The genome of the fire ant *Solenopsis invicta*

Yannick Wurm[a,b,1], John Wang[a,c], Oksana Riba-Grognuz[a,b], Miguel Corona[a], Sanne Nygaard[d], Brendan G. Hunt[e], Krista K. Ingram[f], Laurent Falquet[b], Mingkwan Nipitwattanaphon[a], Dietrich Gotzek[g], Michiel B. Dijkstra[a], Jan Oettler[h], Fabien Comtesse[a], Cheng-Jen Shih[i], Wen-Jer Wu[i], Chin-Cheng Yang[i], Jerome Thomas[j], Emmanuel Beaudoing[j], Sylvain Pradervand[j], Volker Flegel[b], Erin D. Cook[e], Roberto Fabbretti[b], Heinz Stockinger[b], Li Long[b], William G. Farmerie[k], Jane Oakey[l], Jacobus J. Boomsma[d], Pekka Pamilo[m], Soojin V. Yi[e], Jürgen Heinze[h], Michael A. D. Goodisman[e], Laurent Farinelli[n], Keith Harshman[j], Nicolas Hulo[o], Lorenzo Cerutti[o], Ioannis Xenarios[b,o,2], DeWayne Shoemaker[p,2], and Laurent Keller[a,2]

[a]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; [b]Vital-IT Group, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; [c]Biodiversity Research Center, Academia Sinica, Nangang Taipei 115, Taiwan; [d]Centre for Social Evolution, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark; [e]Georgia Institute of Technology, School of Biology, Atlanta, GA 30332-0230; [f]Department of Biology, Colgate University, Hamilton, NY 13346; [g]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012; [h]Biologie I, Universität Regensburg, 93040 Regensburg, Germany; [i]Department of Entomology, National Taiwan University, Taipei 10617, Taiwan; [j]Lausanne Genomic Technologies Facility, University of Lausanne, 1015 Lausanne, Switzerland; [k]Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, FL 32610; [l]Biosecurity Queensland, Brisbane QLD 4108, Queensland, Australia; [m]Department of Biosciences, University of Helsinki, 00014, Helsinki, Finland; [n]Fasteris SA, 1228 Plan-les-Ouates, Switzerland; [o]Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland; and [p]US Department of Agriculture Agricultural Research Service, Center for Medical, Agricultural, and Veterinary Entomology, Gainesville, FL 32608

Ants have evolved very complex societies and are key ecosystem members. Some ants, such as the fire ant *Solenopsis invicta*, are also major pests. Here, we present a draft genome of *S. invicta*, assembled from Roche 454 and Illumina sequencing reads obtained from a focal haploid male and his brothers. We used comparative genomic methods to obtain insight into the unique features of the *S. invicta* genome. For example, we found that this genome harbors four adjacent copies of vitellogenin. A phylogenetic analysis revealed that an ancestral vitellogenin gene first underwent a duplication that was followed by possibly independent duplications of each of the daughter vitellogenins. The vitellogenin genes have undergone subfunctionalization with queen- and worker-specific expression, possibly reflecting differential selection acting on the queen and worker castes. Additionally, we identified more than 400 putative olfactory receptors of which at least 297 are intact. This represents the largest repertoire reported so far in insects. *S. invicta* also harbors an expansion of a specific family of lipid-processing genes, two putative orthologs to the *transformer/feminizer* sex differentiation gene, a functional DNA methylation system, and a single putative telomerase ortholog. EST data indicate that this *S. invicta* telomerase ortholog has at least four spliceforms that differ in their use of two sets of mutually exclusive exons. Some of these and other unique aspects of the fire ant genome are likely linked to the complex social behavior of this species.

social insect | caste differences | nonmodel organism | de novo genome assembly

The major organizing principle of societies of bees, wasps, termites, and ants is a reproductive division of labor, whereby one or a few individuals (the queens and males) specialize in reproduction and the majority of individuals (the workers and soldiers) participate in cooperative tasks such as building the nest, collecting food, rearing the young, and defending the colony. This social organization provides numerous advantages and forms the basis for the tremendous ecological success of social insects. For example, they are found in almost every type of terrestrial environment, make up as much as 50% of animal biomass in some habitats, and play crucial roles as predators, pollinators, and soil processors (1).

In addition to being critically important members of many terrestrial ecosystems, many ant species are also highly successful invasive pests. One such notorious invasive ant species is the fire ant, *Solenopsis invicta* (Fig. 1*A*). This species was inadvertently introduced to the southern United States from South America in the 1930s (2, 3). *S. invicta* is now of profound economic impor-

tance, with annual losses to households, businesses, governments, and institutions of $5,000 million across the United States (4). For example, *S. invicta* aggressively uses its very potent sting, inflicting pain and inducing hypersensitivity reactions in humans (Fig. 1*B*). Furthermore, it forms large colonies at high densities, is capable of damaging agricultural machinery, and thus interfering with crop production and harvesting (5, 6). The many existing methods of fire ant control have failed to halt the spread of this exotic species and have hurt its indigenous competitors. There is thus an urgent need to develop effective and safe alternative management techniques (7). Despite extensive quarantine and extermination efforts, *S. invicta* is now also found in many other countries including Australia, China, and Taiwan (8–10).

The past decade has seen the development of several tools for studying ants at the molecular level (11, 12). Although these tools have provided insights into the genetics of caste differentiation (13, 14) and the effects of social context (15, 16), they are somewhat limited, given that they survey only a small subset of the genome. We therefore undertook whole-genome sequencing and de novo assembly of the fire ant *S. invicta*.

Results and Discussion

Genome Assembly. We report the draft sequence of the genome of the fire ant *S. invicta*, obtained by combined Roche 454 and Illumina technologies for a sequencing cost of approximately $230,000. Our assembly strategy was as follows: We first assembled only Illumina 352-bp insert paired-end reads (Table S1*A*) and subsequently chopped up the resulting assembly into sequences equivalent to the length of Roche 454 reads. These artificial reads

Fig. 1. (*A*) *S. invicta* males (larger, with wings) depart on mating flight while workers (smaller, wingless) patrol (photo by Yannick Wurm). (*B*) A fire ant researcher was stung by his study subject (photo by Daniel P. Wojcik, US Department of Agriculture Agricultural Research Service).

then were combined with Roche 454 shotgun-sequenced reads, resulting in 352.7 Mb of assembled data split among 90,231 contigs with an N50 size of 14,674 bp (N50 is the length such that 50% of the assembled sequence lies in blocks of length N50 or greater). Using 8- and 20-kb insert paired-end Roche 454 reads, 31,250 of the contigs were joined to form 10,543 scaffolds with an N50 size of 720,578 bp (Table 1 and Table S1*F*). These scaffolds represent a total of 352.7 Mb of sequence including 41.3 Mb of un-determined "N" bases found within scaffolds between consecutive contigs. The remaining 58,981 contigs that could not be placed in scaffolds represent a total of 43.4 Mb of sequence and are sig-nificantly shorter than those that were placed in scaffolds (maxi-mum size: 2,002 bp). Among these nonscaffolded contigs, 95% were clustered to each other by blastclust as having more than 50% sequence identity over half the sequence length or very sig-nificant similarity (blastn $E < 10^{-30}$) to one or more genome scaffolds. These contigs likely are highly repetitive elements (Fig. S1), consistent with the estimation that 23% of the *S. invicta* ge-nome consists of highly repetitive or foldback elements (17).

We assessed the accuracy and completeness of the *S. invicta* assembly by comparing it with an independently sequenced and assembled set of ESTs putatively representing 12,488 genes. Among these putative gene transcripts, 98.2% mapped to the ge-nome assembly (blastn $E < 10^{-50}$). A total of 580 putative tran-scripts contained two nonoverlapping 200-bp regions that mapped to two different scaffolded contigs. In the 383 cases in which exons from the same putative gene mapped to different contigs within the same scaffold, scaffolding of contigs was consistent with exon order and orientation. In the remaining 197 cases, putative exons map-

ped to contigs from different scaffolds. We manually inspected 50 of these to determine whether there was evidence for the scaf-folds overlapping and whether the 8- and 20-kb insert libraries provided evidence for the scaffolds being adjacent. In 46 of the 50 cases, the scaffolds were probably either adjacent or the smaller scaffold filled a gap in the larger scaffold. Inconsistent mapping of the four remaining putative transcripts possibly reflected problems with EST assembly because three of the four transcripts had highly significant blast similarity to at least two normally unrelated genes. Overall, these results confirm that the genome assembly is essen-tially complete for gene content and that scaffolding is reliable.

Although the scaffolded *S. invicta* sequence represents 352.7 Mb, unresolved repeats bring the Roche 454 Newbler software estimated genome size to 484.2 Mb. This difference is likely due to the difficulty of resolving repeats. Three conflicting estimates of the haploid genome size of *S. invicta* have been reported pre-viously: 591 Mb (0.62 pg) via reassociation kinetics (17), 753.3 Mb via flow cytometry (18), and 463 Mb via Feulgen image analysis densitometry (0.47 pg reported in ref. 19). The latter estimate is most consistent with our results. Discrepancies in genome size estimations have been previously reported (19, 20). Such variation may be due to technical issues, differences in examined cell types, endoparasitic load, transposon activity, or possibly other genetic differences between individuals or populations.

Gene Content. A combination of ab initio, EST-based, and se-quence similarity-based methods generated an official gene set of 16,569 protein-coding genes. There were significant differences in the guanine-cytosine contents of exons (45.0%), introns (30.9%), the 2,000-bp surrounding genes (1,000 bp up- and downstream; 33.5%), and the genome in general (36.1%) (*t* tests, Bonferonni-corrected, all *P* values $<10^{-10}$). These results are similar to those in the honey bee (21). Blastp search of fire ant proteins against protein databases indicate that 47% of *S. invicta* genes have the strongest similarity to apoid sequences, and another 22% have the strongest similarity to *Nasonia* (Fig. 2), which is consistent with ants being more closely related to bees than to chalcidoid wasps (23). An additional 13% of *S. invicta* genes have the highest sim-ilarity to nonhymenopteran sequences, suggesting that they may be evolving slowly in *S. invicta* or have been lost in *Apis mellifera* and *Nasonia*. Finally, 18% of *S. invicta* proteins have no significant similarity ($E > 10^{-5}$) to non-*Solenopsis* sequences in the GenBank nonredundant protein database (25), suggesting that they may be fast-evolving or ant-specific. Similarly, 17% of the proteins in the *Nasonia vitripennis* official gene set have no significant similarity to non-*Nasonia* sequences in the nonredundant protein database.

Functional Categories. *S. invicta* appears to harbor a typical insect gene repertoire. For example, the *S. invicta* genome includes a complete set of small RNA-processing genes with orthologs to *Argonaute*, *Drosha*, *Pasha*, *Dicer-1*, *Dicer-2*, *Loquacious*, and *R2D2*. Domain analyses of the *S. invicta*, *N. vitripennis*, *Drosophila melanogaster*, and *A. mellifera* proteomes reveal several putative gene duplications in fire ants (Dataset S1). We highlight here the

Table 1. Genome assembly statistics

	Scaffolds*	Scaffolded contigs	Nonscaffolded contigs	All contigs
Number	10,543	31,250	58,981	90,231
N50 size (bp)	720,578	18,166	983	14,674
Maximum size (bp)	6,355,204	192,021	2,002	192,021
Mean size (bp)	33,452	9,965	735	3,931
Minimum size (bp)	1,997	397	200	200
Total consensus (bp)	352,687,102	311,407,343	43,332,432	354,739,775

*Scaffolds include gaps between adjacent contigs. The estimated lengths of these gaps are included in scaffold size estimations. True sizes likely are slightly different.
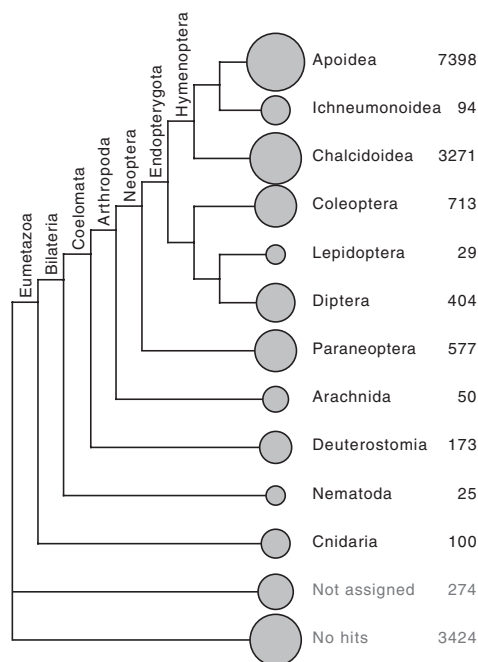
Fig. 2. Taxonomic distribution of best blastp hits of *S. invicta* proteins to the nonredundant (nr) protein database ($E < 10^{-5}$). Results were first plotted using MEGAN software (22) and then branches with fewer than 20 hits were removed, branch lengths were reduced for compactness, and tree topology was adjusted to reflect consensus phylogenies (23, 24).

significance of these duplication events in vitellogenins, odor perception genes, and a family of lipid-processing genes. We also discuss additional features of interest in the fire ant genome relevant to the complex social biology of this species, including sex determination genes, DNA methylation genes, telomerase, and the insulin and juvenile hormone pathways.

Vitellogenins. In contrast to other insects that mainly have only one or two vitellogenins, the fire ant genome harbors four adjacent copies of vitellogenin (*Vg1-4*) (Fig. 3*A*), all of which are at least partially supported by EST reads. A phylogenetic analysis reveals that an ancestral vitellogenin gene first underwent duplication, followed by possibly independent duplications of each of the daughter vitellogenins, thus giving rise to *Vg1* and *Vg4* and to *Vg2* and *Vg3*. All of these duplications occurred after the ancestor of fire ants split from wasps and bees (Fig. 3*B*). The single vitellogenin found in *A. mellifera* is a multifunctional protein (26) involved in the regulation of life span (27, 28) and division of labor (29). Quantitative RT-PCR shows that *Vg1* and *Vg4* are preferentially expressed in workers and *Vg2* and *Vg3* in queens (Fig. 3*C*, *SI Materials and Methods*, and Table S1*G*). Vitellogenin expression in *S. invicta* workers is surprising because they lack ovaries. Given the superorganism properties of ant societies, the expression patterns suggest that vitellogenins underwent neo- or subfunctionalization after duplication to acquire caste-specific functions.

Odor Perception. Consistent with studies in other insects, we find a single *S. invicta* ortholog to *DmOr83b*, a broadly expressed olfactory receptor (OR) required to interact with other ORs for *Drosophila* and *Tribolium castaneum* olfaction (30–32). Beyond *OR83b*, OR number varies greatly between insect species. Blast searches and GeneWise searches using an HMM profile constructed with aligned ORs from *N. vitripennis* (33) and *Pogonomyrmex barbatus* identified more than 400 loci in the *S. invicta* genome with significant sequence similarity to ORs. Preliminary work on gene model reconstruction identified 297 intact full-length proteins. Many *S. invicta* ORs are in tandem arrays (Fig. S2*A*) and derive from recent expansions. *S. invicta* may thus harbor the largest identified insect OR repertoire because there are 10 ORs in *Pediculus humanus* (34), 60 in *Drosophila*, 165 in *A. mellifera*, 225 in *N. vitripennis* (33), and 259 in *T. castaneum* (32). The large numbers of *N. vitripennis* and *T. castaneum* ORs are thought to be due to current or past difficulties in host and food finding. As has been suggested for *A. mellifera* (35), the large number of *S. invicta* ORs may result from the importance of chemical communication in ants. The odorant-binding proteins (OBPs) are another family of genes also known to play roles in chemosensation in *Drosophila* (36). Intriguingly, the social organization of *S. invicta* colonies is completely associated with se-
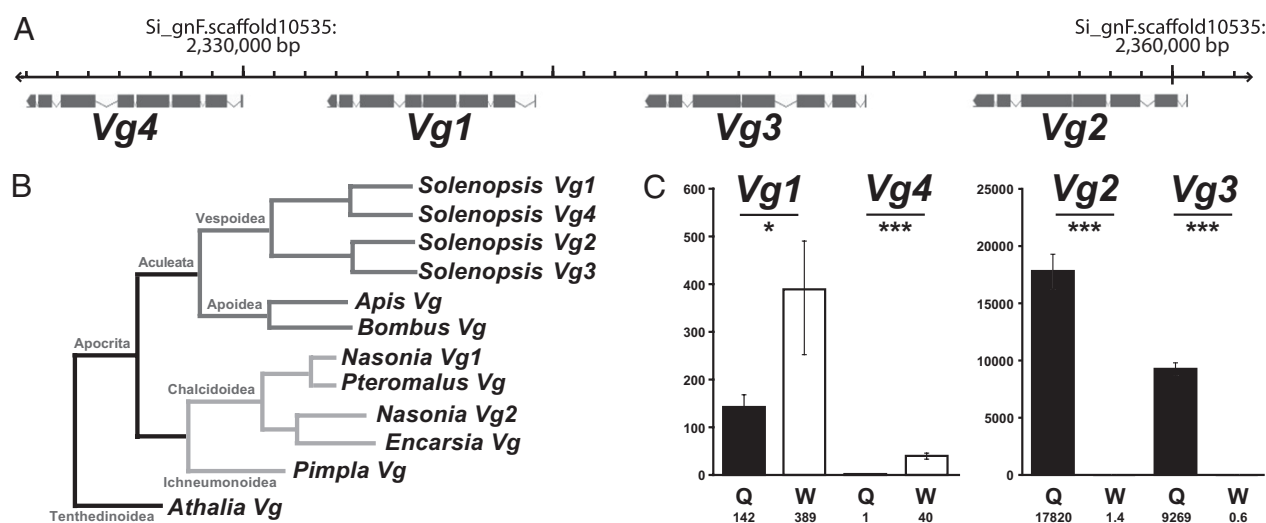

Fig. 3. *S. invicta* vitellogenins. (*A*) Four vitellogenins are located within a single 40,000-bp region of the *S. invicta* genome. (*B*) Parsimony tree of known hymenopteran vitellogenin protein sequences suggests that two rounds of vitellogenin duplication occurred after the split between ants and other hymenopterans including bees and wasps. (*C*) Quantitative RT-PCR of the four putative *S. invicta* vitellogenins on whole bodies of major workers (W) and mated queens (Q) (*n* = 10). The *y* axis indicates mRNA concentrations for the different vitellogenins. Values depicted by each bar are shown below the *x*-axis labels. Error bars represent SEs. Expression differences between queens and workers were significant (Bonferroni-corrected two-tailed *t* tests: *$P < 0.05$, ***$P < 10^{-10}$).

quence variation at the OBP gene *Gp-9* (37, 38). We find 12 OBP domains in the *S. invicta* genome, 2 of which are differentially expressed between workers of alternate *Gp-9* genotypes (15). Further analyses will be required to determine the extent to which these genes are directly involved in the morphological and behavioral differences between queens and workers of the two alternate social organizations of *S. invicta*.

Lipid Metabolism. An unusually high number of genes in *S. invicta* have the following protein domains related to fatty-acid metabolism: Ketoacyl-synt (PF00109), Ketoacylsynt_C (PF02801), and Acyl_transf_1 (PF00698) (Dataset S1). Although some are likely pseudogenes, nine *S. invicta* genes are complete and carry all three domains. In comparison, *A. mellifera* and *D. melanogaster* have only two such genes whereas *N. vitripennis* has six (respective odds ratios: 2.3, 3.4, and 0.8). The expansion of fatty-acid metabolism-related genes in *S. invicta* could stem from the fact that young *S. invicta* queens accumulate as much as 60% of their body mass in the form of lipids within the 2 weeks following eclosion from the pupae (39) as a means of rearing a first worker brood for the duration of a claustral phase during which queens neither feed nor forage (40). Alternatively, such lipid-processing genes may help produce the cuticular hydrocarbons that are involved in kin recognition in ants (41).

Sex Determination. Hymenopterans, including wasps, bees, and ants, have a haplo-diploid sex determination system whereby diploid eggs normally develop into females, and haploid eggs develop into males. In *N. vitripennis*, female development is initiated by maternally derived *transformer/feminizer* mRNA in the zygote (42). In contrast, sex is determined by the *complementary sex determiner* (*csd*) gene in *A. mellifera* (43, 44): Eggs that are heterozygous at this locus develop into females, whereas hemizygous haploids and homozygous diploids develop into males. The *csd* gene is thought to be a recent *Apis* innovation (43), having arisen through a duplication of the *transformer/feminizer* gene. The sex determination mechanism in ants is unknown, but the occurrence of diploid males in some *S. invicta* populations suggests a *csd*-like mechanism (45, 46). The genome of *S. invicta* contains two linked sequences with similarity to *transformer/feminizer* (Fig. S3*A*), but unlike the *A. mellifera* sex-determining locus, the *S. invicta* genes are coded on opposite strands. Members of the Apis transformer/feminizer protein family contain two distinct domains: An N-terminal SDP_N domain and a C-terminal Apis_CSD domain. One of the *S. invicta* sequences (*Tra-A*) contains both domains (SDP_N: $E = 3.5 \times 10^{-11}$; Apis_CSD: $E = 1.4 \times 10^{-7}$). The other (*Tra-B*) contains a partial SDP_N domain ($E = 9.5 \times 10^{-4}$) and a nonsignificant match to Apis_CSD. Alternative splicing of *transformer/feminizer* mRNA is known to play a crucial role in sex determination for many insects (47). Intriguingly, the *S. invicta Tra-B* transcript appears to have two different spliceforms, with only one spliceform including the SDP_N domain. This longer form appears to be the dominant transcript in males, whereas both forms are equally expressed in queens and workers (Fig. S3*B*). A phylogenetic analysis of the *transformer/feminizer* homologs from several hymenopterans shows that the *S. invicta* sequences cluster together, consistent with independent *transformer/feminizer* duplication in the ant and honey bee lineages (Fig. S3*C*).

Methylation. *S. invicta* harbors a complete set of genes known to be involved in DNA methylation, maintenance of methylation patterns, and tRNA methylation in eukaryotes, including DNMT3, DNMT1, and TRDMT1 (previously known as DNMT2) (48). A negative correlation between $CpG_{O/E}$, a statistical measure of DNA methylation, and enrichment of sequence obtained after methylated DNA immunoprecipitation (MeDIP) from a pool of queen and worker prepupae suggested the existence of functional methylation in *S. invicta* (Table S1 *B* and *C*). DNA methylation was confirmed by sequencing of bisulfite-converted genomic DNA from nine genes (*SI Materials and Methods*, Fig. S4, Table S1*D*). DNA methylation is hypothesized to play a key role in developmental responsiveness to environmental factors and may play an important role in social insect caste determination (49). However, the primary targets of DNA methylation in insects appear to be genes with ubiquitous expression across tissues and alternate phenotypes (50–52). Putatively methylated genes identified from MeDIP analysis in *S. invicta* were enriched for biological processes related to cellular metabolism and transcription (*SI Materials and Methods*, Table S1*E*), as is the case with methylated genes in *A. mellifera* (50).

Telomerase Reverse Transcriptase. Ants show remarkable intraspecific life-span variation with queens of some species living to the astonishing age of more than 20 y and workers typically dying within several months or at most a few years and males within a few months (53). Aging is associated with a decline in telomere repair and consequent telomere shortening (54) in many animals. Similarly, in the long-lived ant *Lasius niger* somatic tissues of the short-lived males have dramatically shorter telomeres than those of the much longer-lived queens and workers (55). The telomere sequence of most ants and other insects is composed of TTAGG repeats (56, 57). Consistent with this, the ends of *S. invicta* chromosomes showed strong hybridization signal to labeled $(TTAGG)_n$ probe (Fig. S5). Furthermore, *S. invicta* sequences harbor many more degenerate TTAGG repeats than vertebrate-like TTAGGG repeats. Finally, in contrast with dipteran insect species that lack telomerase, but similarly to other nondipteran insect species whose genomes have been sequenced, *S. invicta* has a single putative telomerase ortholog with RNA-binding (TRBD) and reverse transcriptase (TERT) domains. Interestingly, EST data derived from mixed-stage, mixed-caste, whole-body libraries indicate that this *S. invicta* telomerase ortholog has at least four strongly supported spliceforms that differ in their use of two sets of mutually exclusive exons. These alternative spliceforms may permit fine-tuning of telomerase activity, perhaps in caste- or tissue-specific manners.

Insulin/Insulin-Like Growth Factor Signaling. Insulin and insulin-like growth factor (IGF) signaling is a key integrative pathway regulating aging and fertility in animals (58, 59). In *A. mellifera*, insulin and IGF signaling are involved in the regulation of caste determination (60, 61), division of labor (62), and queen longevity (28) and may play similar roles in other social insects. The family of insulin-like peptides (ILPs) underwent many clade and species-specific duplications, leading to 37 members in *Caenorhabditis elegans*, 27 in *Bombyx mori*, and 7 in *Drosophila* and *Anopheles*. In contrast, *S. invicta* and *A. mellifera* have only two orthologous ILPs, one of which also occurs in *N. vitripennis* (Fig. S2*B*). Both *S. invicta* and *A. mellifera* also have two insulin/IGF1 receptors. Phylogenetic analyses suggest that these two receptors resulted from an ancient duplication with subsequent losses in Diptera and *Nasonia*. Interestingly, we find that one of the putative insulin/IGF1 receptors belongs to our list of genes putatively subjected to dense methylation in *S. invicta* (*SI Materials and Methods*).

Juvenile Hormone. Juvenile Hormone (JH) regulates larval molting and reproductive development in many insects (63). Increases in JH titer have also been shown to induce soldier-caste differentiation in termites (64) and behavioral changes in *A. mellifera* workers (65, 66). Interestingly, *S. invicta* harbors >30 putative juvenile hormone binding protein (JHBPs; PF06585, Dataset S1) encoding genes, more than half of which are located in a single 600,000-bp region. By contrast, there are only 16 such JHBP domains in *Nasonia* and 19 in *A. mellifera*. Similarly, the number of genes that encode juvenile hormone epoxide hydrolases (JHEHs), enzymes that degrade JH, is much higher in *S. invicta* than in

A. mellifera (one) and *Nasonia* (two). Four of the six *S. invicta* JHEH encoding genes are adjacent, suggesting recent duplications. Because JH titer determines fecundity of *S. invicta* queens (65), the expansions of both JHBP and JHEH gene families in *S. invicta* may reflect strong selection occurring after the death of the mother queen with many nonreproductive queens competing to reproduce quickly and become "replacement" queens (40, 67).

In conclusion, this study reveals that a combination of Roche 454 and Illumina sequencing can be used to obtain a good quality genome even when the genome is relatively large and contains a high proportion of repetitive elements. Comparison with other genomes shows that the fire ant genome has many unique properties probably associated with the complex social life of this species. Finally, the sequencing of the fire ant genome provides the foundation for future evolutionary, biomedical, sociogenetic, and pest-management studies of this important pest species and facilitates comparisons with other social species.

Materials and Methods

Computation was performed at the Vital-IT (http://www.vital-it.ch) center for high-performance computing of the Swiss Institute of Bioinformatics. Analyses were assisted by custom Ruby/Bioruby (68, 69), Perl (70), and R (71) scripts. The details of the sequencing, assembly, annotation, and analyses are given in *SI Materials and Methods*.

1. Hölldobler B, Wilson EO (1990) *The Ants* (The Belknap Press of Harvard University Press, Cambridge, MA).
2. Shoemaker DD, DeHeer C, Krieger MJB, Ross KG (2006) Population genetics of the invasive fire ant *Solenopsis invicta* in the U.S.A. *Ann Entomol Soc Am* 99:1213–1233.
3. Ross KG, Shoemaker DD (2008) Estimation of the number of founders of an invasive pest insect population: The fire ant *Solenopsis invicta* in the USA. *Proc Biol Sci* 275:2231–2240.
4. McDonald M (2006) Reds under your feet. *New Sci* 189:50–51.
5. Lard CF, et al. (2006) An economic impact of imported fire ants in the United States of America. PhD thesis (Texas A&M University, College Station, TX).
6. Vinson SB (1986) *Economic Impact and Control of Social Insects* (Praeger, New York).
7. Williams DF, deShazo RD (2004) Biological control of fire ants: An update on new techniques. *Ann Allergy Asthma Immunol* 93:15–22.
8. Henshaw MT, Kunzmann N, Vanderwoude C, Sanetra M, Crozier RH (2005) Population genetics and history of the introduced fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae), in Australia. *Aust J Entomol* 44:37–44.
9. Yang CC, Shoemaker DD, Wu WJ, Shih CJ (2008) Population genetic structure of the red imported fire ant, *Solenopsis invicta*, in Taiwan. *Insectes Soc* 55:54–65.
10. Xu Y-J, Huang J, Lu Y-Y, Zeng L, Liang G-W (2009) Observation of nuptial flights of the red imported fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae) in mainland China. *Sociobiology* 54:831–840.
11. Wurm Y, et al. (2009) Fourmidable: A database for ant genomics. *BMC Genomics* 10:5.
12. Wang J, et al. (2007) An annotated cDNA library and microarray for large-scale gene-expression studies in the ant *Solenopsis invicta*. *Genome Biol* 8:R9.
13. Goodisman MA, Isoe J, Wheeler DE, Wells MA (2005) Evolution of insect metamorphosis: A microarray-based study of larval and adult gene expression in the ant *Camponotus festinatus*. *Evolution* 59:858–870.
14. Gräff J, Jemielity S, Parker JD, Parker KM, Keller L (2007) Differential gene expression between adult queens and workers in the ant *Lasius niger*. *Mol Ecol* 16:675–683.
15. Wang J, Ross KG, Keller L (2008) Genome-wide expression patterns and the genetic architecture of a fundamental social trait. *PLoS Genet* 4:e1000127.
16. Wurm Y, Wang J, Keller L (2010) Changes in reproductive roles are associated with changes in gene expression in fire ant queens. *Mol Ecol* 19:1200–1211.
17. Li J, Heinz KM (2000) Genome complexity and organization in the red imported fire ant *Solenopsis invicta* Buren. *Genet Res* 75:129–135.
18. Johnston JS, Ross LD, Beani L, Hughes DP, Kathirithamby J (2004) Tiny genomes and endoreduplication in Strepsiptera. *Insect Mol Biol* 13:581–585.
19. Ardila-Garcia AM, Umphrey GJ, Gregory TR (2010) An expansion of the genome size dataset for the insect order Hymenoptera, with a first test of parasitism and eusociality as possible constraints. *Insect Mol Biol* 19:337–346.
20. Tsutsui ND, Suarez AV, Spagna JC, Johnston JS (2008) The evolution of genome size in ants. *BMC Evol Biol* 8:64.
21. Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443:931–949.
22. Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomic data. *Genome Res* 17:377–386.
23. Sharkey MJ (2007) Phylogeny and classification of Hymenoptera. *Zootaxa* 1668:521–548.
24. Wiegmann BM, et al. (2009) Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol* 7:34.
25. Pruitt KD, Tatusova T, Maglott DR (2005) NCBI Reference Sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* 33(Database issue):D501–D504.
26. Amdam GV, Norberg K, Hagen A, Omholt SW (2003) Social exploitation of vitellogenin. *Proc Natl Acad Sci USA* 100:1799–1802.
27. Seehuus SC, Norberg K, Gimsa U, Krekling T, Amdam GV (2006) Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci USA* 103:962–967.
28. Corona M, et al. (2007) Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci USA* 104:7128–7133.
29. Nelson CM, Ihle KE, Fondrk MK, Page RE, Amdam GV (2007) The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biol* 5:e62.
30. Larsson MC, et al. (2004) Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* 43:703–714.
31. Benton R, Sachse S, Michnick SW, Vosshall LB (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors *in vivo*. *PLoS Biol* 4:e20.
32. Engsontia P, et al. (2008) The red flour beetle's large nose: An expanded odorant receptor gene family in *Tribolium castaneum*. *Insect Biochem Mol Biol* 38:387–397.
33. Robertson HM, Gadau J, Wanner KW (2010) The insect chemoreceptor superfamily of the parasitoid jewel wasp *Nasonia vitripennis*. *Insect Mol Biol* 19(Suppl 1):121–136.
34. Kirkness EF, et al. (2010) Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc Natl Acad Sci USA* 107:12168–12173.
35. Robertson HM, Wanner KW (2006) The chemoreceptor superfamily in the honey bee, *Apis mellifera*: Expansion of the odorant, but not gustatory, receptor family. *Genome Res* 16:1395–1403.
36. Xu P, Atkinson R, Jones DNM, Smith DP (2005) *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* 45:193–200.
37. Ross KG, Keller L (1998) Genetic control of social organization in an ant. *Proc Natl Acad Sci USA* 95:14232–14237.
38. Keller L, Ross KG (1998) Selfish genes: A green beard in the red fire ant. *Nature* 394:573–575.
39. Keller L, Passera L (1989) Size and fat content of gynes in relation to the mode of colony founding in ants (Hymenoptera; Formicidae). *Oecologia* 80:236–240.
40. Tschinkel WR (2006) *The Fire Ants* (The Belknap Press of Harvard University Press, Cambridge, MA).
41. Blomquist GJ, Bagnères AG (2010) Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology (Cambridge University Press, Cambridge, UK).
42. Verhulst EC, Beukeboom LW, van de Zande L (2010) Maternal control of haplodiploid sex determination in the wasp *Nasonia*. *Science* 328:620–623.
43. Hasselmann M, et al. (2008) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. *Nature* 454:519–522.
44. Gempe T, et al. (2009) Sex determination in honeybees: Two separate mechanisms induce and maintain the female pathway. *PLoS Biol* 7:e1000222.
45. Glancey BM, Romain MKS, Crozier RH (1976) Chromosome numbers of the red and black imported fire ants, *Solenopsis invicta* and *S. richteri*. *Ann Entomol Soc Am* 69:469–470.
46. Ross KG, Fletcher DJC (1985) Genetic origin of male diploidy in the fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae), and its evolutionary significance. *Evolution* 39:888–903.
47. Verhulst EC, van de Zande L, Beukeboom LW (2010) Insect sex determination: It all evolves around *transformer*. *Curr Opin Genet Dev* 20:376–383.

48. Jurkowski TP, et al. (2008) Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *RNA* 14:1663–1670.

49. Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319:1827–1830.

50. Elango N, Hunt BG, Goodisman MAD, Yi SV (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proc Natl Acad Sci USA* 106:11206–11211.

51. Foret S, Kucharski R, Pittelkow Y, Lockett GA, Maleszka R (2009) Epigenetic regulation of the honey bee transcriptome: Unravelling the nature of methylated genes. *BMC Genomics* 10:472.

52. Hunt BG, Brisson JA, Yi SV, Goodisman MAD (2010) Functional conservation of DNA methylation in the pea aphid and the honeybee. *Genome Biol Evol* 2:719–728.

53. Keller L, Genoud M (1997) Extraordinary lifespans in ants: A test of evolutionary theories of ageing. *Nature* 389:958–960.

54. Mariani E, et al. (2003) Different rates of telomere shortening and telomerase activity reduction in CD8 T and CD16 NK lymphocytes with ageing. *Exp Gerontol* 38:653–659.

55. Jemielity S, et al. (2007) Short telomeres in short-lived males: What are the molecular and evolutionary causes? *Aging Cell* 6:225–233.

56. Lorite P, Carrillo JA, Palomeque T (2002) Conservation of (TTAGG)(n) telomeric sequences among ants (Hymenoptera, Formicidae). *J Hered* 93:282–285.

57. Frydrychová R, Grossmann P, Trubac P, Vítková M, Marec F (2004) Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* 47:163–178.

58. Finch CE, Ruvkun G (2001) The genetics of aging. *Annu Rev Genomics Hum Genet* 2: 435–462.

59. Dillin A, Crawford DK, Kenyon C (2002) Timing requirements for insulin/IGF-1 signaling in *C. elegans*. *Science* 298:830–834.

60. Wheeler DE, Buck N, Evans JD (2006) Expression of insulin pathway genes during the period of caste determination in the honey bee, *Apis mellifera*. *Insect Mol Biol* 15: 597–602.

61. de Azevedo SV, Hartfelder K (2008) The insulin signaling pathway in honey bee (*Apis mellifera*) caste development: Differential expression of insulin-like peptides and insulin receptors in queen and worker larvae. *J Insect Physiol* 54:1064–1071.

62. Ament SA, Corona M, Pollock HS, Robinson GE (2008) Insulin signaling is involved in the regulation of worker division of labor in honey bee colonies. *Proc Natl Acad Sci USA* 105:4226–4231.

63. Riddiford LM (2008) Juvenile hormone action: A 2007 perspective. *J Insect Physiol* 54: 895–901.

64. Scharf ME, Buckspan CE, Grzymala TL, Zhou X (2007) Regulation of polyphenic caste differentiation in the termite *Reticulitermes flavipes* by interaction of intrinsic and extrinsic factors. *J Exp Biol* 210:4390–4398.

65. Robinson GE, Vargo EL (1997) Juvenile hormone in adult eusocial Hymenoptera: Gonadotropin and behavioral pacemaker. *Arch Insect Biochem Physiol* 35:559–583.

66. Sullivan JP, Fahrbach SE, Robinson GE, Robinson GE (2000) Juvenile hormone paces behavioral development in the adult worker honey bee. *Horm Behav* 37:1–14.

67. Brent CS, Vargo EL (2003) Changes in juvenile hormone biosynthetic rate and whole body content in maturing virgin queens of *Solenopsis invicta*. *J Insect Physiol* 49:967–974.

68. Aerts J, Law A (2009) An introduction to scripting in Ruby for biologists. *BMC Bioinformatics* 10:221.

69. Goto N, et al. (2010) BioRuby: Bioinformatics software for the Ruby programming language. *Bioinformatics* 26:2617–2619.

70. Stajich JE, et al. (2002) The Bioperl toolkit: Perl modules for the life sciences. *Genome Res* 12:1611–1618.

71. R Development Core Team (2007) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).