Current Biology, Vol. 14, 1284–1290, July 27, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.cub.2004.07.025

Clustered Organization of Reproductive Genes in the *C. elegans* Genome

Michael A. Miller,[1,*] Asher D. Cutter,[2] Ikuko Yamamoto,[3] Samuel Ward,[4] and David Greenstein[3]

[1]Department of Cell Biology
University of Alabama at Birmingham
Birmingham, Alabama 35294

[2]Department of Ecology and Evolutionary Biology
University of Arizona
Tucson, Arizona 85721

[3]Department of Cell and Developmental Biology
Vanderbilt University School of Medicine
Nashville, Tennessee 37232

[4]Department of Molecular and Cellular Biology
University of Arizona
Tucson, Arizona 85721

Summary

Defining the forces that sculpt genome organization is fundamental for understanding the origin, persistence, and diversification of species [1, 2]. The genomic sequences of the nematodes *Caenorhabditis elegans* and *Caenorhabditis briggsae* provide an excellent opportunity to explore the dynamics of chromosome evolution [3, 4]. Extensive chromosomal rearrangement has accompanied divergence from their common ancestor, an event occurring roughly 100 million years ago (Mya) [4]; yet, morphologically, these species are nearly indistinguishable and both reproduce primarily by self-fertilization. Here, we show that genes expressed during spermatogenesis (sperm genes) are nonrandomly distributed across the *C. elegans* genome into three large clusters located on two autosomes. In addition to sperm genes, these chromosomal regions are enriched for genes involved in the hermaphrodite sperm/oocyte switch and in the reception of sperm signals that control fertilization. Most loci are present in single copy, suggesting that cluster formation is largely due to gene aggregation and not to tandem duplication. Comparative mapping indicates that the *C. briggsae* genome differs dramatically from the *C. elegans* genome in clustering. Because clustered genes have a direct role in reproduction and thus fitness, their aggregated pattern might have been shaped by natural selection, perhaps as hermaphroditism evolved.

Results and Discussion

Evolutionary theory predicts that self-fertilization will result in distinctive genomic patterns [5, 6]. Several of these patterns have been identified in *C. elegans*, including reductions in levels of genetic variation [7], a shifted association between genetic variation and recombination rate [8], and high rates of chromosomal rearrangement [4, 9]. However, theory does not predict explicitly whether self-fertilization might also generate patterns of functionally related loci throughout the genome, even for genes involved in reproduction. Because neutral processes are unlikely to create aggregations of functionally related genes, the presence of reproductive gene clusters could reflect a response to selection.

Sperm Genes Are Clustered

We used a DNA microarray dataset comprising 92% of genes in the genome [10] to identify 887 loci whose mRNA transcripts are enriched during spermatogenesis (sperm genes) and 523 loci whose transcripts are enriched during oogenesis (oocyte genes). Using the frequencies of these loci within nonoverlapping chromosomal windows of 0.5 recombination units (cM), we identified three large, nonrandom clusters of sperm genes ($P < 10^{-6}$) but no significant clusters of oocyte genes ($P > 10^{-3}$). One sperm gene cluster (SGC, see bars in Figure 1) containing 48 sperm genes is located on chromosome II and two containing 49 (IV-A) and 86 (IV-B) sperm genes, respectively, are located on chromosome IV (Figure 1). Qualitatively identical results were obtained for a range of window sizes in both map (cM) and physical (bp) units (data not shown).

Among the genes found within the SGCs are 28 loosely clustered, nearly identical loci that encode the major sperm proteins (MSPs, Figure 1) [11]. The MSPs comprise a highly abundant, sperm-specific protein family with dual functions—these proteins act as extracellular signals that control the rate of fertilization [12, 13] and as the cytoskeleton in amoeboid motility of nematode sperm [14]. Although *msp* loci are a significant component of the SGCs, strong clustering is still observed on chromosome IV after they are removed from the dataset ($P < 10^{-10}$ for IV-A and IV-B; clustering weakens on chromosome II to $P < 10^{-3}$).

Reproduction-Related Genes Are Aggregated near the Sperm Gene Clusters

The transcription profiles of genes within SGC-containing regions were further evaluated by using additional DNA microarray experiments [15–17]. Genes that showed enriched expression in various tissues were identified, and the number of these genes observed in each SGC-containing region (bars in Figure 1) was compared to the number expected by chance (Figure 2). In addition to sperm genes, these regions contain a disproportionate number of genes whose transcripts are expressed in the hermaphrodite germ line [15] (Figure 2). Germline genes expressed during the L4 stage, the time of spermatogenesis, are particularly abundant in all SGC-containing regions ($P < 0.001$). Germline genes expressed in adults during oogenesis are strongly overrepresented in SGC region IV-B ($P < 0.001$) and modestly overrepresented in SGC regions II and IV-A ($P < 0.05$, Figure 2). Oocyte genes are modestly overrepresented in regions II and IV-B ($P < 0.05$), but not in region

*Correspondence: mamiller@uab.edu

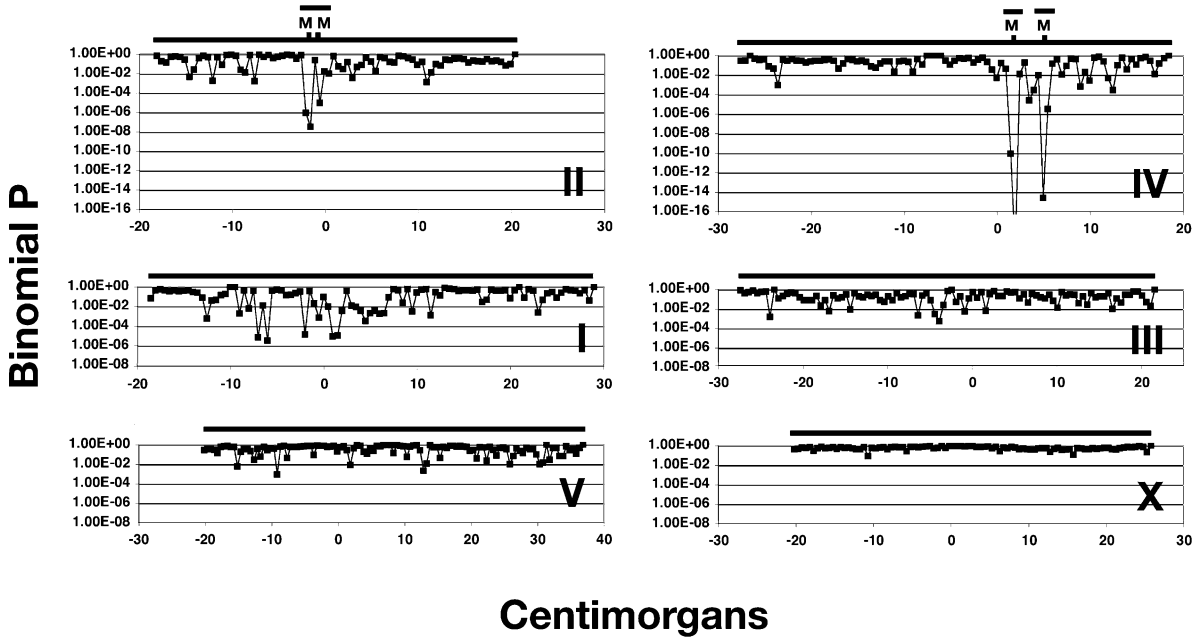

Figure 1. Sperm-Enriched Genes Are Nonrandomly Distributed across the Genome

The probability that the observed number of sperm-enriched genes lies in a given 0.5 cM interval is plotted as a function of recombination distance (centimorgans) for each chromosome. Similar results were observed with 250 kb chromosome intervals. An M marks the positions of *msp* genes. Black bars represent chromosomal windows analyzed in Figure 2.

IV-A (Figure 2). In contrast to genes expressed in the germ line, the clusters contain fewer loci associated with muscle [16] or dauer development [17] than expected by chance in all but one case (Figure 2). From these gene transcription profiles, the SGC-containing regions contain an overrepresentation of genes that are expressed in the germ line, to the exclusion of nonreproductive gene classes.

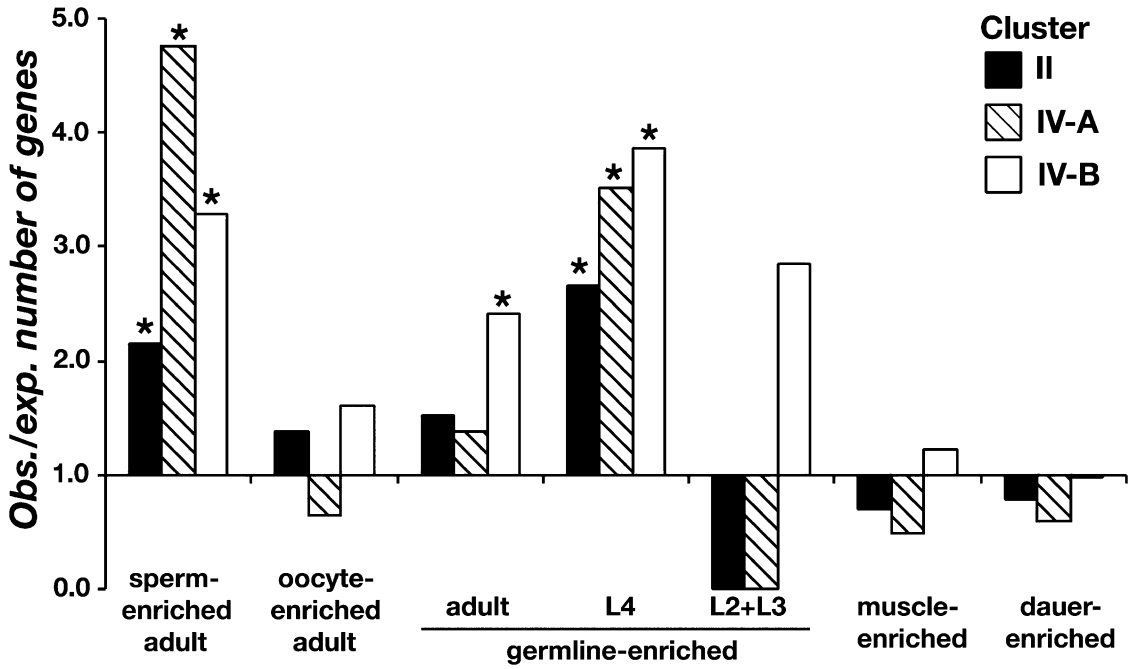

Figure 2. Observed versus Expected Distribution of Genes within SGC Regions that Have Common mRNA Expression Patterns

DNA microarrays comprising either 63% or 92% of total loci in the genome were used to identify genes with common transcriptional profiles [10, 15–17]. The ratio of the number of observed to the number expected by chance is plotted for each microarray dataset. Chromosomal windows containing clusters II (–3.2 to –0.5 cM), IV-A (+1.0 to +2.3 cM), and IV-B (+3.9 to +5.0) are plotted separately. *$P < 0.001$.

Table 1. The Positions of Sperm Signaling and Sperm/Oocyte Switch Genes Are Biased toward Regions Containing the SGCs

Class Description (n > 10 Total Loci)	Number of Genes <3 cM from SGCs	>3 cM from SGCs	*P*
Reproductive processes			
Sperm-enriched (mRNA) loci	307	580	3.3×10^{-31}
Oocyte-enriched (mRNA) loci	98	425	0.66
Oocyte maturation and ovulation*	10	5	4.0×10^{-6}
Oocyte MAPK regulation*	8	5	1.3×10^{-4}
Hermaphrodite sperm/oocyte switch*	10	7	4.2×10^{-5}
Hermaphrodite spermatogenesis onset*	3	9	0.63
Somatic sex determination	3	10	0.74
Distal tip cell signaling*	3	15	0.76
Vulval development	2	23	0.15
Nonreproductive processes			
Ubiquitous cell cycle regulators	2	15	0.42
Ubiquitous cytoskeletal components	3	23	0.31
Embryonic polarity and patterning*	3	17	0.61
Chemotaxis and thermotaxis	3	28	0.17
mab (male abnormal)	4	17	0.96
daf (dauer formation abnormal)	4	17	0.96
ced (cell death abnormal)	1	18	0.12
unc (uncoordinated)	8	46	0.38
lin (lineage abnormal)	8	36	0.83
mec (mechanosensory abnormal)	4	12	0.58
eat (eating abnormal)	3	15	0.76
emb (embryogenesis abnormal)	8	27	0.62
mig (cell migration abnormal)	1	12	0.28
egl (egg-laying defective)	11	36	0.50
sup (suppressor)	3	28	0.17

See Table S1 for references. *P*, probability that gene distribution is random; *, functionally related classes that primarily consist of genes expressed in the germ line.

Genes with increased expression in the adult germ line and during oogenesis include many genes that function in embryos (maternal effect genes) [10, 15], which could lead to underestimating the clustering of genes that function in the oocytes themselves. Furthermore, not all genes involved in reproduction exhibit significantly increased germline expression by microarray analysis [10, 15]. Therefore, we used functional data—based on genetic or biochemical analyses—to identify genes that fall into common reproductive and nonreproductive classes. The extent to which genes in each class (n > 10 characterized loci per class) are overrepresented in SGC-containing regions was determined and statistically analyzed (Table 1). A larger chromosomal bin size (within 3.0 cM from all SGCs) was used in this analysis to compensate for the fewer total loci with functional data relative to microarray expression data. Among the loci in the vicinity of the SGCs, we detected significantly more genes involved in the MSP sperm signaling mechanism than expected, given their genomic abundances. The MSPs bind to receptors on oocytes and somatic sheath cells to stimulate oocyte meiotic maturation, mitogen-activated protein kinase (MAPK) activation, gonadal sheath cell contraction, and spermathecal valve dilation [12, 13]; these events directly precede fertilization. Clustered sperm signaling genes include the *msps*, the MSP/ephrin receptor *vab-1* [13], the ephrin *efn-2* [13], and other downstream regulators that function in oocytes and somatic gonadal cells (Tables 2 and S1).

In addition to sperm signaling genes, genes involved in the hermaphrodite sperm/oocyte switch are also overrepresented in SGC-containing regions (Table 1). This switch from sperm to oocyte production occurs during hermaphrodite larval development and depends on the activity of *fem-3*, which is regulated posttranscriptionally [18]. In contrast to genes involved in sperm signaling and the sperm/oocyte switch, genes in other reproductive classes are not overrepresented here, nor are genes in nonreproductive classes (Table 1). Out of six functionally related classes consisting primarily of genes expressed in the germ line, only three exhibit an SGC distribution bias (Table 1), indicating that germline expression and clustering do not always correlate. Collectively, microarray and functional data indicate that genes expressed during spermatogenesis are clustered together with genes involved in the hermaphrodite sperm/oocyte switch and in the reception of MSP signals that control oocyte meiotic maturation and ovulation.

Empirical Verification of Reproductive Function for Uncharacterized Genes

Given that known genetic regulators of the sperm signaling mechanism are overrepresented in SGC regions, additional regulators are predicted to be among the uncharacterized loci in the same regions. To test this prediction, we analyzed the mutant phenotypes of 12 SGC-linked loci that had not been implicated previously in the sperm signaling mechanism. These loci were selected

Table 2. Clustered Signal Transduction Genes Function in Sperm Signaling Pathways

Gene and Description	exp	om	mk	vd	sc	*C. elegans* Chromosome (cM)	*C. briggsae* Supercontig Number
II major sperm proteins	sp	+	+	+	+	II −2.7 to −1.0	0058, 4206, 2234, 4470, 0090, 0010, 2887
vab-1 Eph receptor kinase	o, sh	−	−	−	+	II −3.09	3052
nmr-1 glutamate receptor*	o, sh	−	−	−	−	II −1.43	0058
clr-1 receptor phosphatase*	oe	−		−		II −1.37	0058
ptp-2 tyrosine phosphatase*	oe, glf	+	+	+		II −0.91	0058
ptc-1 patched receptor *	o	−		−		II +0.62	0058
let-23 EGF receptor	stf			+		II +1.05	0022
mel-11 phosphatase subunit	st			+		II +1.16	0022
IV-A major sperm proteins	sp	+	+	+	+	IV +1.5 to +1.8	See above
soc-2 leucine-rich repeats*	ne		+			IV +1.60	4152
rme-2 LDL receptor*	o			+		IV +1.89	0071
lin-45 raf kinase	glf		+			IV +3.20	0110
lip-1 MAPK phosphatase	gl		−			IV +3.23	0143
itr-1 inositol receptor	o, sh, st	−	−	+	+	IV +3.47	0143
IV-B major sperm proteins	sp	+	+	+	+	IV +4.4 to +4.5	See above
oma-1 zinc finger protein	o	+	+	+		IV +3.94	0058, 4260
cav-1 caveolin*	o	−	−	−		IV +4.43	3857
tsp-12 tetraspanin*	oe	−	−	−		IV +4.53	0063
unc-43 CaMKII kinase*	o, sh	+	+	+	−	IV +4.58	0081
efn-2 ephrin	glf	−	−	−		IV +4.68	0081
lin-3 EGF ligand	oe			+		IV +4.82	0143
let-60 ras	glf		+			IV +5.17	0143
inx-8 & *9* gap junction*	sh	−		−	−	IV +5.17	0143
gex-3 GTPase effector*	oe	−		−	+	IV +5.85	0063
ark-1 tyrosine kinase	stf			−		IV +6.41	3052

Functions, expression data or inferred site of action (exp), and genomic positions of clustered sperm signaling genes are shown. *, one of twelve signaling genes selected by position (the receptor *cam-1* is not shown because a reproductive function was not detected). Abbreviations: sp, sperm expressed; o, oocyte expressed; sh, sheath expressed; st, spermatheca expressed; gl, germline expressed; glf, germline function; stf, spermatheca function; oe, oocyte-enriched transcript; ne, non-enriched transcript; om, oocyte maturation; mk, MAPK regulation; vd, spermathecal valve dilation; sc, somatic sheath cell contraction; +, positive regulator; −, negative regulator; cM, centimorgans. Underlined *C. briggsae* supercontigs are described in the text. See Table S1 for references and data.

arbitrarily from those genes in the region that encode predicted signaling proteins and have existing mutant alleles. For 11 of the 12 genes, phenotypic analysis of the mutants is consistent with a role in MSP-responsive signaling pathways (Table 2). Their wild-type gene products are required for regulating oocyte maturation, gonadal sheath contraction, spermathecal valve dilation, or oocyte MAPK activation (Tables 2 and S1). *cav-1, ptc-1*, and *rme-2* encode proteins that have been localized to oocytes [19-21]; *inx-8* and *inx-9* (both genes are treated as one in this analysis because they are paralogs) encode gap junctional proteins that have been localized to sheath cells surrounding maturing oocytes [22]; *clr-1, gex-3, ptp-2*, and *tsp-12* have oocyte-enriched mRNA transcripts [10, 15]; and *nmr-1* and *unc-43* function in oocytes and sheath cells (M.A.M., unpublished data). A summary of function and expression data for these 11 genes and additional clustered signaling genes from Table 1 is shown in Table 2. These phenotypic studies demonstrate that the SGC-containing regions harbor an abundance of genes encoding both negative and positive regulators of MSP sperm signaling pathways.

Genes in *C. elegans* SGCs IV-A and IV-B Are Dispersed in *C. briggsae*

We used comparative mapping to determine whether clustered *C. elegans* reproductive genes are clustered in the *C. briggsae* genome [4]. The distribution of SGC loci on *C. elegans* chromosome II is largely conserved in both species (Figure 3A and Table S1), although the *msp* genes are in different relative positions. Of 70 *C. elegans* loci that were sampled from SGC region II, 81% had *C. briggsae* homologs that are found in the syntenic supercontig 0058, which corresponds to the largest syntenic region [4].

In contrast to the chromosome II cluster, the positions of *C. elegans* genes in SGC regions IV-A and IV-B are not conserved in *C. briggsae*. The *C. briggsae* homologs of these genes lie in over 50 different supercontigs (Figure 3A and Table S1). Of 78 genes sampled from *C. elegans* region IV-A, most of which have gamete-enriched transcription profiles [10, 15], only 37% of their *C. briggsae* homologs are found in the syntenic supercontig 4152. Two conserved segments from *C. elegans* regions surrounding cluster IV-B are found on the same *C. briggsae* supercontig (0143), but several hundred SGC-linked genes located between these segments are located elsewhere in *C. briggsae* (Figures 3A and 3B, and Table S1). Furthermore, *C. briggsae msp* genes are not found in supercontigs that are syntenic to the *C. elegans* chromosomal regions that contain clusters IV-A or IV-B (Figure 3B and Table 2). These observations indicate that the *C. briggsae* genome does not contain clusters corresponding to *C. elegans* clusters IV-A and IV-B, although both species contain similar distributions of reproductive genes on chromosome II. Mapping *C. briggsae* homologs of *C. elegans* reproductive genes suggests that the *C. briggsae* genome contains smaller clusters located in chromosomal regions distinct from *C. elegans*

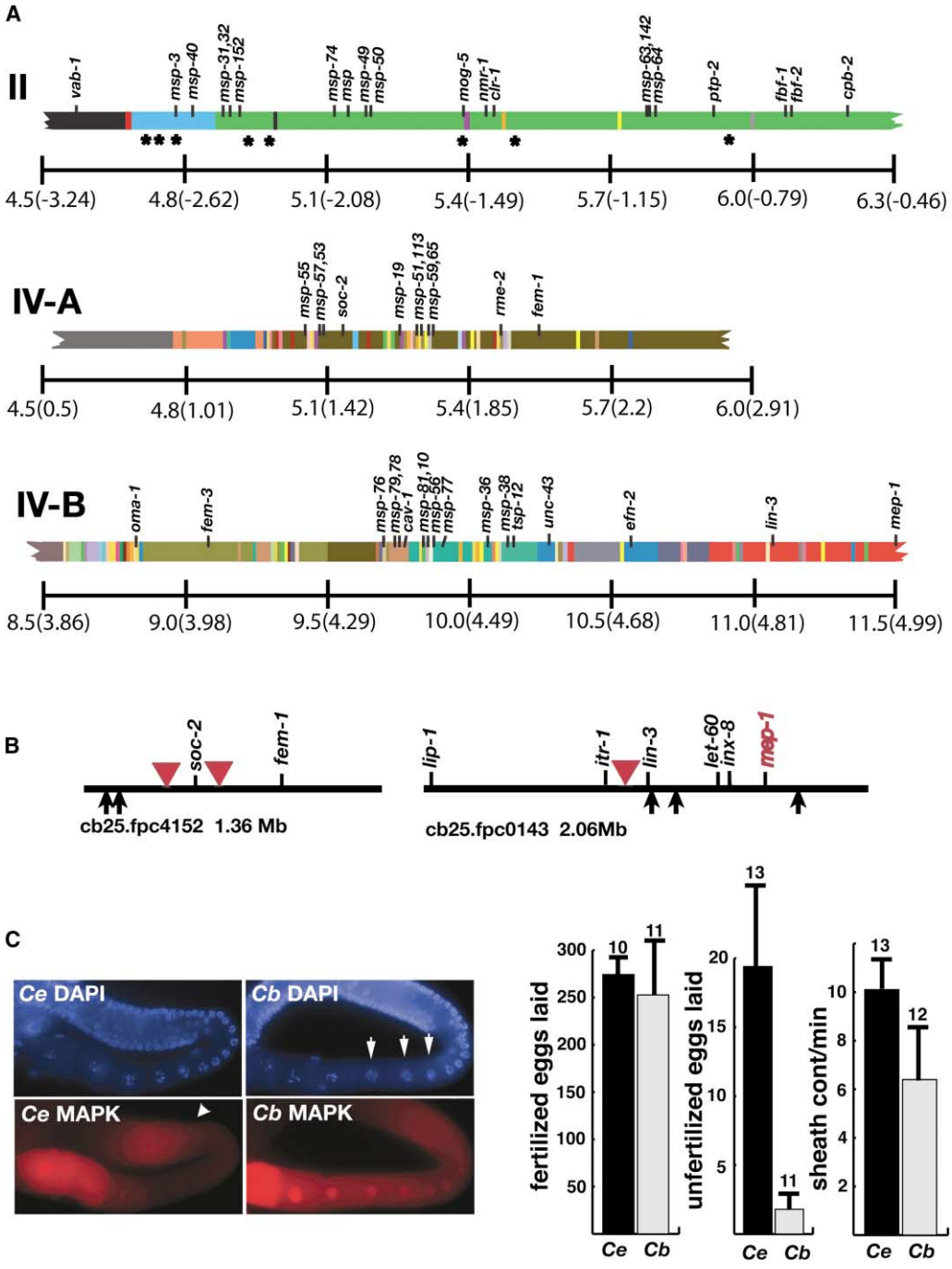

Figure 3. Differences in Reproductive Gene Clustering in *C. elegans* and *C. briggsae* Correlate with Differences in Regulation

(A) Genes in clustered regions of the *C. elegans* genome are distributed differently in *C. briggsae*. Colors represent distinct *C. briggsae* supercontigs mapped onto *C. elegans* genomic regions by using the positions of homologous genes. If multiple *C. briggsae* homologs of a *C. elegans* gene exist on different supercontigs, then sequence breaks were minimized, if possible. *msp* genes were not included because sequence analyses suggest that they are paralogous (see Figure S2). *C. elegans* chromosome positions are shown below in Mb (cM). The positions of selected *C. elegans* reproductive genes are shown above each chromosome region. Asterisks indicate the relative positions of *C. briggsae msp* genes.

(B) *C. briggsae* supercontigs that exhibit synteny with clustered chromosome IV regions lack most reproductive genes, including the *msps*. Red arrowheads mark the relative positions of *C. elegans msp* genes in these *C. briggsae* supercontigs. The positions of black genes are conserved in both species, whereas red genes have different locations. Black arrows represent *C. elegans* sperm-enriched genes whose positions are conserved in the two genomes.

(C) DAPI staining indicates that *C. briggsae* (*Cb*) oocytes progress through pachytene and diplotene meiotic stages more gradually (arrows) than *C. elegans* (*Ce*). MAPK activation, assayed by using the MAPK-YT antibody [12], is downregulated in *C. elegans* as oocytes exit pachytene (arrowhead). By contrast, this does not occur in *C. briggsae*. Other differences include brood size variability, reproductive efficiency, and the rate of sheath cell contraction.

(Figure S1). A caveat here is that reproductive roles for *C. briggsae* genes are inferred from data on their *C. elegans* homologs. The differences in reproductive gene distribution between *C. elegans* and *C. briggsae* have been caused by numerous intra- and interchromosomal rearrangements [9].

Evolutionary Implications of Clustered Reproductive Genes

The difference between *C. elegans* and *C. briggsae* in gene clustering is consistent with two alternative explanations for the timing of cluster formation. Either *C. elegans* SGCs IV-A and IV-B originated in a lineage-specific fashion, or they originated in a common ancestor and were subsequently lost in *C. briggsae*. Two arguments support the former explanation. First, phylogenetic analysis suggests that the *msp* genes, which are an integral part of the SGCs and *C. elegans* reproduction, have expanded independently and invaded new chromosome positions in the *C. elegans* genome (Figure S2). The dynamic nature of the *msp* gene family is evidenced by their varying abundance among diverse nematodes [23], the large number of *msp* pseudogenes present in the genomes of both *C. elegans* and *C. briggsae*, and the observation that only one out of 28 *msp* genes in *C. elegans* shares the same nearest gene neighbors as in *C. briggsae*. Moreover, the *msp* genes of *C. briggsae* and more basal species contain an intron, whereas all expressed *C. elegans msps* are intronless (two pseudogenes on chromosomes II and V contain an intron). This dynamism indicates that it is unlikely that their present distribution in *C. elegans* is indicative of their distribution in a distant ancestor. Second, given the high rate of chromosomal rearrangement in *Caenorhabditis* [4, 9] and the reasonable assumption that rearrangements occur with equal frequency in the lineages leading to *C. elegans* and *C. briggsae*, we expect that most ancestral blocks of loci will be scrambled by now in both lineages in the absence of selection for their maintenance. High rearrangement rates limit the utility of distantly related outgroups such as *Brugia malayi*, which last shared a common ancestor with *C. elegans* roughly 300 Mya and does not have clusters corresponding to the *C. elegans* SGCs (Supplemental Experimental Procedures). Therefore, genomic sequence of basal Caenorhabditid species is necessary to confidently date the origins of the *C. elegans* reproductive clusters.

The distinct genomic patterns of *C. elegans* and *C. briggsae* are in contrast with the strong morphological resemblance of their reproductive tissues. Despite superficial similarities, however, important differences in sperm function and the sperm/oocyte switch exist [24, 25]. Comparing the *C. briggsae* sperm signaling mechanism to *C. elegans* reveals that *C. briggsae* oocytes progress through pachytene and diplotene meiotic stages more gradually than *C. elegans*, without downregulating MAPK activity (Figure 3C). *C. briggsae* hermaphrodites also have more variable brood sizes, lower sheath-contraction rates, and lay fewer unfertilized oocytes after self-derived sperm are depleted (Figure 3C). These observations suggest that *C. elegans* and *C. briggsae* hermaphrodite gonads differ appreciably in genetic regulation.

The evolution of hermaphroditism (hermaphrodite/male) from gonochorism (female/male) represents a major shift in reproductive mode within the *Caenorhabditis* clade [4]. Based on the current *Caenorhabditis* phylogeny, the most parsimonious explanation for breeding system evolution in this group is that hermaphroditism arose independently in *C. elegans* and *C. briggsae* from gonochoristic ancestors [4]. Therefore, might clustering correlate with the acquisition of the hermaphroditic reproductive mode? As nematode females acquired the ability to produce their own sperm through changes in genetic regulation, enhanced sperm gene transcription could lead to reduction in the time required for sperm production. Fitness in self-fertilizing *Caenorhabditis*, defined as the intrinsic rate of population growth, depends on the optimization of the fertilization delay caused by sperm production [26, 27]. If hermaphroditism is a driving force favoring reproductive gene aggregation, we predict that the genomes of both species will differ from their gonochoristic relatives by containing clusters—though they need not be syntenic or contain the same gene complements. Large-scale functional assays for reproductive genes in *C. elegans*' congeners *C. briggsae* and the gonochoristic species *C. remanei, C.* sp. CB5161, and *C. japonica* (for which sequencing efforts have been approved) will help elucidate the degree to which orthologs are involved in reproduction, whether nematode reproductive genes generally display an aggregated distribution, and whether hermaphroditism is a common theme among nematode genomes with reproductive gene clusters.

In addition to the issues of lineage specificity and whether hermaphroditism might have facilitated clustering, we must also consider other mechanisms to explain cluster origin and maintenance. A neutral explanation for the pattern we observe seems unlikely. Although gene duplication and gene conversion may be important processes in the evolution of the *msp* family [11], these processes are not applicable to most other reproductive loci, which are found in single copy. One hypothesis for clustering is that the accretion of reproductive genes reflects selection for linked blocks of adaptive gene combinations, although it is not clear whether population genetic forces can have major effects on gene order [2], particularly in selfing species. A high rate of selfing will allow fixation of chromosomal rearrangements with high probability if they confer a fitness benefit when homozygous, despite a heterozygous cost [5]. However, selfing generally complicates predictions for how selection might act to reduce recombination between loci because the effective recombination rate is low in populations with little heterozygosity [28, 29]. Another hypothesis for clustering is that it reflects selection favoring transcriptional coregulation or coordination [2]. The widespread incidence of operons throughout the nematode genome [30] and local patterns of correlated transcription [16] indicate that this idea is plausible in principle. Although transcriptional interference due to chromosome silencing has been invoked to explain the lack of sperm genes on the X chromosome [15], it is not obvious whether such a phenomenon could work across $>$1 Mb regions of autosomal DNA with genes that exhibit highly variable levels of expression. Clearly, additional

work is required to delineate the relative importance of these scenarios for the origin and maintenance of the *C. elegans* reproductive gene clusters. The nonrandom patterns of sperm gene distribution observed in *Drosophila* [31] and mouse [32] genomes suggest that reproductive gene clustering may not be restricted to self-fertilizing nematodes.

Supplemental Data

Supplemental Data including additional data, detailed Experimental Procedures, references, and two figures are available at http://www.current-biology.com/cgi/content/full/1284/DC1/.

Acknowledgments

We thank the Sanger Institute and the Genome Sequencing Center, Washington University, St. Louis, for providing the *C. elegans* and *C. briggsae* sequences, and Valerie Reinke for sharing unpublished microarray data. We also thank Robert E. Steele, Ron Ellis, and three anonymous reviewers for their helpful comments on the manuscript. Some strains were provided by the *Caenorhabditis* Genetics Center and International *C. elegans* Gene Knockout Consortium. This work was supported by grants from the National Institutes of Health to D.G. and S.W., an American Cancer Society postdoctoral fellowship and UAB Comprehensive Cancer Center Junior Faculty Development Grant to M.A.M., and Department of Defense National Defense Science and Engineering Graduate fellowship and University of Arizona National Science Foundation Integrative Graduate Education and Research Traineeship Genomics Initiative fellowship to A.D.C.

Received: January 11, 2004
Revised: May 30, 2004
Accepted: June 2, 2004
Published: July 27, 2004

References

1. Eichler, E.E., and Sankoff, D. (2003). Structural dynamics of eukaryotic chromosome evolution. Science *301*, 793–797.
2. Hurst, L.D., Pal, C., and Lercher, M.J. (2004). The evolutionary dynamics of eukaryotic gene order. Nat. Rev. Genet. *5*, 299–310.
3. The *C. elegans* Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. Science *282*, 2012–2018.
4. Stein, L.D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M.R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., et al. (2003). The genome sequence of *Caenorhabditis briggsae*: A platform for comparative genomics. PLoS Biol. 1(2), e45 10.1371/journal.pbio.0000045.
5. Charlesworth, B. (1992). Evolutionary rates in partially self-fertilizing species. Am. Nat. *140*, 126–148.
6. Charlesworth, D., and Wright, S.I. (2001). Breeding systems and genome evolution. Curr. Opin. Genet. Dev. *11*, 685–690.
7. Graustein, A., Gaspar, J.M., Walters, J.R., and Palopoli, M.F. (2002). Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. Genetics *161*, 99–107.
8. Cutter, A.D., and Payseur, B.A. (2003). Selection at linked sites in the partial selfer *Caenorhabditis elegans*. Mol. Biol. Evol. *20*, 665–673.
9. Coghlan, A., and Wolfe, K.H. (2002). Fourfold faster rate of genome rearrangement in nematodes than in Drosophila. Genome Res. *16*, 857–867.
10. Reinke, V., Gil, I.S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. Development *131*, 311–323.
11. Ward, S., Burke, D.J., Sulston, J.E., Coulson, A.R., Albertson, D.G., Ammons, D., Klass, M., and Hogan, E. (1988). Genomic organization of major sperm protein genes and pseudogenes in the nematode *Caenorhabditis elegans*. J. Mol. Biol. *199*, 1–13.
12. Miller, M.A., Nguyen, V.Q., Lee, M.H., Kosinski, M., Schedl, T., Caprioli, R.M., and Greenstein, D. (2001). A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation. Science *291*, 2144–2147.
13. Miller, M.A., Ruest, P., Kosinski, M., Hanks, S., and Greenstein, D. (2003). An Eph receptor sperm-sensing control mechanism for oocyte meiotic maturation in *Caenorhabditis elegans*. Genes Dev. *17*, 187–200.
14. Bottino, D., Mogilner, A., Roberts, T., Stewart, M., and Oster, G. (2002). How nematode sperm crawl. J. Cell Sci. *115*, 367–384.
15. Reinke, V., Smith, H.E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S.J., Davis, E.B., Scherer, S., Ward, S., et al. (2000). A global profile of germline gene expression in *C. elegans*. Mol. Cell *6*, 605–616.
16. Roy, P.J., Stuart, J.M., Lund, J., and Kim, S.K. (2002). Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. Nature *418*, 975–999.
17. Wang, J., and Kim, S.K. (2003). Global analysis of dauer gene expression in *Caenorhabditis elegans*. Development *130*, 1621–1634.
18. Kuwabara, P.E., and Perry, M.D. (2001). It ain't over till its ova: germline sex determination in *C. elegans*. Bioessays *23*, 596–604.
19. Scheel, J., Srinivasan, J., Honnert, U., Henske, A., and Kurzchalia, T.V. (1999). Involvement of caveolin-1 in meiotic cell-cycle progression in *Caenorhabditis elegans*. Nat. Cell Biol. *1*, 127–129.
20. Kuwabara, P.E., Lee, M.H., Schedl, T., and Jefferis, G.S. (2000). A *C. elegans* patched gene, ptc-1, functions in germ-line cytokinesis. Genes Dev. *14*, 1933–1944.
21. Grant, B., and Hirsh, D. (1999). Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. Mol. Biol. Cell *10*, 4311–4326.
22. Starich, T., Sheehan, M., Jadrich, J., and Shaw, J. (2001). Innexins in *C. elegans*. Cell Commun. Adhes. *8*, 311–314.
23. Scott, A.L., Dinman, J., Sussman, D.J., and Ward, S. (1989). Major sperm protein and actin genes in free-living and parasitic nematodes. Parasitology *98*, 471–478.
24. LaMunyon, C.W., and Ward, S. (1997). Increased competitiveness of nematode sperm bearing the male X chromosome. Proc. Natl. Acad. Sci. USA *94*, 185–189.
25. Haag, E.S., Wang, S., and Kimble, J. (2002). Rapid coevolution of the nematode sex-determining genes fem-3 and tra-2. Curr. Biol. *12*, 2035–2041.
26. Barker, D.M. (1992). Evolution of sperm shortage in a selfing hermaphrodite. Evolution *46*, 1951–1955.
27. Cutter, A.D. (2004). Sperm-limited fecundity in nematodes: how many sperm are enough? Evolution Int. J. Org. Evolution *58*, 651–655.
28. Charlesworth, D., and Charlesworth, B. (1979). The evolutionary genetics of sexual systems in flowering plants. Proc. R. Soc. Lond. B. Biol. Sci. *205*, 513–530.
29. Nordborg, M. (1997). Structured coalescent processes on different time scales. Genetics *146*, 1501–1514.
30. Blumenthal, T., and Gleason, K.S. (2003). *Caenorhabditis elegans* operons: form and function. Nat. Rev. Genet. *4*, 112–120.
31. Boutanaev, A.M., Kalmykova, A., Shevelyov, Y., and Nurminsky, D. (2002). Large clusters of co-expressed genes in the Drosophila genome. Nature *420*, 666–669.
32. Wang, P.J., McCarrey, J.R., Yang, F., and Page, D.C. (2001). An abundance of X-linked genes expressed in spermatogonia. Nat. Genet. *27*, 422–426.