.* **1**, 486 (2006).
26. D. Gerlach, E. V. Kriventseva, N. Rahman, C. E. Vejnar, E. M. Zdobnov, *Nucleic Acids Res.* **37**, D111 (2009).
27. J. A. Lynch, A. E. Brent, D. S. Leaf, M. A. Pultz, C. Desplan, *Nature* **439**, 728 (2006).
28. A. E. Brent, G. Yucel, S. Small, C. Desplan, *Science* **315**, 1841 (2007).
29. M. D. Drapeau, S. Albert, R. Kucharski, C. Prusko, R. Maleszka, *Genome Res.* **16**, 1385 (2006).
30. N. Yapici, Y. J. Kim, C. Ribeiro, B. J. Dickson, *Nature* **451**, 33 (2008).
31. O. Niehuis, A. K. Judson, J. Gadau, *Genetics* **178**, 413 (2008).
32. J. M. Flowers *et al.*, *Mol. Biol. Evol.* **24**, 1347 (2007).
33. J. van den Assem, J. Visser, *Biol. Comport.* **1**, 37 (1976).
34. D. C. S. G. Oliveira, R. Raychoudhury, D. V. Lavrov, J. H. Werren, *Mol. Biol. Evol.* **25**, 2167 (2008).
35. D. M. Rand, R. A. Haney, A. J. Fry, *Trends Ecol. Evol.* **19**, 645 (2004).
36. C. F. Aquadro, V. B. Dumont, F. A. Reed, *Curr. Opin. Genet. Dev.* **11**, 627 (2001).
37. J. C. Dunning-Hotopp *et al.*, *Science* **317**, 1753 (2007).
38. S. J. Chang *et al.*, *J. Virol.* **83**, 4140 (2009).
39. M. Kanehisa *et al.*, *Nucleic Acids Res.* **36**, D480 (2008).
40. D. B. Rivers, Y. A. Yoder, L. Ruggiero, *Trends Entomol.* **2**, 1 (1999).
41. D. C. de Graaf *et al.*, *Insect Mol. Biol.* **19**, 11 (2010).
42. H. M. Robertson, J. Gadau, K. W. Wanner, *Insect Mol. Biol.* **19**, 121 (2010).

43. Genome sequencing, assembly and annotation were funded by the National Human Genome Research Institute (NHGRI U54 HG003273). The whole-genome shotgun project has been deposited at the DNA Databank of Japan (DDBJ)/European Molecular Biology Laboratory (EMBL)/GenBank under accession numbers AAZX00000000 (*N. vitripennis*), ADAO00000000 (*N. giraulti*), and ADAP00000000 (*N. longicornis*). Additional support, acknowledgments, and accession numbers are provided in the supporting online material.

Author List: The *Nasonia* Genome Working Group

John H. Werren,[1]*† Stephen Richards,[2]*† Christopher A. Desjardins,[1] Oliver Niehuis,[3]‡ Jürgen Gadau,[3] John K. Colbourne,[4] Leo W. Beukeboom,[5] Claude Desplan,[6] Christine G. Elsik,[7] Cornelis J. P. Grimmelikhuijzen,[8] Paul Kitts,[9] Jeremy A. Lynch,[10] Terence Murphy,[9] Deodoro C. S. G. Oliveira,[1]§ Christopher D. Smith,[11,12] Louis van de Zande,[5] Kim C. Worley,[2] Evgeny M. Zdobnov,[13,14,15] Maarten Aerts,[16] Stefan Albert,[17] Victor H. Anaya,[18] Juan M. Anzola,[19] Angel R. Barchuk,[20] Susanta K. Behura,[21] Agata N. Bera,[22] May R. Berenbaum,[23] Rinaldo C. Bertossa,[24] Márcia M. G. Bitondi,[25] Seth R. Bordenstein,[26,27] Peer Bork,[28] Erich Bornberg-Bauer,[29] Marleen Brunain,[30] Giuseppe Cazzamali,[8] Lesley Chaboub,[2] Joseph Chacko,[2] Dean Chavez,[2] Christopher P. Childers,[7] Jeong-Hyeon Choi,[4] Michael E. Clark,[1] Charles Claudianos,[31] Rochelle A. Clinton,[32] Andrew G. Cree,[2] Alexandre S. Cristino,[31,33] Phat M. Dang,[34] Alistair C. Darby,[35] Dirk C. de Graaf,[30] Bart Devreese,[16] Huyen H. Dinh,[2] Rachel Edwards,[1] Navin Elango,[36] Eran Elhaik,[37] Olga Ermolaeva,[9] Jay D. Evans,[38] Sylvain Foret,[39] Gerald R. Fowler,[2] Daniel Gerlach,[13,14] Joshua D. Gibson,[3] Donald G. Gilbert,[40] Dan Graur,[37] Stefan Gründer,[41] Darren E. Hagen,[7] Yi Han,[2] Frank Hauser,[8] Da Hultmark,[42] Henry C. Hunter IV,[11] Gregory D. D. Hurst,[35] Shalini N. Jhangian,[2] Huaiyang Jiang,[2] Reed M. Johnson,[43] Andrew K. Jones,[22] Thomas Junier,[13] Tatsuhiko Kadowaki,[44] Albert Kamping,[5] Yuri Kapustin,[9] Bobak Kechavarzi,[45] Jaebum Kim,[46] Jay Kim,[11] Boris Kiryutin,[9] Tosca Koevoets,[5] Christie L. Kovar,[2] Evgenia V. Kriventseva,[47] Robert Kucharski,[48] Heewook Lee,[45] Sandra L. Lee,[2] Kristin Lees,[22] Lora R. Lewis,[2] David W. Loehlin,[1] John M. Logsdon Jr.,[49] Jacqueline A. Lopez,[4] Ryan J. Lozado,[2] Donna Maglott,[9] Ryszard Maleszka,[48] Anoop Mayampurath,[45] Danielle J. Mazur,[49] Marcella A. McClure,[32] Andrew D. Moore,[29] Margaret B. Morgan,[2] Jean Muller,[28] Monica C. Munoz-Torres,[7,50] Donna M. Muzny,[2] Lynne V. Nazareth,[2] Susanne Neupert,[51] Ngoc B. Nguyen,[2] Francis M. F. Nunes,[25,52] John G. Oakeshott,[53] Geoffrey O. Okwuonu,[2] Bart A. Pannebakker,[5,54] Vikas R. Pejaver,[45] Zuogang Peng,[36] Stephen C. Pratt,[3] Reinhard Predel,[51] Ling-Ling Pu,[2] Hilary Ranson,[55] Rhitoban Raychoudhury,[1] Andreas Rechtsteiner,[4,56] Justin T. Reese,[7,57] Jeffrey G. Reid,[2] Megan Riddle,[58]‖ Hugh M. Robertson,[23] Jeanne Romero-Severson,[59] Miriam Rosenberg,[6] Timothy B. Sackton,[60] David B. Sattelle,[22] Helge Schlüns,[61] Thomas Schmitt,[62] Martina Schneider,[8] Andreas Schüler,[29] Andrew M. Schurko,[49] David M. Shuker,[63] Zilá L. P. Simões,[25] Saurabh Sinha,[46] Zachary Smith,[4] Victor Solovyev,[64] Alexandre Souvorov,[9] Andreas Springauf,[41] Elisabeth Stafflinger,[8] Deborah E. Stage,[1] Mario Stanke,[65] Yoshiaki Tanaka,[66] Arndt Telschow,[29] Carol Trent,[58] Selina Vattathil,[2]¶ Eveline C. Verhulst,[5] Lumi Viljakainen,[67] Kevin W. Wanner,[68] Robert M. Waterhouse,[15] James B. Whitfield,[23] Timothy E. Wilkes,[35] Michael Williamson,[8] Judith H. Willis,[69] Florian Wolschin,[70,3] Stefan Wyder,[13] Takuji Yamada,[28] Soojin V. Yi,[36] Courtney N. Zecher,[27] Lan Zhang,[2] Richard A. Gibbs[2]

[1]Department of Biology, University of Rochester, Rochester, NY 14627, USA. [2]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030, USA. [3]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA. [4]The Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN 47405, USA. [5]Evolutionary Genetics–Centre for Ecological and Evolutionary Studies, University of Groningen, 9750 AA Haren, Netherlands. [6]Department of Biology, New York University, New York, NY 10003, USA. [7]Department of Biology, Georgetown University, Washington, DC 20057, USA. [8]Center for Comparative and Functional Insect Genomics, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. [9]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA. [10]Institut für Entwicklungsbiologie, Universität zu Köln, 50923 Köln, Germany. [11]Department of Biology, San Francisco State University, San Francisco, CA 94132, USA. [12]Drosophila Heterochromatin Genome Project, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [13]Department of Genetic Medicine and Development, University of Geneva Medical School, CH-1211 Geneva, Switzerland. [14]Swiss Institute of Bioinformatics, CH-1211 Geneva, Switzerland. [15]Imperial College London, London, SW7 2AZ, UK. [16]Laboratory of Protein Biochemistry and Biomolecular Engineering, Ghent University, B-9000 Ghent, Belgium. [17]BEEgroup and Institute of Pharmaceutical Biology, University of Würzburg, 97082 Würzburg, Germany. [18]Institute for Theoretical Biology, Humboldt University Berlin, 10115 Berlin, Germany. [19]Animal Science and Biology, Texas A&M University, College Station, TX 77843, USA. [20]Departamento de Ciências Biomédicas, Universidade Federal de Alfenas, Alfenas, Minas Gerais, 37130-000, Brazil. [21]Eck Institute for Global Health, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA. [22]Medical Research Council Functional Genomics Unit, Department of Physiology Anatomy and Genetics, University of Oxford, Oxford OX1 3QX, UK. [23]Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [24]Chronobiology–Centre for Behavior and Neurosciences, University of Groningen, 9750 AA Haren, Netherlands. [25]Faculdade de Filosofia,

Downloaded from www.sciencemag.org on January 15, 2010

Ciências e Letras de Ribeirão Preto, Departamento de Biologia, Universidade de São Paulo, Ribeirão Preto, São Paolo 14040-901, Brazil. [26]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA. [27]Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA 02536, USA. [28]European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [29]Institute for Evolution and Biodiversity, University of Münster, 48143 Münster, Germany. [30]Laboratory of Zoophysiology, Ghent University, B-9000 Ghent, Belgium. [31]The Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia. [32]Department of Microbiology and the Center for Computational Biology, Montana State University, Bozeman, MT 59715, USA. [33]Instituto de Física de São Carlos, Departamento de Física e Informática, Universidade de São Paulo, São Carlos, São Paolo 13560-970, Brazil. [34]Subtropical Insects Research Unit, United States Department of Agriculture–Agricultural Research Service (USDA-ARS), U.S. Horticultural Research Lab, Fort Pierce, FL 34945, USA. [35]School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, UK. [36]School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA. [37]Department of Biology and Biochemistry, University of Houston, Houston, TX 77204, USA. [38]Bee Research Lab, USDA-ARS, Beltsville, MD, 20705, USA. [39]Australian Research Council Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, Queensland 4811, Australia. [40]Department of Biology, Indiana University, Bloomington, IN 47405, USA. [41]Institute of Physiology, Rheinisch-Westfaelische Technische Hochschule (RWTH) Aachen University, D-52074 Aachen, Germany. [42]Department of Molecular Biology, Umeå University, S-901 87 Umeå, Sweden. [43]Department of Entomology, University of Nebraska, Lincoln, NE 68583, USA. [44]Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya 464-8601, Japan. [45]School of Informatics, Indiana University, Bloomington, IN 47405, USA. [46]Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [47]Department of Structural Biology and Bioinformatics, University of Geneva Medical School, CH-1211 Geneva, Switzerland. [48]Research School of Biology, Australian National University, Canberra, Australian Capital Territory 2601, Australia. [49]Roy J. Carver Center for Comparative Genomics and Department of Biology, University of Iowa, Iowa City, IA 52242, USA. [50]Department of Genetics and Biochemistry, Clemson University, Clemson, SC 29634, USA. [51]Institute of General Zoology, University of Jena, D-7743 Jena, Germany. [52]Faculdade de Medicina de Ribeirão Preto, Departamento de Genética, Universidade de São Paulo, Ribeirão Preto, São Paolo 14049-900, Brazil. [53]Department of Entomology, Commonwealth Scientific and Industrial Research Organisation, Canberra, Australian Capital Territory 2601, Australia. [54]Institute of Evolutionary Biology–School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, UK. [55]Vector Group, Liverpool School of Tropical Medicine, Liverpool L3 5QA, UK. [56]Department of Molecular, Cell, and Developmental Biology, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. [57]Reese Consulting, 157/10 Tambon Ban Deau, Amphur Muang, Nong Khai, 43000, Thailand. [58]Department of Biology, Western Washington University, Bellingham, WA 98225, USA. [59]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA. [60]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA. [61]School of Marine and Tropical Biology and Centre for Comparative Genomics, James Cook University, Townsville, Queensland 4811, Australia. [62]Department of Evolutionary Biology and Animal Ecology, University of Freiburg, 79104 Freiburg, Germany. [63]School of Biology, University of St Andrews, St Andrews KY16 9TH, UK. [64]Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. [65]Institut für Mikrobiologie und Genetik, Universität Göttingen, 37077 Göttingen, Germany. [66]Division of Insect Sciences, National Institute of Agrobiological Science, Tsukuba, Ibaraki 305-8634, Japan. [67]Department of Biology and Biocenter Oulu, University of Oulu, 90014 Oulu, Finland. [68]Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT 59717, USA. [69]Department of Cellular Biology, University of Georgia, Athens, GA 30602, USA. [70]Department of Biotechnology, Chemistry, and Food Science, Norwegian University of Life Sciences, N-1432 Ås, Norway.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: werr@mail.rochester.edu (J.H.W.); stephenr@bcm.tmc.edu (S.R.)
‡Current address: Verhaltensbiologie, Universität Osnabrück, 49076 Osnabrück, Germany.
§Current address: Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 8193 Bellaterra, Spain.
||Current address: Weill Cornell Medical College, New York, NY 10065, USA.
¶Current address: Department of Epidemiology, University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030, USA.

Supporting Online Material
www.sciencemag.org/cgi/content/full/327/5963/343/DC1
Materials and Methods
SOM Text
Figs. S1 to S25
Tables S1 to S57
References

22 June 2009; accepted 24 November 2009
10.1126/science.1178028

Zebrafish Behavioral Profiling Links Drugs to Biological Targets and Rest/Wake Regulation

Jason Rihel,[1]*† David A. Prober,[1]*‡ Anthony Arvanites,[2] Kelvin Lam,[2] Steven Zimmerman,[1] Sumin Jang,[1] Stephen J. Haggarty,[3,4,5] David Kokel,[6] Lee L. Rubin,[2] Randall T. Peterson,[3,6,7] Alexander F. Schier[1,2,3,8,9]†

A major obstacle for the discovery of psychoactive drugs is the inability to predict how small molecules will alter complex behaviors. We report the development and application of a high-throughput, quantitative screen for drugs that alter the behavior of larval zebrafish. We found that the multidimensional nature of observed phenotypes enabled the hierarchical clustering of molecules according to shared behaviors. Behavioral profiling revealed conserved functions of psychotropic molecules and predicted the mechanisms of action of poorly characterized compounds. In addition, behavioral profiling implicated new factors such as ether-a-go-go–related gene (ERG) potassium channels and immunomodulators in the control of rest and locomotor activity. These results demonstrate the power of high-throughput behavioral profiling in zebrafish to discover and characterize psychotropic drugs and to dissect the pharmacology of complex behaviors.

Most current drug discovery efforts focus on simple in vitro screening assays. Although such screens can be successful, they cannot recreate the complex network interactions of whole organisms. These limitations are particularly acute for psychotropic drugs because brain activity cannot be modeled in vitro (*1*–*3*). Motivated by recent small-molecule screens that probed zebrafish developmental processes (*4*–*7*), we developed a whole organism, high-throughput screen for small molecules that alter larval zebrafish locomotor behavior. We used an automated rest/wake behavioral assay (*3*, *8*) to monitor the activity of larvae exposed to small molecules at 10 to 30 μM for 3 days (Fig. 1A) (*3*). Multiple behavioral parameters were measured, including the number and duration of rest bouts, rest latency, and waking activity (i.e., activity not including time spent at rest) (Fig. 1B) (*3*). We screened 5648 compounds representing 3968 unique structures and 1680 duplicates and recorded more than 60,000 behavioral profiles. Of these, 547 compounds representing 463 unique structures significantly altered behavior relative to controls, according to a stringent statistical cutoff (*3*).

Because the alterations in behavior were multidimensional and quantitative, we assigned a behavioral fingerprint to each compound and applied clustering algorithms to organize molecules according to their fingerprints (Fig. 2A and figs. S1 to S3). This analysis organized the data set broadly into arousing and sedating compounds and identified multiple clusters corresponding to specific phenotypes (Fig. 2, B to F; Fig. 3, A to C; Fig. 4, B and C; and figs. S1 to S4). Clustering allowed us to address three questions: (i) Do structural, functional, and behavioral profiles overlap? (ii) Does the data set predict links between known and unknown small molecules and their mechanisms of action? (iii) Does the data set identify unexpected

[1]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. [2]Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [5]Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA. [6]Developmental Biology Laboratory, Cardiovascular Research Center, Massachusetts General Hospital, Charlestown, MA 02129, USA. [7]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA. [8]Division of Sleep Medicine, Harvard Medical School, Boston, MA 02215, USA. [9]Center for Brain Science, Harvard University, Cambridge, MA 02138, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: schier@fas.harvard.edu (A.F.S.); rihel@fas.harvard.edu (J.R.)
‡Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

Downloaded from www.sciencemag.org on January 15, 2010