Rationale to sequence the genome of the red flour beetle, *Tribolium castaneum*

Susan Brown[1], Robin Denell[1], Richard Beeman[2] and Richard Gibbs[3]

[1]Division of Biology, Ackert Hall, Kansas State University, Manhattan, KS
[2] USDA-ARS-GMPRC, Manhattan, KS
[3] HGSC, Baylor College of Medicine, Houston, TX

Overview: In a world of worthy candidates, there are several compelling reasons to sequence the genome of the red flour beetle *Tribolium castaneum.* First and foremost, Tribolium is one of the most sophisticated genetic model organisms among all higher eukaryotes. Among arthropods, only Drosophila offers greater power and flexibility of genetic manipulation. Second, the Tribolium genome sequence will provide an informative link when direct comparisons between human and fruit fly sequences are unproductive. Third, as a member of the most primitive order of holometabolous insects, the Coleoptera, it is in a key phylogenetic position to inform us about the genetic innovations that accompanied the evolution of higher forms with more complex development. Fourth, Coleoptera is the largest and most species diverse of all eukaryotic orders and Tribolium offers the only genetic model for this profusion of medically and economically important species. Analysis of the Tribolium genome will facilitate the discovery of new pharmaceuticals and antibiotics. In addition, it will lead to a better understanding of resistance mechanisms and improved specificity and efficacy of insecticides and biological agents to control crop pests and disease vectors. Finally, many genetic and genomic tools are already available for Tribolium. Both forward and reverse genetic approaches are available to facilitate functional genetic analysis. A molecular recombination map and a physical map based on a BAC library of ~20 x genome coverage are now being completed, and will furnish the necessary foundation upon which to assemble the genomic sequence, which at ~200 Mbs is only slightly larger than that of Drosophila.

A. Specific biological rationales for the utility of new sequence data.

Improving human health, informing human biology and providing surrogate experimental systems. Bioactive agents with potential impact on human health and biology have long been sought in the plant and fungal biodiversity present in the tropical rain forests. Beetles, with their unparalleled species and habitat diversity, are a major untapped source of antibiotics and biopharmaceuticals. In Tribolium, *p*-benzoquinones, aliphatic hydrocarbons and other potential repellents, irritants, toxicants and antifungal or antibacterial components are produced in large “stink” glands (Blum, 1981). Naturally-occurring polymorphism in antibiotic potency has been demonstrated in at least two Tribolium species (Prendeville and Stevens, 2002), and several mutations have been identified that affect the biosynthesis and secretion of stink gland components. *Tribolium castaneum* was the first animal species ever reported to produce prostaglandin synthetase inhibitors (Howard et al., 1986). Given the importance of anti-inflammatory drugs acting through the prostaglandin pathway, the detection of this class of inhibitory compounds in Tribolium demonstrates the potential for drug discovery via mining the Tribolium genome.

Tribolium species host a large variety of protozoan and bacterial parasites and/or symbionts and thus offer a model system in which to study host/pathogen interactions. Like many arthropods, certain Tribolium species are infected with Wolbachia, a rickettsia-like organism. Strains of some Tribolium species (e.g. *T. confusum*) harbor Wolbachia as obligate intracellular symbionts or parasites, and show classic incompatibility syndrome when mated with Wolbachia-deficient strains (Wade and Stevens, 1985). Other species, such as *T. madens*, are devoid of Wolbachia, and develop lethal infections when artificially inoculated (Fialho and Stevens, 2000). Still others, including *T. castaneum,* appear to be refractory to Wolbachia, in that they are immune to infection and are host-incompetent. Utilization of Wolbachia infection is being examined with respect to insect vectors of human disease, such as mosquitoes, either to impact the population density directly or to drive favorable traits through the population (Dobson, 2003; Rascon et al., 2003) The same approaches may be efficacious with respect to insect agricultural pests. A sequenced Tribolium genome has considerable potential for stimulating the study of Wolbachia-host interactions to reveal the underlying genetic basis for this wide range of responses to Wolbachia challenge.

The Center for Disease Control lists Tribolium as an intermediate host of the human tapeworm parasite *Hymenolepis diminuta* (http://www.dpd.cdc.gov/dpdx/HTML/ImageLibrary/Hymenolepiasis_il.asp?body=G-L/Hymenolepiasis/body_Hymenolepiasis_il8.htm). Differences in susceptibility to tapeworm infection have been observed between different strains of *Tribolium castaneum* (Yan and Norman, 1995). A recent study identified five quantitative trait loci (QTL) which significantly affect beetle susceptibility (Zhong et al., 2003). Availability of the complete genome sequence of *T. castaneum* would greatly facilitate the identification of the genes important for parasite transmission and enhance our understanding of the genetic basis of complex traits such as resistance to parasitism.

Providing a better connection between the sequences of non-human organisms and the human sequence and informing the human sequence. Drosophila homologs of many human genes have been identified. However, there is mounting evidence that gene sequences in Drosophila are diverging rapidly and it is not likely that all homologs will be identified by direct sequence comparisons. For example, analysis of the recently published Anopheles genomic sequence indicates that mosquito and fly genes are markedly less similar than expected for two taxa separated for about 200-250 million years (De Gregorio and Lemaitre, 2002). Furthermore, when Tribolium, Drosophila and human sequences are compared, the Tribolium sequences are often more similar to their human counterparts than are their Drosophila homologs (D.Tautz, personal communication; D. Beeman unpublished). In fact, some Drosophila gene sequences have diverged to the point that BLAST analysis of certain honey bees ESTs found better matches among chordates than in Drosophila (Whitfield et al., 2002). For most of these, Drosophila homologs could be identified, but in some cases the Apis coding sequences have no match in the Drosophila genome, indicating these genes were lost in the Drosophila lineage. These results indicate that data from other insect orders will be required to link human genes to their Drosophila homologs. In fact, the relationship between Drosophila *zen* and human HOX3 genes was elucidated by comparisons including data from Schistocerca and Tribolium (Falciani et al., 1996). More recently, in a chromosomal walk to positionally clone several Tribolium genes, we identified the insect ortholog of a human muscular dystrophy gene, which was not identified by direct comparison with Drosophila sequence (Beeman, unpublished). Thus, analysis of the Tribolium genome will provide informative comparisons for the identification of insect homologs of human genes. Moreover, Tribolium is the most

efficient system in which to perform functional analysis of those genes lost in the Drosophila lineage but conserved in other insects.

Facilitating the ability to do experiments, e.g "direct" genetics or positional cloning, in additional organisms. Genetic screens in Tribolium have produced a wealth of morphological, physiological and developmental mutants. Recently, genome wide screens for embryonic lethality have produced several putative segmentation and homeotic mutants (Maderspacher et al., 1998; Sulston and Anderson, 1996). Currently, there are over 300 mutant strains in the Manhattan, Kansas stock collection, which are easily maintained at room temperature by subculturing every 3-4 months. The embryonic lethal mutations are maintained as heterozygous male lines (Berghammer et al., 1999a). The first maternal effect selfish (*Medea*) genes were discovered in Tribolium (Beeman et al., 1992a), and similar mechanisms of maternal selection of selfish genes have subsequently been discovered only in mammals (Hurst, 1993; Peters and Barker, 1993). Attempts to positionally clone Medea and other interesting mutations are currently underway and would greatly benefit from a completely sequenced genome.

Tribolium castaneum is a sophisticated genetic model species representing the largest and most diverse of all eukaryotic orders, the Coleoptera. This order includes many beneficial and deleterious species, the latter associated with billions of dollars of agricultural losses annually. The sequence of the Tribolium genome will provide a basis for identifying and interpreting the organization and function of genes and gene families in other beetles. Further, components of the biosynthetic pathways identified in Tribolium can be functionally tested. The value of using the *C. elegans* genome sequence to study parasitic nematodes is a direct parallel (Bird et al., 1999).

Insects produce and respond to a plethora of secondary metabolites to identify and locate mates or hosts, as well as to deter enemies and defend against pathogens. These compounds have long been under intense scrutiny for possible application to the biocontrol of insect pests (Campbell et al., 2002). We are just beginning to identify the enzymatic pathways by which these compounds are synthesized, and to understand how they function. Tribolium species produce and emit a rich array of volatiles, including a species-nonspecific aggregation pheromone 4,8-dimethyldecanal, as well as other semiochemicals whose functions are still uncertain (Arnaud et al., 2002). A fully sequenced beetle genome would provide an important resource for chemical ecology studies in Tribolium and a wide variety of other beetles.

As a major, global pest of stored grain and cereal products, peanuts, and many other dried and stored commodities for human consumption, Tribolium has a long history of exposure to pesticides. It has proven to be readily adaptable to all classes of insecticides and fumigants, having developed resistance via oxidative and hydrolytic metabolism, target insensitivity and other mechanisms (Andreev et al., 1994; Beeman and Nanis, 1986; Beeman and Stuart, 1990). These features, in combination with excellent genetic tractability, recommend Tribolium as an ideal subject for identification of new pesticide targets through knowledge of resistance mechanisms (Andreev et al., 1994; Beeman et al., 1992b)

Expanding our understanding of evolutionary processes. Among the insects, Tribolium is in a key phylogenetic position to inform us about the genetic innovations that accompanied the evolution of higher insects. Among the winged insects, relatively more primitive orders are placed in the hemimetabola. These insects do not develop morphologically distinct larval and adult forms, but rather undergo a series of molts in which earlier stages (nymphs) resemble miniature adults. In contrast, higher insects that

undergo complete metamorphosis comprise the holometabola. Larval forms are worm-like, then transform into pupae within which develop adult insects of much different appearance.

Coleoptera occupies a basal position among the holometabola. In comparison, Diptera are one of the most advanced orders. Beetles and flies diverged close to 300 million years ago (Kristensen, 1999). Other insect species whose genomes are being sequenced (honey bee = Hymenoptera), or are likely to be (butterflies and moths = Lepidoptera), diverged more recently. In future comparison to hemimetabolan genomes, the basal position of Tribolium offers considerable potential for illuminating the origins of complete metamorphosis. It also offers an important phylogenetic position for comparisons with the other holometabolous insect orders. Each of these shows a number of specialized morphological characters. Among beetles, for example, the ancestral two pairs of wings have been modified such that the first pair forms wing covers, or elytra. In contrast, in flies the second pair of wings has been replaced by small appendages, halteres, important for balancing during flight. Drosophila is even more highly specialized in several aspects of embryogenesis and larval morphology. Embryonic segmentation occurs simultaneously along the anterior-posterior axis (long-germ development), whereas more primitive insects (and vertebrates) show sequential segmentation from anterior to posterior. Insect larvae typically have heads derived from several segments as well as gnathocephalic segments with appendages that function in feeding. During Drosophila embryogenesis, all of these segments move through the presumptive anterior opening of the digestive tract to occupy internal positions, where they elaborate evolutionarily novel structures. The resulting larva is essentially headless, and also lacks the thoracic limbs characteristic of almost all insects. In contrast, related developmental events in Tribolium are much more generalized. Segmentation (short-germ) is progressive. Larvae have typical head and gnathocephalic segments, and bear thoracic limbs. Since head segments remain in a linear order, studies in Tribolium would contribute to a comprehensive understanding of head and ultimately brain development. Tribolium head development should be very similar to that in hemimetabolous insects including grasshoppers where brain development has been very carefully described (Watson and Schurmann, 2002), but functional approaches are limited compared to Tribolium.

The Lepidoptera and Hymenoptera also show some conserved and many specialized developmental and morphological features. It is an exciting prospect to use a comparative approach at the genomic level to better understand the evolutionary changes that brought about morphological specializations. However, the most informative approaches to understanding the evolution of unique morphological features will utilize genetic as well as genomic approaches in comparisons of insects with relatively primitive versus specialized development. No Lepidopteran or Hymenopteran presently offers the powerful genetic approaches established for Drosophila and Tribolium. In the next section, we document the considerable progress already made in understanding the differences in segmentation and embryonic patterning in these two insects. The availability of genomic approaches in Tribolium will make such comparative studies all the more informative, and will provide an outstanding general model for understanding the evolution of morphological specializations that will be applicable to phylogenetically diverse animals.

Expanding our understanding of basic biological processes relevant to human health, including developmental biology and neurobiology. Advances in cellular,

developmental and neurobiology are predominantly gained by studies of model organisms amenable to a variety of genetic and biochemical approaches. Many of the intricacies of embryonic patterning have been elucidated by genetic and molecular studies in Drosophila. However, as described above many aspects of Drosophila development are highly derived, and evidence suggests that the underlying genetic regulation is equally derived. Comparative studies in several insects including Tribolium (some of which are briefly summarized below to highlight contributions made by Tribolium studies) have provided insight into many developmental processes.

The establishment of anterior-posterior and dorsal-ventral coordinates in early embryos is one area in which Drosophila has been shown to display highly specialized regulatory mechanisms. Rapid (24 hr) embryonic development in Drosophila is facilitated by the maternal contribution of factors that define the egg coordinates. However, there is mounting evidence that the gene encoding the anterior morphogen Bicoid evolved within the Dipteran lineage by gene duplication and divergence (Stauber et al., 2002; Brown et al., 2001). In fact, recent analysis indicates that *hunchback* and *orthodenticle* pattern the anterior region of the Tribolium embryo and may be part of an ancestral patterning system (Schroder, 2003). The dorsal/ventral axis in Tribolium appears to be patterned by gradients of Toll-receptor and Dorsal protein centered on the ventral midline, (Chen et al., 2000; Maxton-Kuchenmeister et al., 1999) which regulate *twist* (Sommer and Tautz, 1994) and *decapentaelegic (dpp)* (Sanchez-Salazar et al., 1996) expression as in Drosophila. However, these gradients are formed by zygotic regulation of transcription in Tribolium, implying that the local activation of maternally supplied Toll-receptor and ventrally limited nuclear localization of Dorsal are derived features of D/V patterning in Drosophila. Expression of *caudal*, *tailless* and *forkhead* homologs in Tribolium suggest that posterior, but not anterior, aspects of the terminal patterning system are conserved (Schroder et al., 2000; Schulz et al., 1998).

Comparative analysis of segmentation in Tribolium has shown that gap and pair-rule genes are expressed in conserved patterns (Brown and Denell, 1996; Sommer and Tautz, 1993; Wolff et al., 1995) but forward and reverse genetic approaches suggest that conserved expression patterns do not always indicate conserved function. For example, a deletion that removes the pair-rule gene *fushi tarazu* does not result in a pair-rule phenotype in Tribolium (Brown et al., 1994a). These results emphasize the importance of comparative studies using genetically tractable insects such as Tribolium. On the other hand, it is likely that segment boundaries are formed by an ancient mechanism as evidenced by the highly conserved expression patterns of the segment polarity genes *wingless* and *engrailed* (Nagy and Carroll, 1994; Brown et al., 1994b) as well as functional studies (Oppenheimer et al., 1999). Genetic analysis of homeotic mutants in Tribolium provided the first evidence for a contiguous cluster of Hox genes (Beeman, 1987), as later observed in vertebrates. Genetic and molecular analysis of Tribolium homeotic genes and mutants continues to correlate changes in regulatory gene interactions and morphological evolution, and show that some Hox gene functions in Drosophila are derived (Brown et al., 2000; DeCamillis et al., 2001; Lewis et al., 2000; Brown et al., 2002b).

Formation of limb fields requires expression of *dpp* in Drosophila, but not in Tribolium (Jockusch et al., 2000). Although several genes required for imaginal leg development in Drosophila are also expressed in developing embryonic legs in Tribolium, their relative expression patterns indicate that there are significant differences in the genetic regulation of leg development (Beermann et al., 2001; Prpic et al., 2001).

The expression patterns observed in Tribolium are similar to other insects while those in Drosophila are specialized, perhaps reflecting development from imaginal discs.

The organization of the central nervous system is highly conserved in arthropods. In Drosophila, the segmentation and homeotic genes are expressed in the developing CNS and are required for proper neuronal differentiation. In Tribolium, segmentation and homeotic genes are also expressed in the developing CNS in segmentally reiterated patterns that closely resemble those of their Drosophila counterparts. Recently, sequence analysis of BAC clones containing proneural genes in Tribolium has provided insight into the evolution of the achaete/scute complex (Wheeler et al., 2003). Comparison of the function of the single Tribolium *achaete/scute* homolog with that of the Drosophila proneural *ac/sc* genes suggests that the Drosophila *ac/sc* genes acquired new developmental roles in specifying the fate of neural precursors while maintaining an ancestral function in their formation. s*cute* also acquired a role in somatic sex determination in the Drosophila lineage (Torres and Sánchez, 1989); (Skaer et al., 2002). Further analysis indicates that *ventral neurons defective (vnd)*, *intermediate neurons defective (ind)* and *muscle segment homoebox (msh)* are expressed in Tribolium in patterns largely similar to those of their Drosophila homologs. However, gaps in the expression of *vnd* indicate that some neurons in the Tribolium CNS must be patterned by different genes (Wheeler et al., 2003). Finally, differences in the genetic regulation of compound eye development are suggested by the expression pattern of *dpp* in Tribolium (Friedrich and Benzer, 2000).

These studies indicate that the Drosophila paradigms, while conserved in many aspects, are not sufficient to explain the developmental strategies (e.g., long vs. short germ, imaginal vs. embryonic limbs) used by most other insects. This comparative approach, though informative, has been limited to the analysis of candidate genes whose functions were first identified in Drosophila. It is important to move beyond the candidate gene approach in understanding the observed differences, which are likely to affect intracellular processes and intercellular interactions such as cell signaling *that will be relevant to human biology.* Studies in Tribolium are likely to be most informative since Tribolium offers the possibility of genetic manipulation in addition to its facility for developmental and molecular studies. One only needs to examine the explosion of information about the genetic regulation of development resulting from the sequenced genomes of Drosophila and other model organisms to appreciate the impact of similar information on Tribolium research.

B. Strategic issues in acquiring new sequence data.

The demand for the new sequence data. The largest accumulation of laboratories working extensively on the developmental genetics of Tribolium are in Kansas (three labs) and Germany (four labs). Under the auspices of the HFSP, members of these research groups have met on a yearly basis since 1999. At the last meeting, held at the University of Tübingen in Sept 2002, we discussed the possibility of a Tribolium genome project. It was the consensus of the meeting that the sequence of the Tribolium genome would benefit not only the research projects of those present, but a wide range of projects involving comparative approaches at the molecular or developmental level. In addition, several PIs working on Tribolium attended the USDA sponsored Comparative Insect Genomics Workshop (Washington DC, Oct 2001) where it was the consensus that Tribolium was the obvious candidate Coleopteran for genome sequencing. Several

recent reviewers have also concluded that Tribolium is the obvious next candidate for a whole genome sequencing project (Evans and Gundersen-Rindal, 2003; Kaufman et al., 2002). A perusal of PubMed shows that in the period 2001-2002, thirty-eight independent laboratories published papers on Tribolium while a search of the SCI database reveals over 500 publications referring to *Tribolium castaneum*. To date, twenty-four laboratories have submitted Tribolium sequences to GenBank. Neither approach fully documents the many cases where investigators have cloned Tribolium genes for the purpose of molecular comparisons. As the sequences of additional genomes are completed, it becomes obvious that each has a unique capacity to inform the human genome, and that complete understanding of the human genome will require comparative analysis of several representative genomes. Thus the size of the research community that will benefit from the complete Tribolium genome sequence far exceeds that of individuals working primarily on Tribolium. Needless to say, the complete genome sequence will greatly enhance the ability to identify and study genes of interest. Several individuals in the Drosophila research community have extended their work to include the examination of Tribolium orthologs and their impact on fly developmental pathways (Lohr et al., 2001; Fujioka et al., 2002; Guichet et al., 1997). We routinely process requests for Tribolium genomic DNA, cDNA and genomic libraries and Tribolium mutants. The enthusiasm of the research community is documented in the letters of support submitted separately.

The suitability of the organism for experimentation. *Tribolium castaneum* is a holometabolous insect that has been raised for more than four decades in the laboratory, and subsists on wheat flour supplemented with 5% yeast. It is easily manipulated and tolerates crowding and inbreeding. The generation time is flexible (3-6 weeks depending on the rearing temperature), and adults experience long reproductive lives. Tribolium eggs are approximately twice the size of Drosophila eggs, and most protocols developed for experimental manipulation in Drosophila, including *in situ* hybridization and immunohistochemistry, work well in Tribolium. Tribolium lacks polytene chromosomes, but mitotic and meiotic spreads are easily obtained (Stuart and Mocelin, 1995). Nine autosomes and X/Y sex chromosomes comprise the chromosomal complement in *Tribolium castaneum,* and recombination occurs in both sexes.

Sophisticated genetic manipulations are possible in Tribolium. For example, gain of function mutations have been reverted by mutagenesis to reveal null phenotypes (Stuart et al., 1993; Shippy et al., 2000; Brown et al., 2000). Mutant alleles of homeotic genes not previously identified by mutation were identified in screens designed to saturate the region of the HOMC uncovered by deficiencies (Brown et al., 2000). For several regions, chromosomal rearrangements have been induced (Beeman et al., 1986) to facilitate stock maintenance.

There are a plethora of additional genomic resources and technologies (e.g. gene transfer, ability to go from molecule to mutation) available in Tribolium that will allow the new sequence information to be effectively used. Tools for reverse genetic analysis in Tribolium include RNA interference and transformation. These methods, which rely on known gene sequences, will facilitate the effective use of new genomic sequence information. Maternal and/or zygotic mRNAs can be depleted in Tribolium by injecting dsRNA into the abdomen of female pupa (Bucher et al., 2002) or freshly laid eggs (0-2 hours) (Brown et al., 1999). Thus Tribolium is highly suited to high through-put genome-wide RNAi screens, as in *C. elegans* (Kamath et al., 2003; Maeda et al., 2001). Such screens would identify embryonic and maternal genes, as well as genes

with functions relevant to basic cell biology that produce phenotypes affecting oogenesis.

Using the piggyBac vector and eye-specific transformation markers, it is now possible to introduce DNA into the Tribolium genome (Berghammer et al., 1999b; Lorenzen et al., 2002). Several researchers are currently designing vector constructs to induce tissue and stage specific expression of introduced genes and dsRNA constructs. In addition, several labs are preparing to perform genome wide screens using transposon-mediated mutagenesis. Having the whole genome sequence available will greatly facilitate the identification and analysis of transposon insertion sites, especially in cases where the affected gene is not in the immediate vicinity of the insertion.

The rationale for the complete sequence of the organism. Genes that regulate critical events in development are often expressed at very low levels and their representation in EST projects is likely to be rare. Having the entire genome sequence to analyze should facilitate the identification of homologous genes expected to be expressed at low levels or in specific tissues. Furthermore, to truly exploit Tribolium sequence as a link between Drosophila and vertebrate model organisms will require comparisons of the entire genome sequence.

Insects display a wide range of developmental, morphological and physiological adaptations to their environment. These are in part due to differences in the regulation of conserved genes and pathways. Full appreciation of these differences requires analysis of the noncoding regions of the genome where the regulatory elements reside. In addition to sequence comparisons, it will be possible to perform functional analysis of noncoding regions by introducing into the Tribolium genome, reporter constructs driven by putative regulatory sequences. However, it is often difficult to identify regulatory elements in noncoding DNA by sequence comparisons between distantly related insects. Indeed, even the Drosophila research community eagerly awaits the sequence of a second Drosophilid for comparative analysis of noncoding regions and the identification of conserved regulatory regions. Similarly, we have identified a member of a different Tribolium species group that should provide informative comparisons of noncoding regions. *Tribolium* is a cosmopolitan genus; species are found infesting stored grain throughout the world. The castaneum species group originated in Indonesia and Australia, while the confusum species group originated in Africa (Hinton, 1948). The canonical species of each group, *Tribolium castaneum* and *Tribolium confusum*, are difficult to distinguish, but cannot interbreed. These two species have enjoyed a long history of use in population biology studies. Preliminary analysis of 16s rRNA genes and regions upstream of the pair-rule gene *even-skipped* indicate that these species are sufficiently distant to provide informative comparisons of noncoding regions (Brown, unpublished; Tautz, personal communication). We have constructed a lambda genomic library of *T. confusum* from which to isolate regions of interest for further study.

The cost of sequencing the genome and the state of readiness of the organism's DNA for sequencing. The size of the Tribolium genome has been estimated by hybridization kinetics (Cot analysis) (Brown et al., 1990) and microdensitometric quantitation of Feulgen-stained spermatids (Alvarez-Fuster et al., 1991). The results of both methods indicate that the haploid complement in Tribolium is approximately 2pg or 200Mb. Unique sequences comprise more than 60% of the genome and repetitive DNA displays a long period interspersion pattern. Several satellite DNA sequences that are conserved between Tribolium species have been identified (Juan et al., 1993), and are clustered in putative centromeric regions (Plohl et al., 1993). The WGS sequencing strategy used to sequence the Drosophilid genomes, which is in place at the Human

Genome Sequencing Center at Baylor College of Medicine, will be used to sequence the Tribolium genome. 7-8 x coverage of the Tribolium genome will be the most informative and should cost between 3-6 million dollars US. A strain of *Tribolium castaneum* that had been inbred for 20 generations (Scott Thomson, University of Wisconsin, Parkside) has been used as a source of DNA for construction of BAC libraries, as well as molecular and physical maps. This strain is currently maintained in several laboratories in the US, and will be used as the source of DNA for sequencing. Southern analysis of this strain using a midrepetitive sequence as probe (Stuart et al., 1996), revealed virtually identical patterns in the 10 individuals sampled. Use of this inbred strain will greatly alleviate problems such as contig stretching, which occurs when attempting to align independent sequencing reads from heterozygous loci and which has made assembly of the *Aedes aegypti* genome somewhat problematic (D. Severson, personal communication).

Other sources of funding for this project. Several resources that will facilitate the assembly and annotation of the Tribolium genome have been funded through a variety of sources. Two BAC libraries were constructed by Exelixis Pharaceuticals Inc., San Francisco, from DNA isolated from the inbred strain of *Tribolium castaneum* described above. These libraries are based on *Eco*R I partial digests of Tribolium genomic DNA cloned into the pBACE3.6 vector. The original library is arrayed on 24 384-well plates and contains inserts whose sizes average 125 Kb. A second *Eco*R I library with inserts averaging 150 Kb was constructed in the same vector and arrayed on 48 plates. At slightly over 27,000 clones, these libraries represent more than 20 x coverage of the Tribolium genome. As part of the NSF funded BAC library initiative, these libraries are being archived and distributed by Dr. Jeff Tomkins at Clemson University Genomics Institute (CUGI), SC.

In addition, Exelixis has agreed to provide a large resource of EST sequence data to facilitate annotating the Tribolium genome sequence (see letter of support). Exelixis has sequenced more than 8,800 ESTs from adult tissue and assembled them into more than 4,600 contigs. With funds from the Human Frontiers Science Program, more than 2,000 additional ESTs from embryonic tissue have been sequenced (D. Tautz, personal communication). These have been assembled into a minimum of 586 non-redundant clones. Using automated *in situ* hybridization, several genes expressed in interesting and informative patterns during embryonic development have been identified, and are being functionally analyzed using RNAi. This high-throughput methodology will be available at the University of Cologne, Germany for similar analysis of potentially interesting ORFs identified in the Tribolium genome (D. Tautz, personal communication). In addition, cDNAs from developmental studies by individual researchers are also available for annotation purposes. It is likely that a significant amount of annotation will be facilitated by comparison with other insect genomes.

The linkage groups of the original recombination map were identified by morphological mutations, while physiological and biochemical markers have since been added. To facilitate genetic analysis and maintenance of mutant stocks several balancers, duplications and deficiencies have been created (Beeman et al., 1986; Beeman et al., 1996). More than 40% of the genome is represented by linkage groups 2 and 3. More than 30 % of the genome is balanced by crossover suppressors, and the extent of suppression has been determined and placed on the genetic map. Recently, recombination maps based on molecular markers have been constructed. A map based on 133 RFLP markers predicts a genome of 570cM (Beeman and Brown, 1999). With funding from HFSP, a higher resolution molecular map of sequence-tagged sites is

being completed. This map contains over 400 markers from BAC end sequences, ESTs and developmentally important genes identified by individual researchers. With funds from NIH, NSF, USDA, Kansas State University, the Max Planck Institute for Developmental Biology and HFSP we are completing a physical map based on the BAC libraries constructed by Exelixis. This map is based on *Hin*D III digests of more than 27,000 BAC clones. The molecular and physical maps are integrated to some extent by the BAC end sequences already included in the molecular map. This integration will improve as more BAC clones are sequenced and mapped. The ends of these BAC clones are being sequenced by D. Tautz's group, (equivilant to ~10 % of the genome) and mapped by R. Beeman and S. Brown. All together, these resources provide a high resolution scaffold on which to assemble the genome sequence. Several BACs have been sequenced by shotgun methods (Brown et al., 2002a; Wheeler et al., 2003), and were assembled with no difficulty. In addition, the sequence of the *Tribolium castaneum* mitochondrial genome has recently been published (Friedrich and Muqim, 2003). Clearly, the rationale to sequence the Tribolium genome is strongly supported by the plethora of genomic resources already available to facilitate the assembly and analysis of the Tribolium genome sequence.

Bioinformatics. The Tribolium genetic database is maintained by R. Beeman (http://bru.usgmrl.ksu.edu/sci/beeman). With support from an NIH IDEA grant (KBRIN) we are constructing a database at KSU to handle the molecular map, the physical map and the EST sequence data from Exelixis and D. Tautz. Initial assembly of the Tribolium genome sequence will be performed at the sequencing center, but it is our goal to make the sequence available to the greater research community as quickly as possible. Toward this end we will use existing software and database management tools, such as the Generic Genome Browser (Stein et al., 2002) as much as possible to construct a database that will required minimal curation. A combination of EST and cDNA mapping, automated gene prediction software and BLAST comparisons will be used to annotate the Tribolium genome. We will work with other database managers to design a database that can be queried in a variety of ways to provide maximal access. Following the lead of the honey bee genome consortium (Gene Robinson, personal communication), we will collaborate with others to design comprehensive arthropod databases for comparative analysis.

Reference List

Alvarez-Fuster,A., Juan,C., and Petitpierre,E. (1991). Genome size in *Tribolium* flour-beetle: inter- and intraspecific variation. Genetic. Res. Camb. *58*, 1-5.

Andreev,D., Rocheleau,T., Phillips,T.W., Beeman,R.W., and ffrench-Constant,R.H. (1994). A PCR diagnostic for cyclodiene insecticide resistance in the red flour beetle, *Tribollium castaneum.* Pesticide Science *41*, 345-349.

Arnaud,L., Lognay,G., Verscheure,M., Leenaers,L., Gaspar,C., and Haubruge,E. (2002). Is dimethyldecanal a common aggregation pheromone of Tribolium flour beetles? J. Chem. Ecol. *28*, 523-532.

Beeman,R.W. (1987). A homoeotic gene cluster in the red flour beetle. Nature *327*, 247-249.

Beeman,R.W. and Brown,S.J. (1999). RAPD-based genetic linkage maps of *Tribolium castaneum*. Genetics *153*, 333-338.

Beeman,R.W., Friesen,K.S., and Denell,R.E. (1992a). Maternal-effect, selfish genes in flour beetles. Science *256*, 89-92.

Beeman,R.W., Johnson,T.R., and Nanis,S.M. (1986). Chromosome rearrangements in *Tribolium castaneum*. J. Hered. *77*, 451-456.

Beeman,R.W. and Nanis,S.M. (1986). Malathion resistance alleles and their fitness in the red flour beetle (Coleoptera: Tenebrionidae). J. Econ. Entomol. *79*, 580-587.

Beeman,R.W. and Stuart,J.J. (1990). A gene for lindene + cyclodiene resistance in the red flour beetle (Coleoptera: Tenebrionidae). J. Econ. Entomol. *83*, 1745-1751.

Beeman,R.W., Stuart,J.J., Denell,R.E., McGaughey,W.H., and Dover,B.A. (1992b). Tribolium as a model insect for study of resistance mechanisms. (Washington, D.C.: ACS), pp. 202-208.

Beeman,R.W., Stuart,J.J., Haas,M.S., and Friesen,K.S. (1996). Chromosome extraction and revision of linkage group 2 in *Tribolium castaneum*. J. Hered. *87*, 224-232.

Beermann,A., Jay,D.G., Beeman,R.W., Hulskamp,M., Tautz,D., and Jurgens,G. (2001). The Short antennae gene of Tribolium is required for limb development and encodes the orthologue of the Drosophila Distal-less protein. Development *128*, 287-297.

Berghammer,A., Bucher,G., Maderspacher,F., and Klingler,M. (1999a). A system to efficiently maintain embryonic lethal mutations in the flour beetle *Tribolium castaneum*. Dev. Genes Evol. *209*, 382-389.

Berghammer,A.J., Klingler,M., and Wimmer,E.A. (1999b). A universal marker for transgenic insects. Nature *402*, 370-371.

Bird,D.M., Opperman,C.H., Jones,S.J., and Baillie,D.L. (1999). THE CAENORHABDITIS ELEGANS GENOME: A Guide in The Post Genomics Age. Annu. Rev. Phytopathol. *37*, 247-265.

Blum,M.S. (1981). Chemical Defenses of Arthropods. (New York: Academic Press).

Brown,S., Fellers,J., Shippy,T., Denell,R., Stauber,M., and Schmidt-Ott,U. (2001). A strategy for mapping *bicoid* on the phylogenetic tree. Curr. Biol. *11*, R43-R44.

Brown,S.J., DeCamillis,M., Gonzalez-Charneco,K., Denell,M., Beeman,R., Nie,W., and Denell,R. (2000). Implications of the *Tribolium Deformed* mutant phenotype for the evolution of Hox gene function. Proc. Natl. Acad. Sci. USA *97*, 4510-4514.

Brown,S.J. and Denell,R.E. (1996). Segmentation and dorsal-ventral patterning in Tribolium. Sem. Cell Dev. Biol. *7*, 553-560.

Brown,S.J., Fellers,J.P., Shippy,T.D., Richardson,E.A., Maxwell,M., Stuart,J.J., and Denell,R.E. (2002a). Sequence of the *Tribolium castaneum* homeotic complex: the region corresponding to the *Drosophila melanogaster* antennapedia complex. Genetics *160*, 1067-1074.

Brown,S.J., Henry,J.K., Black,W.C., and Denell,R.E. (1990). Molecular genetic manipulation of the red flour beetle: Genome organization and cloning of a ribosomal protein gene. Insect Biochem. *20*, 185-193.

Brown,S.J., Hilgenfeld,R.B., and Denell,R.E. (1994a). The beetle *Tribolium castaneum* has a *fushi tarazu* homolog expressed in stripes during segmentation. Proc. Natl. Acad. Sci. USA *91*, 12922-12926.

Brown,S.J., Mahaffey,J.P., Lorenzen,M.L., Denell,R.E., and Mahaffey,J.W. (1999). Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. Evol. & Dev. *1*, 11-15.

Brown,S.J., Patel,N.H., and Denell,R.E. (1994b). Embryonic expression of the single *Tribolium engrailed* homolog. Dev. Genet. *15*, 7-18.

Brown,S.J., Shippy,T.D., Beeman,R.W., and Denell,R.E. (2002b). Tribolium Hox genes repress antennal development in the gnathos and trunk. Mol. Phylogenet. Evol. *24*, 384-387.

Bucher,G., Scholten,J., and Klingler,M. (2002). Parental RNAi in Tribolium (Coleoptera). Curr. Biol. *12*, R85-R86.

Campbell,J.F., Mullen,M.A., and Dowdy,A.K. (2002). Monitoring stored-product pests in food processing plants with pheromone trapping, contour mapping, and mark-recapture. J. Econ. Entomol. *95*, 1089-1101.

Chen,G., Handel,K., and Roth,S. (2000). The maternal NF-(&kgr;)B/Dorsal gradient of *Tribolium castaneum*: dynamics of early dorsoventral patterning in a short-germ beetle. Development *127*, 5145-5156.

De Gregorio,E. and Lemaitre,B. (2002). The mosquito genome: the post-genomic era opens. Nature *419*, 496-497.

DeCamillis,M.A., Lewis,D.L., Brown,S.J., Beeman,R.W., and Denell,R.E. (2001). Interactions of the Tribolium *Sex combs reduced* and *proboscipedia* orthologs in embryonic labial development. Genetics *159*, 1643-1648.

Dobson,S.L. (2003). Reversing Wolbachia-based population replacement. Trends Parasitol. *19*, 128-133.

Evans,J.D. and Gundersen-Rindal,D. (2003). Beenomes to Bombyx: future directions in applied insect genomics. Genome Biol. *4*, 107.

Falciani,F., Hausdorf,B., Schröder,R., Akam,M., Tautz,D., Denell,R., and Brown,S. (1996). Class 3 Hox genes in insects and the origin of *zen*. Proc. Natl. Acad. Sci. USA *93*, 8479-8484.

Fialho,R.F. and Stevens,L. (2000). Male-killing Wolbachia in a flour beetle. Proc. R. Soc. Lond B Biol. Sci. *267*, 1469-1473.

Friedrich,M. and Benzer,S. (2000). Divergent *decapentaplegic* expression patterns in compound eye development and the evolution of insect metamorphosis. J. Exp. Zool. *288*, 39-55.

Friedrich,M. and Muqim,N. (2003). Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castanaeum*. Mol. Phylogenet. Evol. *26*, 502-512.

Fujioka,M., Yusibova,G.L., Patel,N.H., Brown,S.J., and Jaynes,J.B. (2002). The repressor activity of Even-skipped is highly conserved, and is sufficient to activate *engrailed* and to regulate both the spacing and stability of parasegment boundaries. Development *129*, 4411-4421.

Guichet,A., Copeland,J.W.R., Erdélyi,M., Hlousek,D., Závorszky,P., Ho,J., Brown,S., Percival-Smith,A., Krause,H.M., and Ephrussi,A. (1997). The nuclear receptor homologue Ftz-F1 and the homeodomain protein Ftz are mutually dependent cofactors. Nature *385*, 548-552.

Hinton,H.E. (1948). A synopsis of the genus *Tribolium* Macleay with some remarks on the evolution of its species groups. Bull. Ent. Res. *39*, 13-35.

Howard,R.W., Jurenka,R.A., and Blomquist,G.J. (1986). Prostaglandin synthetase inhibitors in the defensive secretion of the red flour beetle *Tribolium castaneum*, (Herbst) (Coleoptera:Tenebrionidae). Insect Biochem. *16*, 757-760.

Hurst,L.D. (1993). *scat+* is a selfish gene analogous to *Medea* of *Tribolium castaneum*. Cell *75*, 407-408.

Jockusch,E.L., Nulsen,C., Newfeld,S.J., and Nagy,L.M. (2000). Leg development in flies versus grasshoppers: differences in *dpp* expression do not lead to differences in the

expression of downstream components of the leg patterning pathway. Development *127*, 1617-1626.

Juan,C., Vazquez,P., Rubio,J.M., Petitpierre,E., and Hewitt,G.M. (1993). Presence of highly repetitive DNA sequences in *Tribolium* flour-beetles. Heredity (Edinburgh) *70*, 1-8.

Kamath,R.S., Fraser,A.G., Dong,Y., Poulin,G., Durbin,R., Gotta,M., Kanapin,A., Le Bot,N., Moreno,S., Sohrmann,M., Welchman,D.P., Zipperlen,P., and Ahringer,J. (2003). Systematic functional analysis of the Caenorhabditis elegans genome using RNAi. Nature *421*, 231-237.

Kaufman,T.C., Severson,D.W., and Robinson,G.E. (2002). The Anopheles genome and comparative insect genomics. Science *298*, 97-98.

Kristensen,N. (1999). Phylogeny of edopterygote insects, the most successful lineage of living organisms. European Journal of Entomology *96*, 237-253.

Lewis,D.L., DeCamillis,M., and Bennett,R.L. (2000). Distinct roles of the homeotic genes *Ubx* and *abd-A* in beetle embryonic abdominal appendage development. Proc. Natl. Acad. Sci. U. S. A *97*, 4504-4509.

Lohr,U., Yussa,M., and Pick,L. (2001). Drosophila fushi tarazu. a gene on the border of homeotic function. Curr. Biol. *11*, 1403-1412.

Lorenzen,M.D., Brown,S.J., Denell,R.E., and Beeman,R.W. (2002). Transgene expression from the *Tribolium castaneum Polyubiquitin* promoter. Insect Mol. Biol. *11*, 399-407.

Maderspacher,F., Bucher,G., and Klingler,M. (1998). Pair-rule and gap gene mutants in the flour beetle *Tribolium castaneum*. Dev. Genes Evol. *208*, 558-568.

Maeda,I., Kohara,Y., Yamamoto,M., and Sugimoto,A. (2001). Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. Curr. Biol. *11*, 171-176.

Maxton-Kuchenmeister,J., Handel,K., Schmidt-Ott,U., Roth,S., and Jackle,H. (1999). *Toll* homologue expression in the beetle Tribolium suggests a different mode of dorsoventral patterning than in Drosophila embryos. Mech. Dev. *83*, 107-114.

Nagy,L.M. and Carroll,S. (1994). Conservation of *wingless* patterning functions in the short-germ embryos of *Tribolium castaneum*. Nature *367*, 460-463.

Oppenheimer,D.I., MacNicol,A.M., and Patel,N.H. (1999). Functional conservation of the wingless-engrailed interaction as shown by a widely applicable baculovirus misexpression system. Curr. Biol. *9*, 1288-1296.

Peters,L.L. and Barker,J.E. (1993). Novel inheritance of the murine severe combined anemia and thrombocytopenia (Scat) phenotype. Cell *74*, 135-142.

Plohl,M., Lucijanic-Justic,V., Ugarkovic,D., Petitpierre,E., and Juan,C. (1993). Satellite DNA and heterochromatin of the flour beetle *Tribolium confusum*. Genome *36*, 467-475.

Prendeville,H.R. and Stevens,L. (2002). Microbe inhibition by Tribolium flour beetles varies with beetle species, strain, sex, and microbe group. J. Chem. Ecol. *28*, 1183-1190.

Prpic,N.M., Wigand,B., Damen,W.G., and Klingler,M. (2001). Expression of *dachshund* in wild-type and Distal-less mutant Tribolium corroborates serial homologies in insect appendages. Dev. Genes Evol. *211*, 467-477.

Rascon,J.L., Styer,L.M., and Scott,T.W. (2003). Wolbachia-induced mortality as a mechanism to modulate pathogen transmission by vector arthropods. J. Med. Entomol. *40*, 125-132.

Sanchez-Salazar,J., Pletcher,M.T., Bennett,R.L., Brown,S.J., Dandamudi,T.J., Denell,R.E., and Doctor,J.S. (1996). The *Tribolium decapentaplegic* gene is similar in sequence, structure, and expression to the *Drosophila dpp* gene. Roux's Arch. Dev. Biol. *206*, 237-246.

Schroder,R. (2003). The genes *orthodenticle* and *hunchback* substitute for *bicoid* in the beetle Tribolium. Nature *422*, 621-625.

Schroder,R., Eckert,C., Wolff,C., and Tautz,D. (2000). Conserved and divergent aspects of terminal patterning in the beetle *Tribolium castaneum*. Proc. Natl. Acad. Sci. USA *97*, 6591-6596.

Schulz,C., Schröder,R., Hausdorf,B., Wolff,C., and Tautz,D. (1998). A *caudal* homologue in the short germ band beetle *Tribolium* shows similarities to both, the *Drosophila* and the vertebrtae *caudal* expression patterns. Dev. Genes Evol. *208*, 283-289.

Shippy,T.D., Guo,J., Brown,S.J., Beeman,R.W., Haas,M.S., and Denell,R.E. (2000). Analysis of *maxillopedia* expression pattern and larval cuticular phenotypes of wild type and mutant Tribolium. Genetics *155*, 721-731.

Skaer,N., Pistillo,D., Gibert,J.M., Lio,P., Wulbeck,C., and Simpson,P. (2002). Gene duplication at the *achaete-scute* complex and morphological complexity of the peripheral nervous system in Diptera. Trends Genet. *18*, 399-405.

Sommer,R.J. and Tautz,D. (1993). Involvement of an orthologue of the *Drosophila* pair-rule gene *hairy* in segment formation of the short germ-band embryo of *Tribolium* (Coleoptera). Nature *361*, 448-450.

Sommer,R.J. and Tautz,D. (1994). Expression patterns of twist and snail in Tribolium (Coleoptera) suggest a homologous formation of mesoderm in long and short germ band insects. Dev. Genet. *15*, 32-37.

Stauber,M., Prell,A., and Schmidt-Ott,U. (2002). A single Hox3 gene with composite *bicoid* and *zerknullt* expression characteristics in non-Cyclorrhaphan flies. Proc. Natl. Acad. Sci. U. S. A *99*, 274-279.

Stein,L.D., Mungall,C., Shu,S., Caudy,M., Mangone,M., Day,A., Nickerson,E., Stajich,J.E., Harris,T.W., Arva,A., and Lewis,S. (2002). The generic genome browser: a building block for a model organism system database. Genome Res. *12*, 1599-1610.

Stuart,J.J., Brown,S.J., Beeman,R.W., and Denell,R.E. (1993). The *Tribolium* homeotic gene *Abdominal* is homologous to *abdominal-A* of the *Drosophila* bithorax complex. Development *117*, 233-243.

Stuart,J.J., De Gortari,M.J., Hall,P.S., Maxwell,M.E., Mocelin,G., Brown,S.J., and Muir,W.M. (1996). Useful DNA polymorphisms are identified by snapback, a midrepetitive element in *Tribolium castaneum*. Genome *39*, 568-578.

Stuart,J.J. and Mocelin,G. (1995). Cytogenetics of chromosome rearrangements in *Tribolium castaneum*. Genome *38*, 673-680.

Sulston,I.A. and Anderson,K.V. (1996). Embryonic patterning mutants in *Tribolium castaneum*. Development *122*, 805-814.

Torres,M. and Sánchez,L. (1989). The *scute* (T4) gene acts as a numerator element of the X:A signal that determines the state of activity of *Sex- lethal* in *Drosophila*. EMBO J. *8*, 3079-3086.

Wade,M.J. and Stevens,L. (1985). Microorganism mediated reproductive isolation in flour beetles (genus Tribolium). Science *227*, 527-528.

Watson,A.H. and Schurmann,F.W. (2002). Synaptic structure, distribution, and circuitry in the central nervous system of the locust and related insects. Microsc. Res. Tech. *56*, 210-226.

Wheeler,S.R., Carrico,M.L., Brown,S.J., and Skeath,J.B. (2003). Isolation of achaete-scute genes from Tribolium castaneum reveals conservation and variation in neural pattern formation and cell fate specification. Development in press.

Whitfield,C.W., Band,M.R., Bonaldo,M.F., Kumar,C.G., Liu,L., Pardinas,J.R., Robertson,H.M., Soares,M.B., and Robinson,G.E. (2002). Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. Genome Res. *12*, 555-566.

Wolff,C., Sommer,R., Schröder,R., Glaser,G., and Tautz,D. (1995). Conserved and Divergent aspects of the *Drosophila* segmentation gene *hunchback* in the short germ band embryo of the flour beetle *Tribolium*. Development *121*, 4227-4236.

Yan,G. and Norman,S. (1995). Infection of Tribolium beetles with a tapeworm: variation in susceptibility within and between beetle species and among genetic strains. J. Parasitol. *81*, 37-42.

Zhong,D., Pai,A., and Yan,G. (2003). Quantitative trait loci for susceptibility to tapeworm infection in the red flour beetle. Genetics *in revision*.