Insect Molecular Biology

Insect Molecular Biology (2010) **19**(3), 337–346 doi: 10.1111/j.1365-2583.2010.00992.x

An expansion of the genome size dataset for the insect order Hymenoptera, with a first test of parasitism and eusociality as possible constraints

A. M. Ardila-Garcia*‡, G. J. Umphrey† and T. R. Gregory*

**Department of Integrative Biology, †Department of Mathematics and Statistics, University of Guelph, Guelph, ON, Canada*

Abstract

Although the Hymenoptera represent a remarkably diverse and socioeconomically important group that is of considerable interest in genome biology, they remain understudied in terms of genome size. This study reports new genome size estimates for 89 species of ants, bees and wasps, representing 17 families and four superfamilies. These are used in a test of the hypothesis that genome sizes are constrained by traits associated with parasitism or eusociality. Not all parasitoid wasps exhibit small genomes, though a relationship based on specific types of parasitism may still occur; by contrast, there was no convincing evidence of a constraint relating to eusociality. The data provided here can be used to guide future research aimed at understanding the evolution of large-scale genomic properties in this order.

Keywords: ants, bees, *C*-value, flow cytometry, Feulgen image analysis densitometry, parasitoid, wasps.

Introduction

With 125 000 described species and perhaps 5–10 times as many awaiting discovery, the insect order Hymenoptera is one of the most species-rich animal groups on Earth (Grimaldi & Engel, 2005). As a result of their enormous diversity, their socioeconomic importance (e.g. as pollinators or biocontrol agents), their intriguing ecological and social characteristics (e.g. parasitism and eusociality) and their haplodiploid sex determination mechanism, the Hymenoptera have long been a subject of considerable interest (Page *et al.*, 2002; Hölldobler & Wilson, 2008).

Not surprisingly, this interest in the Hymenoptera has carried over into modern comparative genomics. 'The completed genome sequence of the honey bee (*Apis mellifera*) has recently been complemented by whole-genome sequences from three parasitoid wasps from the genus *Nasonia*' (Honeybee Genome Sequencing Consortium, 2006; Liolios *et al.*, 2008; *Nasonia* Genome Working Group, 2010; www.genomesonline.org). The choice of these species as sequencing targets highlights the relevance of large-scale genomics for advancing the understanding of eusociality ('sociogenomics'), parasitism, and development (e.g. Pultz & Leaf, 2003; Robinson *et al.*, 2005; Cristino *et al.*, 2006; Honeybee Genome Sequencing Consortium, 2006; Wilson, 2006; Goodisman *et al.*, 2008; *Nasonia* Genome Working Group, 2010).

As informative as comparisons of genes, and especially whole genome sequences, are likely to be, they do not encompass all potential relationships between genomic properties and hymenopteran biology. The size of the genome, measured in terms of total mass of DNA, has also been hypothesized to relate to eusociality, parasitism, and development in insects (e.g. Gregory, 2002; Johnston *et al.*, 2004; Koshikawa *et al.*, 2008). This is based on the associations between genome size, cell size, and cell division rates that have been identified in many taxa (Gregory, 2005). Larger genomes generally mean larger, more slowly dividing cells, and this may be linked to parasitism because of constraints on small body size or rapid developmental rate relative to hosts (e.g. Johnston *et al.*, 2004) and to eusociality because of consequences for neuron size and number (e.g. Roth *et al.*, 1994).

First published online 26 February 2010.

Correspondence: T. Ryan Gregory, Department of Integrative Biology, University of Guelph, Guelph, ON, Canada N1G 2W1. Tel.: +1 519 824 4120, x58053; fax: +1 519 767 1656; e-mail: rgregory@uoguelph.ca

‡Current address: Department of Botany, University of British Columbia, Vancouver, BC, Canada V6T 1Z4.

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society

Surprisingly, the Hymenoptera have not generally featured in studies relating to these hypotheses, even though they represent an ideal group that includes both eusocial and solitary species and parasitoid and non-parasitoid species. In fact, until very recently, only 10 genome size estimates were available for the entire order Hymenoptera (Gregory, 2009; http://www.genomesize.com). Fortunately, this has begun to change: the hymenopteran dataset was recently expanded to include 40 new estimates for ants by Tsutsui *et al.* (2008), though this still leaves other groups largely or entirely unexplored.

The order Hymenoptera has traditionally been divided into two major groups: the (paraphyletic) Symphyta, which includes sawflies and their relatives, and the Apocrita, which includes the more familiar ants, bees and 'wasps' (the latter also being paraphyletic). The present study aimed foremost to expand current knowledge regarding genome size diversity in the Hymenoptera by surveying some of the most familiar and economically important groups within the Apocrita: the superfamilies Apoidea, Chalcidoidea, Ichneumonoidea and Vespoidea. This also allowed a first exploration of potential links between genome size, parasitism and eusociality within this order. Notably, the superfamilies Chalcidoidea and Ichneumonoidea include parasitoids, making it possible to compare these with non-parasitoid groups. Similarly, the inclusion of eusocial and solitary species from the superfamilies Vespoidea and Apoidea allows an early examination of the proposed relationship between genome size and eusociality that has not previously been possible because nearly all of the species studied to date (i.e. mostly ants) are eusocial.

The results of this study indicate that any possible links between genome size, parasitism and eusociality in the Hymenoptera are more complex than previous hypotheses may suggest. They also highlight areas in which future work carried out within a broad phylogenetic and ecological context will be useful in further exploring large-scale genome evolution in this group.

Results

Overall patterns

Genome size estimates for the 89 species reported in this study are provided in Table 1. In combination with previous reports, this brings the total number of hymenopteran species studied to 131 (Gregory, 2009). Reported values range roughly 11-fold, from 0.10 pg in the braconid wasps *Aphidius colemani* and *Peristenus stygicus* to 1.14 pg in the black and yellow mud dauber *Sceliphron caementarium*. (Note that a second estimate for *S. caementarium* was lower; see below). Taking all available data from the present study and the Animal Genome Size Database (Gregory, 2009), the mean genome size for the Hymenoptera studied so far is 0.37 pg $\pm$ 0.18SD. The mean for data in the present study alone was very similar at 0.38 pg $\pm$ 0.20SD.

A test for homogeneity among the four superfamily means was conducted using one-way ANOVA, treating the mean genome size for each family as a replicate. The null hypothesis of equal superfamily means was rejected at the 5% level ($F_{3,14} = 6.7023$, $P = 0.0049$). The parasitoid superfamily Ichneumonoidea displays the smallest range and lowest mean genome size (range: 0.10–0.27 pg; mean: 0.15 $\pm$ 0.05SD, $n = 10$), followed by the non-parasitoid Vespoidea (range: 0.15–0.71 pg; mean: 0.34 pg $\pm$ 0.12SD, $n = 80$), the parasitoid Chalcidoidea (range: 0.18–0.75 pg; mean: 0.43 $\pm$ 0.19SD, $n = 12$), and finally the non-parasitoid Apoidea (range: 0.19–1.14 pg; mean: 0.53 pg $\pm$ 0.22SD, $n = 29$) (Table 1; Gregory, 2009; Fig. 1). A Tukey HSD pairwise comparison showed that the difference between the Apoidea and both the Ichneumonoidea and Vespoidea is significant.

Family means were compared within each superfamily using either a one-way ANOVA with the subsequent application of Tukey's HSD method at the 5% level if appropriate or (for the Ichneumonoidea) a two-sample pooled variance *t*-test. Homogeneity of family means within the Chalcidoidea was rejected at the 5% level ($F_{4,7} = 8.202$, $P = 0.0089$). Tukey's method showed that both the Aphelinidae and Encyrtidae means are higher than the Trichogrammatidae mean. Homogeneity of family means within the Vespoidea was rejected at the 5% level ($F_{3,76} = 3.942$, $P = 0.0114$), with Tukey's method showing that the Mutillidae mean is higher than the Vespidae mean. Homogeneity among family means within the Apoidea could not be rejected at the 5% level ($F_{6,22} = 0.612$, $P = 0.718$). Within the Ichneumonoidea the *t*-test showed that the Ichneumonidae mean is higher than the Braconidae mean ($t_8 = -5.33$, two-tailed $P = 0.0007$), though the Ichneumonidae was represented by a single estimate.

Parasitism and eusociality

Grouping the various species that have been studied as either parasitoid, eusocial, or solitary provides an opportunity to test functional associations between lifestyle and genome size. Across all Hymenoptera, the mean for non-parasitoids taken together (0.39 pg $\pm$ 0.18SD, $n = 109$) is larger than that of parasitoids (0.30 pg $\pm$ 0.20SD, $n = 22$) (*t*-test, $P = 0.04$). Parasitoids do not differ significantly from eusocial species (0.34 pg $\pm$ 0.12SD, $n = 78$) (*t*-test, $P > 0.2$) but the genomes in both groups are significantly smaller than those of non-parasitoid solitary species (0.48 pg $\pm$ 0.26SD, $n = 24$) (ANOVA, $P < 0.01$). There is no significant difference between ants (by far the best studied

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

Table 1. Genome size estimates for 89 species in the order Hymenoptera

Taxonomy	GS	SE	N	Method	Location	Lifestyle
Order Hymenoptera						
Suborder Apocrita						
Superfamily Apoidea						
Family Andrenidae (Adrenid bees)						
Andrena dunningi	0.50	0.009	1M, 3F	FCM	1	S
Family Apidae (Honey bees, bumble bees, carpenter bees and their relatives)						
Apis mellifera	0.24	0.004	5F	FCM	1	EU
Bombus bimaculatus	0.34	0.008	3F	FCM	1	EU
Bombus impatiens	0.47	0.027	4F	FCM	1	EU
Ceratina calcarata	0.68	0.008	2M, 2F	FCM	1	S
Ceratina dupla dupla	0.59	–	1F	FCM	1	S
Melissodes desponsa	0.52	–	1F	FCM	1	?
Melissodes illata	0.37	0.007	3F	FCM	1	?
Xylocopa virginica krombeini	0.69	–	1F	FIAD	4	?
Family Colletidae (Plasterer bees and yellow-faced bees)						
Hylaeus affinis	0.64	0.011	4F	FCM	2	S
Family Crabronidae (Digger wasps and relatives)						
Ectemnius continuus	0.38	–	1F	FCM	1	S
Gorytes atricornis	0.48	–	1F	FCM	1	S
Microbembex monodonta	0.66	–	1F	FIAD	2	S
Trypoxylon politum	0.35	0.005	2F	FIAD	2	S
Family Halictidae (Sweet bees and their relatives)						
Agapostemon splendens	0.66	0.007	6M, 4F	FIAD	4	S
Augochloropsis metallica	0.90	–	1F	FCM	1	S
Halictus ligatus	0.60	0.017	6F	FCM	1	EU
Halictus ligatus	0.49	0.011	8F	FIAD	4	EU
Halictus poeyi	0.40	0.001	3F	FIAD	2	EU
Family Megachilidae (Leafcutter bees, mason bees and their relatives)						
Anthidiellum notatum rufimaculatum	0.48	–	1F	FIAD	4	?
Megachile albitarsis	0.80	–	1F	FIAD	4	?
Megachile rotundata	0.83	–	1F	FCM	1	S
Family Sphecidae (Thread-waisted wasps)						
Ammophila pictipennis	0.41	–	1F	FIAD	5	S
Chalybion californicum	0.54	0.010	4M, 2F	FCM	1	S
Larra bicolor	0.19	0.005	3F	FCM	1	S
Miscophus slossonae	0.19	0.002	2F	FIAD	2	S
Sceliphron caementarium	1.14	0.045	1M, 1F	FCM	1	S
Sceliphron caementarium	0.81	0.009	2F	FIAD	5	S
Stictiella serrata	0.78	–	1F	FIAD	2	S
Superfamily Chalcidoidea						
Family Aphelinidae						
Aphelinus abdominalis	0.65	0.002	7F	FCM	6	P
Encarsia formosa	0.42	0.002	7F	FCM	6	P
Eretmocerus eremicus	0.55	0.008	7M, 1F	FCM	6	P
Eretmocerus mundus	0.75	0.008	5M, 2F	FCM	6	P
Family Encyrtidae						
Copidosoma floridanum	0.57	0.01	3M	FIAD*	8	P
Leptomastix dactylopii	0.56	0.008	6F	FCM	6	P
Family Eulophidae						
Diglyphus isaea	0.23	0.007	5M, 1F	FCM	6	P
Family Trichogrammatidae						
Trichogramma platneri	0.18	0.007	3M, 7F	FCM	7	P
Trichogramma pretiosum	0.19	0.007	1M, 10F	FCM	7	P
Superfamily Ichneumonoidea						
Family Braconidae (Braconid wasps)						
Aphidius colemani	0.10	0.001	4M, 8F	FCM	6	P
Aphidius ervi	0.14	0.002	5M, 2F	FCM	6	P
Aphidius ervi	0.11		3M (pooled)	FIAD*	8	P
Dacnusa sibirica	0.16	0.001	3M, 3F	FCM	6	P
Macrocentrus grandii	0.11	–	3M (pooled)	FIAD*	8	P
Peristenus digoneutis	0.13	–	3M (pooled)	FIAD*	8	P
Peristenus stygicus	0.10	–	3M (pooled)	FIAD*	8	P
Praon sp.	0.13	–	2M (pooled)	FIAD*	8	P
Family Ichneumonidae (Ichneumonid wasps)						
Neotheronia bicincta	0.27	0.006	3F	FIAD	5	P

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

Table 1. *Continued*

Taxonomy	GS	SE	N	Method	Location	Lifestyle
Superfamily Vespoidea						
Family Formicidae (Ants)						
Amblyopone pallipes	0.37	–	1F	FCM	1	EU
Aphaenogaster fulva	0.42	0.002	3F	FCM	1	EU
Aphaenogaster (rudis-texana group, N16*)*	0.43	–	1F	FCM	1	EU
Aphaenogaster (rudis-texana group, N17*)*	0.46	0.011	4F	FCM	1	EU
Aphaenogaster (rudis-texana group, N22b*)*	0.44	0.001	3F	FCM	1	EU
Aphaenogaster treatae	0.50	0.019	3F	FCM	1	EU
Atta texana	0.27	0.009	6F	FCM	3	EU
Camponotus floridanus	0.23	0.002	3F	FIAD	2	EU
Dolichoderus mariae	0.18	<0.001	3F	FCM	1	EU
Dolichoderus taschenbergi	0.23	0.008	3F	FCM	1	EU
Dorymyrmex bureni	0.18	0.003	3F	FIAD	2	EU
Forelius pruinosus	0.22	0.003	2F	FIAD	2	EU
Lasius (Acanthomyops) latipes	0.27	0.004	2M, 9F	FCM	1	EU
Lasius minutus	0.23	0.001	4F	FCM	1	EU
Monomorium viride	0.50	0.019	3F	FIAD	2	EU
Odontomachus brunneus	0.33	0.014	3F	FIAD	2	EU
Paratrechina longicornis	0.18	0.005	3F	FIAD	2	EU
Pheidole dentata	0.24	0.003	3F	FIAD	2	EU
Pheidole floridana	0.21	0.004	5F	FIAD	2	EU
Ponera pennsylvanica	0.55	0.014	6F	FCM	1	EU
Pseudomyrmex ejectus	0.29	–	1F	FIAD	2	EU
Pseudomyrmex gracilis	0.35	0.013	5F	FCM, FIAD	1, 2	EU
Solenopsis invicta	0.47	0.005	4F	FIAD	2	EU
Solenopsis molesta	0.38	0.004	5F	FCM	1	EU
Tapinoma sessile	0.37	0.007	10F	FCM	1	EU
Temnothorax ambiguus	0.31	0.003	9F	FCM	1	EU
Temnothorax texanus	0.32	0.017	5F	FCM	1	EU
Tetramorium caespitum	0.27	0.009	3F	FCM	1	EU
Trachymyrmex septentrionalis	0.25	0.002	3M, 6F	FIAD	2	EU
Family Mutillidae (Velvet ants)						
Dasymutilla archboldi	0.46	0.015	2F	FIAD	2	S
Dasymutilla occidentalis	0.46	–	1F	FIAD	2	S
Dasymutilla pyrrhus	0.49	–	1F	FIAD	2	S
Family Scoliidae (Scoliid wasps)						
Campsomeris plumipes fossulana	0.21	0.004	4F	FIAD	4	S
Family Vespidae						
Dolichovespula arenaria	0.32	–	1F	FCM	1	EU
Eumenes fraternus	0.22	–	1F	FIAD	5	S
Eumenes smithii	0.30	–	1F	FIAD	2	S
Euodynerus cf. hidalgo	0.30	–	1F	FIAD	4	S
Polistes dominula	0.29	0.004	6F	FCM	1	EU
Polistes dorsalis	0.48	0.009	7F	FIAD	2, 4, 5	EU
Polistes fuscatus	0.41	0.006	4F	FCM	1	EU
Polistes fuscatus	0.31	0.019	2F	FIAD	5	EU
Polistes metricus	0.24	0.005	4F	FIAD	2, 4, 5	EU
Symmorphus canadensis	0.23	–	1F	FCM	1	S
Vespula germanica	0.23	0.003	1M, 3F	FCM	1	EU
Vespula maculifrons	0.22	0.003	4F	FCM	1	EU
Vespula squamosa	0.17	0.005	3F	FIAD	4	EU
Vespula vulgaris	0.18	–	1F	FCM	1	EU
Zethus slossonae	0.15	–	1F	FIAD	4	S

Haploid genome size (GS, in pg), standard error (SE), and number of individuals (N; M = male, F = female) are indicated. Two methods were used, flow cytometry (FCM) using brain tissue versus *Drosophila melanogaster* and Feulgen image analysis densitometry (FIAD) using haemocytes versus *Tenebrio molitor* or sperm (FIAD*) vs. *D. melanogaster*. Lifestyle, including solitary (S), eusocial (EU), or parasitoid (P), is listed.
Locations of collection: (1) Guelph, ON; (2) Archbold Biological Station, Lake Placid, FL; (3) Brownsville, TX; (4) Orlando, FL; (5) Kissimmee, FL; (6) Biobest Canada; (7) Beneficial Insectary Canada; and (8) Miodrag Grbic, laboratory colony (University of Western Ontario).

eusocial group) and non-ant eusocial species (*t*-test, $P = 0.5$).

Four aculeate families consisting of eusocial species were sampled in the present analysis, two in the superfamily Apoidea and two in the Vespoidea (Fig. 2). In the case of the Apoidea, data from several non-eusocial families are also available for comparison. In general, there is substantial overlap in genome size ranges between euso-

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

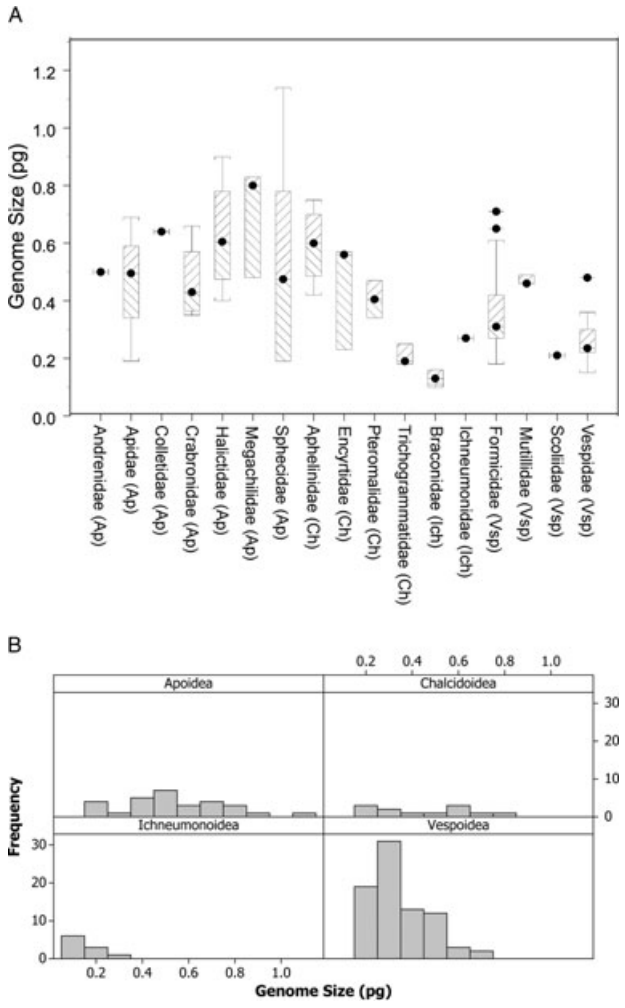

Figure 1. (A) Box plot of genome sizes for families of Hymenoptera arranged according to superfamilies (Ap = Apoidea, Ch = Chalcidoidea, Ich = Ichneumonoidea, Vsp = Vespoidea), including data from the present study and the Animal Genome Size Database (Gregory, 2009). Box plot specifications follow Pagano & Gauvreau (1993); boxes extend from first to third quartiles, with medians marked within the boxes. Whiskers extend to smallest and largest observations that fall within 1.5 box lengths from box ends, and observations beyond this are marked as outliers. (B) Histograms of reported genome size data in four superfamilies of Hymenoptera.

cial families and their close non-eusocial relatives, with no evidence of demarcation between these categories when examined phylogenetically (Fig. 2). Only one non-eusocial species has been studied within the superfamily Vespoidea, the solitary wasp *Campsomeris plumipes* (family Scoliidae), which has a genome size of 0.21 pg and is well within the range of the two related families that contain eusocial species, the Formicidae (ants) and Vespidae (wasps) (Fig. 2). Furthermore, there were no significant differences in mean genome sizes between the nine eusocial and five solitary species sampled in the family Vespidae (0.28 pg ± 0.10SD vs. 0.24 pg ± 0.06SD; *t*-test, $P > 0.4$). In fact, comparing three vespid subfamilies reveals that the eusocial Polistinae (0.34 pg ± 0.10SD, $n = 4$) exhibit the largest genomes, with no significant difference between the eusocial Vespinae (0.22 pg ± 0.06SD, $n = 5$) and solitary Eumeninae (0.24 pg ± 0.06, $n = 5$).

Overall, the seven bee families analysed showed similar genome size ranges (Fig. 2). The eusocial honey bees *A. mellifera* (0.24 pg) and *Apis cerana* (0.19 pg) display some of the smallest reported hymenopteran genome sizes (Jordan & Brosemer, 1974; Table 1), but in fact both the lowest and highest estimated genome sizes for bees were found within the solitary family Sphecidae. Within both the Apidae and Halictidae, solitary species exhibited larger genomes than those of eusocial species, but this was based on only a few species per family (Table 1).

Intraspecific differences and/or methodological discrepancies

In the few species for which both flow cytometry (FCM) and Feulgen image analysis densitometry (FIAD) were used, estimates obtained with FIAD tended to be lower (Table 1). This could relate to differences in tissue and/or standard used (haemocytes or spermatozoa in FIAD, brain in FCM). This source of error does not affect the general patterns described above, but it highlights the importance of further standardization in future large-scale insect studies. In the case of the wasp *S. caementarium*, FCM provided an estimate of 0.81 pg for samples from Ontario, whereas FIAD gave an estimate of 1.14 pg for specimens from Florida. This could represent differences in methodology, but it is worth noting that apparent intraspecific variation was also reported in a few ants by Tsutsui *et al.* (2008). Finally, it should be noted that discrepancies remain regarding the estimated genome size of the fire ant *Solenopsis invicta*, which was reported as 0.77 pg using reassociation kinetics (Li & Heinz, 2000), 0.62 pg with FCM (Johnston *et al.*, 2004), and 0.48 pg by FIAD (present study). The latter value is the most similar to estimates from other members of the genus obtained using FCM (*Solenopsis xyloni*, 0.48 pg; Tsutsui *et al.*, 2008; *Solenopsis molesta*, 0.38 pg; present study).

Although few estimates were available from multiple labs for the same species, some additional analyses were possible using a pooled dataset including the 89 new estimates reported here and 42 others from the Animal Genome Size Database (Gregory, 2009); all but seven of the other estimates are for ant species as reported by Tsutsui *et al.* (2008). Multiple regression analyses were performed to test the significance of the extra Sum of Squares for the location of laboratory analysis (Laboratory Location, Guelph vs. Other), with either Superfamily or

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

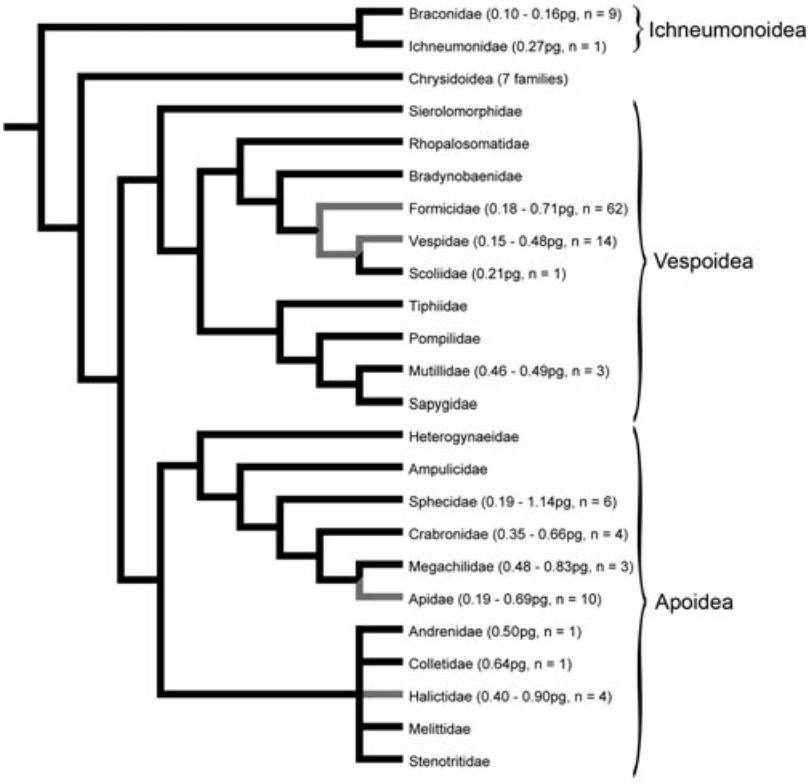

Figure 2. Phylogeny of the Aculeata (superfamilies Apoidea, Chrysidoidea, and Vespoidea) and associated distribution of known genome size diversity. Eusocial lineages are indicated by grey branches. No estimates are yet available from the superfamily Chrysidoidea. Phylogeny based on Grimaldi & Engel (2005).

Family as an additional categorical variable in the model. The *F*-test for the addition of Laboratory Location was not significant at the 5% level when either Superfamily ($F_{1,126} = 0.917$, $P = 0.34$) or Family ($F_{1,112} = 1.512$, $P = 0.22$) was included in the model. The reduction in the multiple coefficient of determination is very minor if Laboratory Location is dropped from the model (with Superfamily as a categorical variable, $r^2 = 0.3269$ with Laboratory Location included and $r^2 = 0.3220$ with Laboratory Location excluded; with Family as a categorical variable, $r^2 = 0.4894$ with Laboratory Location included and $r^2 = 0.4825$ with Laboratory Location excluded).

In order to further test the compatibility of large, independent datasets, comparisons were made between the non-overlapping ant genome size datasets from the present study and Tsutsui *et al.* (2008) using both a pooled variance two-sample *t*-test and a Wilcoxon rank-sum test. Neither test was significant at the 5% level, however, both were significant at the 10% level (*t*-test, $t_{60} = 1.8701$, $P = 0.0663$; Wilcoxon rank-sum test, $z = 1.9044$, $P = 0.0569$). These relatively minor differences could be accounted for primarily by taxonomic and geographic differences in sampling, rather than methodology (most estimates in both studies were obtained using the same general protocol; DeSalle *et al.*, 2005), indicating that pooling independent datasets in future analyses is not problematic if best practices have been followed.

Discussion

Genome size, parasitism and eusociality

At present, genome size estimates are available for 131 species of Hymenoptera (Gregory, 2009), most of these having been reported very recently (Tsutsui *et al.*, 2008; present study). While this encompasses a minuscule portion of the order as a whole (around 0.1% of described species), it does provide the first opportunity to examine patterns both relative to other insects and within the order.

With the exception of some very poorly studied parasitic orders (Phthiraptera, Strepsiptera; Johnston *et al.*, 2004, 2007), the Hymenoptera exhibit the smallest average genome size and some of the smallest absolute values found to date in any insect order (Gregory, 2009). Thus, not only do they conform to the hypothetical 2 pg threshold proposed for holometabolous insects as a result of developmental constraints (Gregory, 2002), but additional factors appear to limit genome expansion in these insects.

Constraints on genome size relating specifically to body size, developmental rate, and flight have been proposed for other insects (e.g. Finston *et al.*, 1995; Gregory *et al.*, 2003; Johnston *et al.*, 2004; Gregory & Johnston, 2008; Ardila-Garcia & Gregory, 2009), and could apply equally to the Hymenoptera (but note that regularly flying eusocial bees and wasps do not exhibit smaller genomes than ants; Table 1). However, the present study focused on two

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

additional features for which the order is best known: parasitism and eusociality. Both of these have been suggested as possible correlates of genome size in insects on the basis of comparisons of small numbers of species from other orders (Johnston *et al*., 2004; Koshikawa *et al*., 2008). However, the present study is the first to explore these parameters within a single order using a large sample of species.

Sampling in the present study was aimed at broadly surveying common species rather than targeting particular clades or ecological characteristics, but this nonetheless provided an opportunity for some phylogenetically informed comparisons. According to one phylogenetic hypothesis (Whitfield, 1998; Grimaldi & Engel, 2005), the Ichneumonoidea is part of a clade that also includes the monophyletic, non-parasitoid group known as the Aculeata (including the superfamilies Apoidea, Vespoidea, and Chrysidoidea). (But see Pennacchio & Strand, 2006, in which the Ichneumonoidea is slightly more distantly related to the Aculeata). Parasitoids in the superfamily Ichneumonoidea exhibit smaller genomes than the aculeate superfamilies Apoidea or Vespoidea (the Chrysidoidea were not sampled). In fact, some of the smallest genome sizes ever reported for insects are found in the family Braconidae within the superfamily Ichneumonoidea. However, the estimate of 0.27 pg for *Neotherrinia bicincta* (family Ichneumonidae) shows that this is not true of all ichneumonoids as this is more than twice as large as the mean of 0.12 pg for braconids. Moreover, the present data show that small genomes are not typical of all parasitoids (Table 1). Specifically, the families Aphelinidae and Encyrtidae (but not Eulophidae or Trichogrammidae), which are part of the more distantly related superfamily Chalcidoidea, possess genome sizes intermediate between the non-parasitoid Vespoidea and Apoidea. Therefore, it is clear that parasitism alone is not sufficient to constrain genomes to a small size, given the comparatively large values observed in the Chalcidoidea (up to 0.75 pg, twice the hymenopteran average). Nor does a non-parasitic lifestyle necessarily result in the evolution of larger genomes, as shown by the wide range and small minimum values in the Apoidea. In short, any potential link between genome size and ecological lifestyle in this order must be more complex than a simple dichotomy between parasitism and non-parasitism.

There are several potential reasons for this. First, ectoparasitism is thought to be ancestral in the Apocrita (Whitfield, 1998; Pennacchio & Strand, 2006), with shifts to endoparasitism in several groups and loss of parasitism in others (shifting to predation, pollen and nectar feeding, or gall-forming). Ample opportunities exist to investigate potential impacts on genome size of such ecological shifts at several phylogenetic scales in the future, most notably: (1) through comparisons between the Symphyta and Apocrita; (2) among some families of the Aculeata where shifts have occurred; and (3) between gall-forming wasps versus parasitoid relatives (e.g. Cynipidae versus other Cynipoidea). However, it must be noted that there is an important distinction between the evolution of a feature that imposes a new constraint and creates selective pressure for genome size reduction versus the removal of a constraint that simply relaxes selective pressures and may then allow genomes to expand. Given this and the occurrence of other potential constraints independent of parasitism, it should not be surprising that many non-parasitic species display small genome sizes.

Second, the term 'parasitoid' is very general and encompasses many distinct lifestyles that may differ in their impacts on genomic properties. This includes both endoparasitoids versus ectoparasitoids, idiobionts (preventing any further host development) versus koinobionts (allowing host development to proceed), parasitism of widely different arthropod hosts with different body sizes, developmental rates, and immune characteristics, and specializations for attacking different host life stages (egg, larva, pupa or adult) (for reviews, see Whitfield, 1998; Pennacchio & Strand, 2006). Current information is not sufficient to address patterns at this finer scale, but this clearly represents an important area for further exploration. Again, because multiple shifts between endo- and ectoparasitism and between idiobiont and koinobiont lifestyles have occurred within the Hymenoptera (Whitfield, 1998; Dowton & Austin, 2001; Pennacchio & Strand, 2006), this can be conducted in a phylogenetically informed manner, including across the order as well as within specific families which include many different parasitic types (e.g. the hyperdiverse Braconidae).

The genome sizes of the eusocial species studied to date are smaller than those of solitary, non-parasitic species (and they do not differ on average from those of parasitoids). However, once again, the evidence that eusociality necessarily imposes a constraint on genome size is very limited. Indeed, comparisons of related families of eusocial and non-eusocial species in this study did not reveal any patterns consistent with such a constraint (Fig. 2). This contrasts with the recent report of Koshikawa *et al.* (2008), who compared termites (Isoptera) and cockroaches (Blattaria) within the context of differences in social complexity. Such a comparison is difficult to interpret due to small sample size and because it represents a single phylogenetic comparison across groups that differ in many ways besides their degree of sociality. The present study has provided a preliminary examination involving multiple comparisons of sister taxa within a single order, but this will need to be expanded to include additional lineages and more species per lineage before any conclusions can be drawn regarding links between genome size and sociality.

© 2010 The Authors

Though not as great as in some other insect orders, the range in genome size reported here for the Hymenoptera requires explanation in molecular as well as ecological terms. First, it remains to be determined what types of sequences (e.g. transposable elements) are responsible for the differences in genome size observed across species of Hymenoptera. The species currently subject to complete genome sequencing (*A. mellifera* and *Nasonia* spp.) exhibit small genomes, and *A. mellifera* is notable for lacking major transposable element (TE) families, with most of its existing TEs being *mariner* elements and only rare, inactive remnants of other types present (Honeybee Genome Sequencing Consortium, 2006). It is possible that selection against retrotransposon insertion is strong in haplo-diploid insects (Honeybee Genome Sequencing Consortium, 2006), though it remains to be seen whether such elements are similarly uncommon in Hymenoptera with larger genomes. The data presented in Table 1 can serve to highlight species that would be most suitable for addressing this issue.

Similarly, it will be important to explore possible karyotypic patterns with regard to genome size diversity among Hymenoptera. Diploid chromosome numbers of Hymenoptera range more than 50-fold, from $2n = 2$ to $2n = 106$, or even $2n = 120$ in some populations (both extremes occur in ants; Crosland & Crozier, 1986; Mariano *et al*., 2004, 2008). Interestingly, it has been suggested that chromosome numbers are lower in parasitic Hymenoptera and higher in eusocial species (Sherman, 1979; Gokhman, 2009). Although genome sizes range far less than chromosome numbers in Hymenoptera, it will be useful to compare genome sizes among species with differing chromosome numbers at varying phylogenetic scales to determine whether this extreme karyotypic evolution has been associated with gains or losses of DNA. Moreover, it should be possible to examine DNA content on a chromosome-by-chromosome basis to determine how the karyotypic distribution of genomic DNA varies among groups (Gokhman, 2009). Finally, it will be useful to consider genome size with regard to specific components of chromosomes such as satellite DNA, at least some of which appears to be conserved and transcribed in wasps (Palomeque & Lorite, 2008).

Finally, other genomic properties could be analysed in comparison with genome size as data continue to accrue. Specifically, hymenopteran genomes are of special interest in terms of various features, including recombination rate, methylation, and AT-content (Gadau *et al*., 2000; Honeybee Genome Sequencing Consortium, 2006; Kronforst *et al*., 2008). In terms of the former, it was noted that the honey bee exhibits a remarkably high recombination rate relative to other eukaryotes, and even as compared to several other hymenopterans (Gadau *et al*., 2000; Honeybee Genome Sequencing Consortium, 2006). Gadau *et al.* (2000) noted that this is unrelated to haplodiploidy or chromosome number, and suggested that it is an adaptive feature that generates diversity in social species with sterile workers. Based on the data in the present study and Gadau *et al.* (2000), recombination rates among the few species analysed also seem to be independent of genome size, though it may be important to consider DNA amount (especially repetitive sequences) per chromosome in this regard. Kronforst *et al.* (2008) indicated that methylation is quite variable among Hymenoptera (from 1 to 19% of the genome). Given that methylation is used to silence TEs in addition to performing a role in development, it would be interesting to consider methylation patterns in terms of genome size variability. Finally, the honeybee genome sequence revealed a high AT-content compared to other sequenced insect genomes (Honeybee Genome Sequencing Consortium, 2006). AT-content is correlated with genome size in vertebrates (Vinogradov, 1998), once again making a comparison between these parameters of possible interest within the Hymenoptera and across insects more generally.

In light of the above, it is clear that the Hymenoptera will represent a very important target for genomic research over the coming years. Knowledge about the large-scale genomic properties of Hymenoptera, including genome sizes, will be an important component of this research. Many questions remain regarding the factors that influence genome size diversity in this group and how this may be associated with the genetic and behavioural characteristics that have stimulated so much interest in the Hymenoptera. Thus, the data presented here are meant to instigate, not complete, genome size research in this remarkably diverse order.

Experimental procedures

Sources of specimens

Specimens were hand-collected in Guelph, Ontario and the surrounding area between June and September 2006 and 2007, and around Orlando, Florida and the Archbold Biological Station (Lake Placid, Florida) in May 2007. Samples from the ant *Atta texana* were collected in Brownsville, Texas. In addition, parasitoid wasp samples were donated by Beneficial Insectary, Biobest Canada, and Miodrag Grbic (University of Western Ontario). Specimens were identified by Gary Umphrey (University of Guelph), Matthias Buck (University of Guelph), Jason Gibbs (York University, Ontario), Stuart Fullerton (University of Central Florida) and Mark Deyrup (Archbold Biological Station). Specimen vouchers were stored at room temperature, either in 99% ethanol (ants) or pinned (others).

Genome size estimation

Genome size estimates for hymenopterans collected in Ontario were obtained using FCM according to DeSalle *et al.* (2005). For this, neural tissue from unknowns was co-prepared with stan-

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

dards consisting of *Drosophila melanogaster* Oregon R strain females (1C = 0.18 pg; Rasch *et al*., 1971). A single head from an unknown and one from *D. melanogaster* were ground together in a 2 ml Kontes Dounce grinder kept on ice and containing 0.5 ml of Galbraith buffer (per 1 l dH_2O: 8.8 g of $Na_2C_6H_5O_7$, 4.2 g of 3-[N-morpholino]-propane sulphonate, 1.99 g $MgCl_2$, 1.0 ml Triton X-100, 100 μl 10 mg/ml RNase A, adjusted to pH 7.2) using a Type A pestle. The suspension was then passed through 30 μm nylon mesh (Spectramesh) and stained with propidium iodide at a concentration of 50 μg/ml and analysed using a BD FACSCalibur flow cytometer with a 488 nm laser.

To better accommodate extended work in the field, samples collected in Florida were prepared as air-dried smears of haemolymph that could be shipped to the lab. These were then analysed with FIAD using haemocytes from *Tenebrio molitor* as a standard (1C = 0.52 pg; Juan & Petitpierre, 1989). Four parasitoid wasps were also measured by FIAD using the same protocol, but using spermatozoa against a *D. melanogaster* standard. Feulgen procedures followed Hardie *et al.* (2002), with slides post-fixed overnight in 85 methanol: 10 formalin: 5 glacial acetic acid, hydrolized for 2 h in 5N HCl, stained for 2 h in fresh Schiff reagent, and rinsed with a series of bisulphite and distilled water changes before being analysed using the Bioquant image analysis software package.

Acknowledgements

We thank Beneficial Insectary, Biobest Canada, Matthias Buck, Mark Deyrup, Stuart Fullerton, Jason Gibbs, and Miodrag Grbic for providing samples and/or assisting with identification. Supported by a Natural Sciences and Engineering Research Council of Canada grant to TRG and a University of Guelph student travel award to AMAG.

References

Ardila-Garcia, A.M. and Gregory, T.R. (2009) An exploration of genome size diversity in dragonflies and damselflies (Insecta: Odonata). *J Zool* **278**: 163–173.

Cristino, A.S., Nunes, F.M.F., Lobo, C.H., Bitondi, M.M.G., Simoes, Z.L.P., Da Fontoura Costa, L. *et al.* (2006) Caste development and reproduction: a genome-wide analysis of hallmarks of insect eusociality. *Insect Mol Biol* **15**: 703–714.

Crosland, M.W.J. and Crozier, R.H. (1986) *Myrmecia pilosula*, an ant with only one pair of chromosomes. *Science* **231**: 1278.

DeSalle, R., Gregory, T.R. and Johnston, J.S. (2005) Preparation of samples for comparative studies of arthropod chromosomes: visualization, *in situ* hybridization, and genome size estimation. *Methods Enzymol* **395**: 460–488.

Dowton, M. and Austin, A.D. (2001) Simultaneous analysis of 16S, 28S, COI and morphology in the Hymenoptera: Apocrita – evolutionary transitions among parasitic wasps. *Biol J Linnean Soc* **74**: 87–111.

Finston, T.L., Hebert, P.D.N. and Foottit, R.B. (1995) Genome size variation in aphids. *Insect Biochem Mol Biol* **25**: 189–196.

Gadau, J., Page, R.E., Werren, J.H. and Schmid-Hempel, P. (2000) Genome organization and social evolution in Hymenoptera. *Naturwissenschaften* **87**: 87–89.

Gokhman, V.E. (2009) *Karyotypes of Parasitic Hymenoptera*. Springer, Berlin.

Goodisman, M.A.D., Kovacs, J.L. and Hunt, B.G. (2008) Functional genetics and genomics in ants (Hymenoptera: Formicidae): the interplay of genes and social life. *Myrmecol News* **11**: 107–117.

Gregory, T.R. (2002) Genome size and developmental complexity. *Genetica* **115**: 131–146.

Gregory, T.R. (2005) Genome size evolution in animals. In *The Evolution of the Genome* (Gregory, T.R., ed.), pp. 3–87. Elsevier, San Diego, CA.

Gregory, T.R. (2009) Animal genome size database. http://www.genomesize.com.

Gregory, T.R. and Johnston, J.S. (2008) Genome size diversity in the family Drosophilidae. *Heredity* **101**: 228–238.

Gregory, T.R., Nedved, O. and Adamowicz, S.J. (2003) C-value estimates for 31 species of ladybird beetles (Coleoptera: Coccinellidae). *Hereditas* **139**: 121–127.

Grimaldi, D. and Engel, M.S. (2005) *Evolution of the Insects*. Cambridge University Press, Cambridge.

Hardie, D.C., Gregory, T.R. and Hebert, P.D.N. (2002) From pixels to picograms: a beginners' guide to genome quantification by Feulgen image analysis densitometry. *J Histochem Cytochem* **50**: 735–749.

Hölldobler, B. and Wilson, E.O. (2008) *The Superorganism*. W.W. Norton & Co., New York, NY.

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

Johnston, J.S., Ross, L.D., Beani, L., Hughes, D.P. and Kathirithamby, J. (2004) Tiny genomes and endoreduplication in Strepsiptera. *Insect Mol Biol* **13**: 581–585.

Johnston, J.S., Yoon, K.S., Strycharz, J.P., Pittendrigh, B.R. and Clark, J.M. (2007) Body lice and head lice (Anoplura: Pediculidae) have the smallest genomes of any hemimetabolous insect reported to date. *J Med Entomol* **44**: 1009–1012.

Jordan, R.A. and Brosemer, R.W. (1974) Characterization of DNA from three bee species. *J Insect Physiol* **20**: 2513–2520.

Juan, C. and Petitpierre, E. (1989) C-banding and DNA content in seven species of Tenebrionidae (Coleoptera). *Genome* **32**: 834–839.

Koshikawa, S., Miyazaki, S., Cornette, R., Matsumoto, T. and Miura, T. (2008) Genome size of termites (Insecta, Dictyoptera, Isoptera) and wood roaches (Insecta, Dictyoptera, Cryptocercidae). *Naturwissenschaften* **95**: 859–867.

Kronforst, M.R., Gilley, D.C., Strassmann, J.E. and Queller, D.C. (2008) DNA methylation is widespread across social Hymenoptera. *Curr Biol* **18**: R287–R288.

Li, J. and Heinz, K.M. (2000) Genome complexity and organization in the red imported fire ant *Solenopsis invicta* Buren. *Genet Res* **75**: 129–135.

Liolios, K., Mavrommatis, K., Tavernarakis, N. and Kyrpides, N.C. (2008) The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucl Acids Res* **36**: D475–D479.

Mariano, C.S.F., Delabie, J.H.C., Ramos, L.S., Lacau, S. and Pompolo, S.G. (2004) *Dinoponera lucida* Emery (Formicidae: Ponerinae): the highest number of chromosomes known in Hymenoptera. *Naturwissenschaften* **91**: 182–185.

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society,

Mariano, C.S.F., Pompolo, S.G., Campos Barros, L.A., Mariano-Neto, E., Campiolo, S. and Delabie, J.H.C. (2008) A biogeographical study of the threatened ant *Dinoponera lucida* Emery (Hymenoptera: Formicidae: Ponerinae) using a cytogenetic approach. *Insect Conserv Div* **1**: 161–168.

Nasonia Genome Working Group (2010) Functional and evolutionary insights from the genome of three parasitic *Nasonia* species. *Science* **327**: 343–348.

Pagano, M. and Gauvreau, K. (1993) *Principles of Biostatistics*. Duxbury Press, Belmont, CA.

Page, R.E., Gadau, J. and Beye, M. (2002) The emergence of hymenopteran genetics. *Genetics* **160**: 375–379.

Palomeque, T. and Lorite, P. (2008) Satellite DNA in insects: a review. *Heredity* **100**: 564–573.

Pennacchio, F. and Strand, M.R. (2006) Evolution of developmental strategies in parasitic Hymenoptera. *Annu Rev Entomol* **51**: 233–258.

Pultz, M.A. and Leaf, D.S. (2003) The jewel wasp *Nasonia*: querying the genome with haplo-diploid genetics. *Genesis* **35**: 185–191.

Rasch, E.M., Barr, H.J. and Rasch, R.W. (1971) The DNA content of sperm of *Drosophila melanogaster*. *Chromosoma* **33**: 1–18.

Robinson, G.E., Grozinger, C.M. and Whitfield, C.W. (2005) Sociogenomics: social life in molecular terms. *Nat Rev Genet* **6**: 257–271.

Roth, G., Blanke, J. and Wake, D.B. (1994) Cell size predicts morphological complexity in the brains of frogs and salamanders. *Proc Natl Acad Sci USA* **91**: 4796–4800.

Sherman, P.W. (1979) Insect chromosome numbers and eusociality. *Am Nat* **113**: 925–935.

Tsutsui, N.D., Suarez, A.V., Spagna, J.C. and Johnston, J.S. (2008) The evolution of genome size in ants. *BMC Evol Biol* **8**: 64.

Vinogradov, A.E. (1998) Genome size and GC-percent in vertebrates as determined by flow cytometry: the triangular relationship. *Cytometry* **31**: 100–109.

Whitfield, J.B. (1998) Phylogeny and evolution of host-parasitoid interactions in Hymenoptera. *Annu Rev Entomol* **43**: 129–151.

Wilson, E.O. (2006) How to make a social insect. *Nature* **443**: 919–920.

© 2010 The Authors
Journal compilation © 2010 The Royal Entomological Society, **19**, 337–346