Insect. Soc. 56 (2009) 44–48

Research article

Construction and characterization of a BAC-library for a key pollinator, the bumblebee *Bombus terrestris* L.

L. Wilfert [1,5], M. Muñoz Torres [2], C. Reber-Funk [1], R. Schmid-Hempel [1,6], J. Tomkins [3], J. Gadau [4] and P. Schmid-Hempel [1]

[1] *Institute of Integrative Biology (IBZ), ETH Zurich, CH-8092 Zurich, Switzerland; e-mail: lena.wilfert@ed.ac.uk*
[2] *Clemson University Genomics Institute, Clemson University, Clemson, SC 29634, USA*
[3] *Clemson Environmental Genomics Lab, Clemson University, Clemson, SC 29634, USA*
[4] *School of Life Sciences, Arizona State University, Tempe, AZ 85287–4501, USA*
[5] *current address: Institute of Evolutionary Biology, University of Edinburgh, UK*
[6] *contact for data repository of this BAC library*

Received 12 March 2008; revised 13 September 2008; accepted 17 September 2008.
Published Online First 13 October 2008

Abstract. The primitively social bumblebee *Bombus terrestris* is an ecological model species as well as an important agricultural pollinator. As part of the ongoing development of genomic resources for this model organism, we have constructed a publicly available bacterial artificial chromosome (BAC) library from males of a field-derived colony. We have shown that this library has a high coverage, which allows any particular sequence to be retrieved from at least one clone with a probability of 99.7 %. We have further demonstrated the library's usefulness by successfully screening it with probes derived both from previously described *B. terrestris* genes and candidate genes from another bumblebee species and the honeybee. This library will facilitate genomic studies in *B. terrestris* and will allow for novel comparative studies in the social Hymenoptera.

Keywords: Genomics, gene discovery, breeding, sex determination, pollination.

Introduction

Many species of social insects are not only commercially important but have become model species for a wide range of basic questions such as social evolution, caste determination, sex ratio strategies, foraging behaviour, social parasitism, ecological physiology, sensory ecology, evolutionary parasitology, ecological immunology, as well as pollination and community ecology. Bumblebees have served as subjects of study in all of these aspects (e.g. Goulson, 2003). Indeed, only the honeybee appears to rival the breadth of research done on this group. Despite the significant contributions that social insects have made, the genomic resources – so common and well developed for many other groups of insects or vertebrates – are quite limited. With a few exceptions (Hoffman and Goodisman, 2007; Wang et al., 2007), important genomic resources (genome sequence, libraries of genomic DNA and expressed sequences) are available only for the honeybee. Here, we report on a substantial genomic resource for a social insect outside the honeybee – a high-coverage bacterial artificial chromosome (BAC) library of genomic DNA for the bumblebee *Bombus terrestris* L.

B. terrestris is a common European bumblebee species that is of importance as an ecological and evolutionary model organism (Goulson, 2003). The genomic tools promise to open up new avenues of research for ecological study organisms such as the bumblebee. For example, the evolution of sociality and caste determination is a research topic that has recently benefited from the application of genomic tools. Pereboom et al. (2005) have identified genes that are differentially expressed in the development of queens and workers. This allows for comparative studies of caste determination on a detailed molecular level. Much recent research on bumblebees has focused on ecological immunity (Sadd and Schmid-Hempel, 2006) and host-parasite interactions (i.e. Baer and Schmid-Hempel, 2001; Ruiz-Gonzalez and Brown, 2006). This research has been complemented by quantitative genetic studies identifying quantitative trait loci (QTL) explaining a part of the phenotypic variation of fitness-relevant traits such as the strength of innate immune mechanisms (Wilfert et al., 2007b) and susceptibility to a protozoan parasite (Wilfert et al., 2007a). Such

Table 1. Hybridization results of six gene probes using the *Bombus terrestris* BAC library. Probes were generated by PCR. Positive clones were fingerprinted by *Hind*III and assembled into contigs.

Gene probe/Accession Number	Primer	No. of positive clones	Contig	No. of clones in contig
Arginine Kinase (Kawakita et al., 2003) AF492888	F-GTT GAC CAA GCY GTY TTG GA R-CAT GGA AAT AAT ACG RAG RTG	52	A	17
Defensin AY425599	F-GTG GCT CTT CTC TTT GTG GCT G R-CAC TCT TCT TTG TCT ATC GGC ACG	21	B	8
Dscam (Graveley et al., 2004) AY686596	F-TTG GCT TTC ACT TCT GGC GG R-TGC GGT CCA CTT CCT TGA TG	27	C	22
Elongation Factor 1 alpha (Kawakita et al., 2003) AF492955	F-GGA CAC AGA GAT TTC ATC AAR AA R-TTG CAA AGC TTC RTG RTG CAT TT	27	D	6
Long-wave Rhodopsin (Mardulyn and Cameron, 1999) AF091722	F-AAT TGC TAT TAY GAR ACN TGG GT R- ATA TGG AGT CCA NGC CAT RAA CCA	12	E	10
Relish XM_624623	F-TGG ACG CTT TTC AGA ATT GG R-GAG CTT CCA GAA TGA GAT ATT CG	26	A	17

tools raise the opportunity to study the maintenance of genetic variation for fitness-relevant traits involved in host-parasite interactions in natural populations (Schmid-Hempel, 2001). To comprehensively test such hypotheses by studying the molecular signature of evolution, we need to identify the genes underlying the quantitative variation in host resistance.

Identifying single genes in turn requires access to the physical genome. The genetic map used to identify QTLs, by contrast, is based on recombination distances. In *B. terrestris*, one centimorgan of genetic distance represents an average of 226 Kb of the physical genome (Wilfert et al., 2006). To integrate these two approaches, large-insert libraries such as bacterial artificial chromosome (BAC) libraries are extremely valuable tools because they allow the physical mapping of genes based on information from genetic linkage maps, expressed sequences or heterologous candidate genes. As previous studies have shown, this joint approach is very successful. For example, the gene underlying the sex determination locus has been isolated with the help of a BAC-library in the honeybee (Tomkins et al., 2002; Beye et al., 2003). In order to facilitate the identification and cloning of genes, and to eventually facilitate physical mapping and genome assembly, we have here constructed a high coverage BAC-library of *B. terrestris* as one of the first such tools outside of *Apis*.

Materials and methods

High-molecular weight DNA from fresh haploid male pupae was prepared following a procedure adapted for honeybees (Tomkins et al., 2002). DNA was partially digested with the restriction enzyme *Hind*III. Size selection was performed on the fragments via two consecutive rounds of pulse field electrophoresis (PFGE). Fragments were then ligated into the vector pIndigoBAC 536 (Peterson et al., 2000). Vectors were transformed into *E. coli* DH10B cells using electroporation. Recombinant colonies were picked using a Genetix Q-bot and stored individually in 384 well plates at -80° C. Additionally high-density colony filters for hybridization-based screening of the BAC library were prepared using a Genetix Q-Bot. Clones were arrayed in double spots using a 4 x 4 array with 6 fields, on 11.5 x 22.5 cm – Hybond N+ filters (Amersham). This pattern allows 18,432 clones to be represented per filter. Colony filters were grown and processed using standard techniques for alkaline lysis (Sambrook et al., 1989).

Library screening

To screen the library, we developed PCR products from three *B. terrestris*-specific sequences (*Arginine Kinase, Elongation Factor 1 Alpha* and *Longwave-Rhodopsin*), one from a sequence from *B. ignitus* (*Defensin*) and two from *Apis mellifera* genes (*Relish* and *Dscam*) and used them as probes for hybridization. Primers and accession numbers of these genes are detailed in Table 1. The PCR products obtained were 200 – 400 bp in length; approximately 100 ng of product were labeled independently using 30uCi of alpha-P32-dCTP following manufacturer's instructions with the DECAprime II random priming DNA labeling kit (Ambion, Inc. ABI, Foster City, CA, USA). Hybridization of colony filters was performed using standard techniques (Sambrook et al., 1989) with the following modifications: hybridizations were performed for at least 16 h at 60° C (with *B. terrestris* probes) and 55° C (for heterologous probes); filters were washed twice at corresponding temperatures (30 min per wash) in a 2X SSC/0.1 % SDS solution first, and in a 1X SSC/0.1 % SDS solution the second time. Hybridized BAC-filters were imaged in a Storm Scanner (GE Healthcare, Piscataway, NJ, USA) and positive hits were scored with HybSweeper (Lazo et al., 2005).

BAC-DNA preparation and fingerprinting analyses

Positive clones were fingerprinted using techniques established by Chen et al. (2002) and Marra et al. (1997). Briefly, DNA from BAC clones was prepared from 900 µL cultures of Terrific Broth (GIBCO)-Chloramphenicol (12.5 ug/uL) in 96-well format, inoculated with 1.5 µL of BAC freezer stocks. After 18 h, cultures were treated with a modified alkaline lysis method. Samples for fingerprinting were digested with the restriction endonuclease *Hind*III, electrophoresed on 1 % agarose gels for 15 h at 60 V, and stained with Sybr Gold (Invitrogen) for 1 h. Gels were imaged in a Storm Scanner (GE Healthcare, Piscataway, NJ, USA). Fingerprinting data were scored using Image3 software (v.3.10, www.sanger.ac.uk/software/Image/).

To determine average insert size of the library, 192 clones were randomly selected from the library, DNA was digested using the endonuclease *Not*I (NEB, Ipswich, MA, US) and analyzed by PFGE.

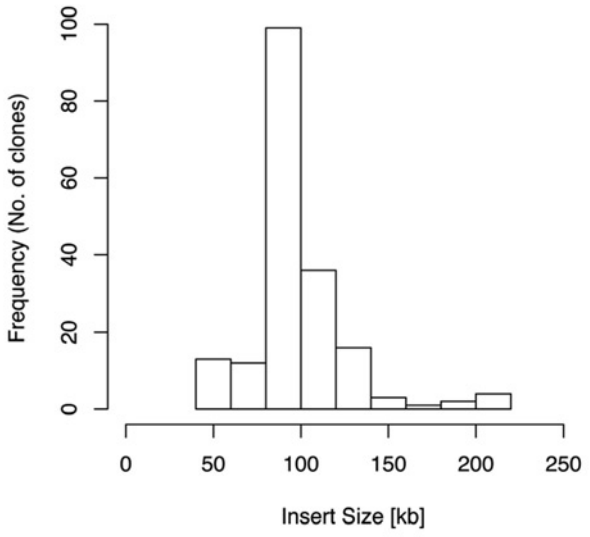

Figure 1. Histogram of insert size distribution of BAC clones ($n = 186$) of the bumblebee BAC-library.

Results

The bumblebee BAC library consists of 36'864 clones. The average insert size (n = 186) was 102.9 ± 28.5 Kb (see Fig. 1) with a range of 40 – 220 Kb. PFGE analysis revealed that the library contains 3.1 % empty clones (6 of 192 clones assayed for insert size). The bumblebee genome has been estimated as being 625 Mb in size (Wilfert et al., 2006). The library thus has an expected coverage of 6x genome equivalents, allowing any one particular *B. terrestris* sequence to be recovered with a probability of 99.7 % from at least one clone.

B. terrestris is a model organism for the evolution of the innate immune system. We therefore screened the BAC library with probes derived from candidate genes involved in antimicrobial defense pathways (NF-κB-like transcription factor *Relish, Defensin*) and in parasite recognition / phagocytosis (*Dscam*). As a positive control, we screened for genes that had previously been used to infer the phylogeny of the genus *Bombus* (Kawakita et al., 2004) (*Long-wave Rhodopsin*, *Arginine Kinase,* and *Elongation Factor 1 alpha*). An average of 27.5 ± 13.3 positive clones for each of six gene probes was retrieved, indicating a high redundancy of the library (Table 1).

The BAC clones identified by hybridization are expected to contain some false positives. To obtain the clones most likely to contain the genes of interest, positive BAC clones were fingerprinted with Hin*d*III and then assembled into contigs using the Fingerprinted Contigs software (FPC (Soderlund et al., 2000)) at high stringency (using a tolerance value of 7 and a minimum cutoff value of 1e-10). This analysis allowed us to anchor genetic markers of the six genes analyzed here onto corresponding physical regions of the genome, represented by BAC clones. Markers generated from *Dscam*, *Defensin*, *Long-Wave Rhodopsin* and *Elongation Factor 1 alpha* were represented in four separate contigs containing 22, 8, 10 and 6 BAC clones respectively; the markers generated for *Arginine kinase* and *Relish* identified 9 and 8 clones respectively, which were represented in a single contig containing 17 BAC clones. These results provide supporting evidence for the quality of the library and constitute a first step in describing the location of these genes on the *B. terrestris* genome.

The FPC analysis allowed us to drastically reduce the number of candidate clones that will become targets for sequencing, i.e. from 52 clones identified in the hybridization screen to 9 clones for *Arginine Kinase.* A common cause of high numbers of false positives in hybridization screens is the use of degenerate primers, which leads to some degree of unspecific binding, to obtain clones containing candidate genes known only in related species. We have used this approach to identify candidate BAC clones for the isolation of immune genes in *B. terrestris* based on sequence information from the honeybee *A. mellifera.* Additionally, several of the genes we screened for – *Arginine Kinase*, *Defensin* and *Elongation Factor 1 alpha* – are known to be double copy genes in *A. mellifera* (Consortium, 2006) and *Nasonia vitripennis* (Genome Assembly 1.0, personal communication; Stephen Richards, Human Genome Sequencing Center, Baylor College of Medicine), this may also inflate the number of positive clones. FPC analysis is a powerful tool to deal with these issues common in genomic studies of non-model organisms: only those clones sharing statistically significant similar band patterns will become candidate clones for further analysis.

Based on the average insert size and the number of BAC clones generated, we have estimated a 6x coverage of the bumblebee genome. The mean number of positive clones after FPC analysis for the genes used in this screen indicates an average 10x coverage. This discrepancy can be explained by over- and underrepresentation of certain regions of the genome. Such biases in genome coverage include the use of a single enzyme, a partial restriction digestion of genomic DNA, and the unavailability of certain regions of genomes such as centromeres, highly repetitive sequences and telomeres, due to their lack of recognition sites for common restriction enzymes (Mahtani and Willard., 1998; Chew et al., 2002; Yuan et al., 2008). With an average 10x coverage of coding sequences, this library is likely to prove a valuable tool in the genomic analysis of *B. terrestris* and related social Hymenoptera.

Discussion

We here describe the construction and characterization of a BAC-library for the bumblebee *Bombus terrestris.* This high-quality library may serve as an important resource for genomic studies of bumblebees, such as gene isolation

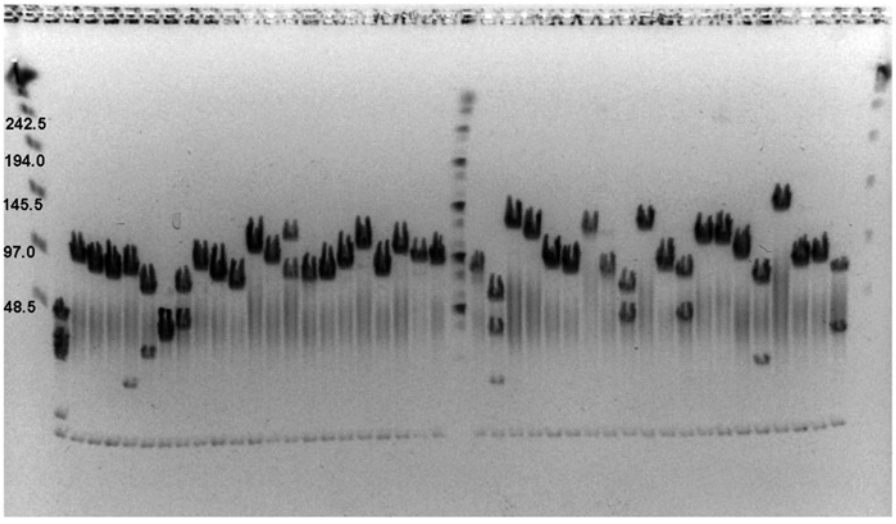

Figure 2. Analysis of BAC clones by PFGE. Randomly picked recombinant BAC clones from our *Bombus terrestris* library were digested with *Not*I to release the cloned genomic insert. Sizes were separated on a 1% agarose CHEF gel (0.5X TBE) and stained with Ethidium bromide. This gel shows the results for 42 BAC clones; the marker loaded in either end-well is Lambda Ladder (NEB). The marker in lane 23 is Midrange II (NEB). Fragment sizes for Lambda Ladders are indicated in Kb on the left.

and genome mapping. We have screened the BAC library with several probes, demonstrating the library's usefulness as a genomic tool for *B. terrestris*. We could retrieve not only clones positive for already described sequences specific to described *B. terrestris* genes (*Long-wave Rhodopsin*, *Arginine Kinase,* and *Elongation Factor 1 alpha* (Kawakita et al., 2004)) but also for probes derived from another species from the genus *Bombus* (*Defensin* from *B. ignitus*) and from the distantly related honeybee *A. mellifera* (*Dscam* (Graveley et al., 2004) and *Relish*).

Comparative research into the social Hymenoptera stands to gain much by combining information from the sequenced genome of the honeybee and genomic information from related species (The Honeybee Genome Consortium, 2006). Understanding the genetics of sex determination in the haplo-diploid Hymenoptera may prove to be a case in point. In the social Hymenoptera, sex is determined by a single complementary locus that triggers male development in hemizygous and homozygous embryos (Cook and Crozier, 1995), while other mechanisms are used in many families of the Hymenoptera (Heimpel and de Boer, 2008). In honeybees, the genetics of complementary sex determination (CSD) was first investigated using linkage mapping (Hunt and Page, 1994). To identify the responsible gene, the identified genetic region was fine-mapped (Hasselmann et al., 2001). With the help of a honeybee BAC library, the *csd* gene was then identified and demonstrated to be functional (Beye et al., 2003). Similarly, the sex determination locus in *B. terrestris* has been genetically mapped to an approximate location (Gadau et al., 2001). Using information from the honeybee and the BAC library we here describe, it will be possible to rapidly investigate the molecular and genetic nature of sex determination in bumblebees. This BAC library thus is not only a valuable tool for investigating the bumblebee genome, but vastly increases the potential for informative comparative studies in the social Hymenoptera.

B. terrestris BAC resources (library and high density filters) may be ordered from the Clemson University Genomics Institute (http://www.genome.clemson.edu/). The use of this BAC library should make reference to this paper. To maximize the information gained from this resource, a data repository for the BAC library is managed by R. Schmid-Hempel, ETH Zürich (rsh@env.ethz.ch).

Acknowledgements

The construction of the BAC library was financed by an ETH Research Grant (nr. TH-19/03–2 to PSH and JG). Authors L. Wilfert and M. Muñoz Torres contributed equally to this work.

References

Baer B. and Schmid-Hempel P. 2001. Unexpected consequences of polyandry for parasitism and fitness in the bumblebee, *Bombus terrestris. Evolution* **55:** 1639–1643

Beye M., Hasselmann M., Fondrk M.K., Page R.E. and Omholt S.W. 2003. The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* **114:** 419–429

Chen M.S., Presting G., Barbazuk W.B., Goicoechea J.L., Blackmon B., Fang F.C., Kim H., Frisch D., Yu Y.S., Sun S.H., Higingbottom S., Phimphilai J., Phimphilai D., Thurmond S., Gaudette B., Li P., Liu J.D., Hatfield J., Main D., Farrar K., Henderson C., Barnett L., Costa R., Williams B., Walser S., Atkins M., Hall C., Budiman M.A., Tomkins J.P., Luo M.Z., Bancroft I., Salse J., Regad F., Mohapatra T., Singh N.K., Tyagi A.K., Soderlund C., Dean R.A. and Wing R.A. 2002. An integrated physical and genetic map of the rice genome. *Plant Cell* **14:** 537–545

Chew J.S., Oliveira C., Wright J.M. and Dobson M.J. 2002. Molecular and cytogenetic analysis of the telomeric (TTAGGG)n repetitive sequences in the Nile tilapia, *Oreochromis niloticus* (Teleostei: Cichlidae). *Chromosoma* **111:** 45–52

Cook J.M. and Crozier R.H. 1995. Sex determination and population biology in the hymenoptera. *Trends Ecol Evol* **10:** 281–286

Gadau J., Gerloff C.U., Kruger N., Chan H., Schmid-Hempel P., Wille A. and Page R.E. 2001. A linkage analysis of sex determination in

Bombus terrestris (L.) (Hymenoptera : Apidae). *Heredity* **87:** 234–242

Goulson D. 2003. *Bumblebees – their Behaviour and Ecology.* Oxford University Press, New York 235 pp

Graveley B.R., Kaur A., Gunning D., Zipursky S.L., Rowen L. and Clemens J.C. 2004. The organization and evolution of the Dipteran and Hymenopteran Down syndrome cell adhesion molecule (Dscam) genes. *RNA-Publ. RNA Soc.* **10:** 1499–1506

Hasselmann M., Fondrk M.K., Page R.E. and Beye M. 2001. Fine scale mapping in the sex locus region of the honey bee (*Apis mellifera*). *Insect Mol. Biol.* **10:** 605–608

Heimpel G.E. and de Boer J.G. 2008. Sex determination in the Hymenoptera. *Annu. Rev. Entomol.* **53:** 209–230

The Honeybee Genome Sequencing Consortium. 2006. Insights into social insects from the genome of the honeybee *Apis mellifera. Nature* **443:** 931–949

Hoffman E.A. and Goodisman M.A.D. 2007. Gene expression and the evolution of phenotypic diversity in social wasps. BMC Biol. 5:23, doi:10.1186/1741-7007-5-23

Hunt G.J. and Page R.E. 1994. Linkage analysis of sex determination in the honey-bee (*Apis mellifera*). *Mol. Gen. Genet.* **244:** 512–518

Kawakita A., Sota T., Ascher J.S., Ito M., Tanaka H. and Kato M. 2003. Evolution and phylogenetic utility of alignment gaps within intron sequences of three nuclear genes in bumble bees (*Bombus*). *Mol. Biol. Evol.* **20:** 87–92

Kawakita A., Sota T., Ito M., Ascher J.S., Tanaka H., Kato M. and Roubik D.W. 2004. Phylogeny, historical biogeography, and character evolution in bumble bees (*Bombus* : Apidae) based on simultaneous analysis of three nuclear gene sequences. *Mol. Phylogenet. Evol.* **31:** 799–804

Lazo G.R., Lui N., Gu Y.Q., Kong X.Y., Coleman-Derr D. and Anderson O.D. 2005. Hybsweeper: a resource for detecting high-density plate gridding coordinates. *Biotechniques* **39:** 320-+

Mahtani M.M. and Willard. H.F. 1998. Physical and genetic mapping of the human X chromosome centromere: repression of recombination. *Genome Res.* **8:** 100–110

Mardulyn P. and Cameron S.A. 1999. The major opsin in bees (Insecta : Hymenoptera): A promising nuclear gene for higher level phylogenetics. *Mol. Phylogenet. Evol.* **12:** 168–176

Marra M.A., Kucaba T.A., Dietrich N.L., Green E.D., Brownstein B., Wilson R.K., McDonald K.M., Hillier L.W., McPherson J.D. and Waterston R.H. 1997. High throughput fingerprint analysis of large-insert clones. *Genome Res.* **7:** 1072–1084

Pereboom J.J.M., Jordan W.C., Sumner S., Hammond R.L. and Bourke A.F.G. 2005. Differential gene expression in queen-worker caste determination in bumble-bees. *Proc. R. Soc. B-Biol. Sci.* **272:** 1145–1152

Peterson D., Tomkins J., Frisch D., Wing R. and Paterson A. 2000. Construction of bacterial artificial chromosome (BAC) libraries: an illustrated guide. *J. Agric. Genomics* **5:** http://wheat.pw.usda.-gov/jag/

Ruiz-Gonzalez M.X. and Brown M.J.F. 2006. Males vs workers: testing the assumptions of the haploid susceptibility hypothesis in bumblebees. *Behav. Ecol. Sociobiol.* **60** 501–509

Sadd B.M. and Schmid-Hempel P. 2006. Insect immunity shows specificity in protection upon secondary pathogen exposure. *Curr. Biol.* **16:** 1206–1210

Sambrook J., Fritsch E.F. and Maniatis T. 1989. *Molecular Cloning.* Cold Spring Harbor Laboratory Press, New York. 1659 pp

Schmid-Hempel P. 2001. On the evolutionary ecology of host-parasite interactions: addressing the question with regard to bumblebees and their parasites. *Naturwissenschaften* **88:** 147–158

Soderlund C., Humphray S., Dunham A. and French L. 2000. Contigs built with fingerprints, markers, and FPCV4.7. *Genome Res.* **10:** 1772–1787

Tomkins J., Luo M., Fang G., Main D., Goicoechea J., Atkins M., Frisch D., Page R.E., Guzman-Novoa E., Yu Y., Hunt G.J. and Wing R. 2002. New genomic resources for the honey bee (*Apis mellifera* L.): development of a deep-coverage BAC library and a preliminary STC database. *Genet. Mol. Res.* **1:** 306–316

Wang J., Jemielity S., Uva P., Wurm Y., Graff J. and Keller L. 2007. An annotated cDNA library and microarray for large-scale gene-expression studies in the ant *Solenopsis invicta. Genome Biol.* **8**

Wilfert L., Gadau J., Baer B. and Schmid-Hempel P. 2007a. Natural variation in the genetic architecture of a host-parasite interaction in the bumblebee *Bombus terrestris. Mol. Ecol.* **16:** 1327–1339

Wilfert L., Gadau J. and Schmid-Hempel P. 2006. A core linkage map of the bumblebee *Bombus terrestris. Genome* **49:** 1215–1226

Wilfert L., Gadau J. and Schmid-Hempel P. 2007b. The genetic architecture of immune defense and reproduction in male *Bombus terrestris* bumblebees. *Evolution* **61:** 804–815

Yuan G., Qi C., Junsong P., Zheng L., Huanle H., Aizhong W., Rentao S. and Run C. 2008. Construction of a BAC library from cucumber (*Cucumis sativus* L.) and identification of linkage group specific clones. *Prog. Nat. Sci.* **18:** 143–147