Identifying the genetic loci responsible for development of multivulval phenotypes in model organism *C. briggsae*

Molly E. Gallagher

Abstract

Thanks to recent advances in sequencing technology, researchers are rapidly accumulating vast numbers of genomic sequences for countless plant and animal species. However in many cases, little is known regarding the functional significance of any observed sequence differences. The goal of this work is to understand how two sister species of nematode worms have evolved genomic functional and regulatory differences, despite the high degree of similarity between their anatomy and life history traits. *C. elegans* is a model organism widely used in the field of molecular genetics. In *C. elegans*, there are four genes known to influence the morphology of the egg-laying structure (vulva) such that single loss-of-function mutations can result in the formation of a multivulval phenotype. The genome of *C. briggsae,* a sister species to *C. elegans*, has been less thoroughly studied.

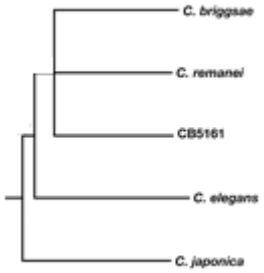


Figure 1

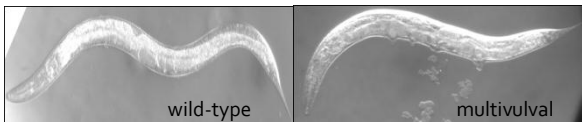


Figure 2

In this study, eight mutant strains of *C. briggsae* possessing the multivulval phenotype have been isolated from a mutagenesis screen. Each strain has been tested against the others using complementation, in order to determine whether the multivulval trait results from mutations within the same gene or represents distinct genes. Current results suggest that mutations in at least seven unique genes have been isolated. Thus the number of genes that can easily mutate to the multivulval phenotype in *C. briggsae* is larger than that found in *C. elegans*. Currently, each mutation is in the process of being 'mapped' on the genome, meaning its approximate location on one of the six chromosomes is being determined using molecular methods. After the mapping process is completed, each mutation can be sequenced for further analysis. This will confirm which of the *C. briggsae* strains bear mutations in one of the four genes identified in *C. elegans*, and which represent different loci previously unknown to influence *C. briggsae* vulval morphology.

Introduction

C. elegans and *C. briggsae* are androdioecious species, meaning they have a rare reproductive system characterized by populations both of males and of hermaphrodites. The hermaphroditic animals breed with the males but are also capable of self-crossing, and will do so if kept in isolation. The worms have five somatic chromosomes. Males have a single X chromosome, whereas hermaphrodites have two X chromosomes. The hermaphrodites possess an egg-laying structure called the vulva. The vulva is formed from six vulval precursor cells (VPCs) located on the ventral side of the animal, three of which will undergo cell divisions in the wild-type animals. The adjacent anchor cell (AC) helps to direct organogenesis by producing signaling molecules such as epidermal growth factor (EGF). The EGF molecule is a ligand that attaches to receptors on the precursor cells triggering a cascade of reactions involved in cell division. Multivulval phenotypes may result from genetic disruption of the EGF gene resulting in overexpression, or from interference with any element of the downstream molecular pathway. Multivulval animals are easily distinguishable from their wild-type counterparts under a compound microscope. Only one of the three or four egg-laying structures will develop normally, forming a tube of cells connected internally to the egg-producing gonad. The others will appear as obvious protrusions on the animal's ventral side.

Figure 3 shows the cell fates of precursor cells in wild-type animals as compared to those of vulvaless and multivulval individuals. The solid-colored cells will not undergo cell division, while shaded cells will continue to divide.

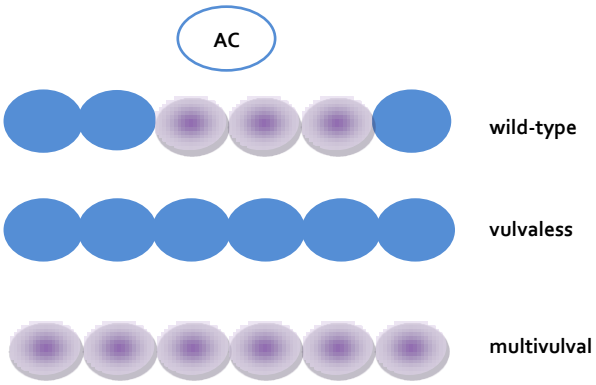


Figure 3

The proximal cause of multivulval phenotype is excess inappropriate cell division. This pattern relates closely to the development of cancerous growths, which are also the result of uncontrolled cell division. By studying how genetic changes affect the molecular pathways involved in the organogenesis of the egg-laying system, we inform our understanding of the mechanisms of malignant tumor formation.

The underlying genetic basis for the development of the multivulval phenotype in *C. elegans* is well understood, and the molecular pathways and networks that mediate vulval organogenesis have been thoroughly studied. Though the two species are morphologically and developmentally similar, the genetic regulations involved in *C. briggsae* multivulval mutants are unknown. Detailed comparison of the two can help us understand the evolution of complex regulatory networks.

Methods

In 2010 Steven Billups conducted a mutagenesis screen using a laboratory strain of *C. briggsae* designated AF16. Animals were exposed to the mutagen ethane methyl sulfonate (EMS) (Sulston and Hodgkin). EMS induces single point-mutations in DNA, usually in the form of a transition. A transition mutation is the exchange of one purine for another or one pyrimidine for another; adenine←→ guanine or cytosine←→ thymine. Following exposure to EMS, single hermaphrodites are isolated and allowed to-self cross for two generations. In the F2 generation, ¼ of the offspring will be homozygous for any recessive genetic mutations induced by EMS. Since the mutations are phenotypically obvious, mutant animals can be identified and selected using a dissection microscope. From the screen it was possible to isolate a total of eight strains of *C. briggsae* animals characterized by mutant multivulval phenotypes.

Complementation Tests

After recovering the eight mutant strains of *C. briggsae* characterized by a multivulval phenotype, the first step was to determine whether all eight mutations were unique. In *C. elegans*, a single loss-of-function mutation is known to be sufficient to produce the multivulval phenotype in only four genes, *pry-1, lin-31, lin-13,* and *lin-1*(Ferguson et al 1987; Jacobs et al). Other genes in *C. elegans* produce the multivulval phenotype only on conjunction with one another; these are known as syn-muv genes. These syn-muv genes function to restrict production of EGF, which is overexpressed in their absence (Andersen). There may be other relevant genes involved in *C. briggsae*, but we expected that at least some of the eight mutations occurred within homologues of the abovementioned four genes, and that less than eight total genes were involved.

Ideally, a cross could be done between a male homozygous for one multivulval mutation and a hermaphrodite homozygous for another. Unfortunately, the male multivulval animals are unlikely to mate; reproduction in these mutants occurs almost exclusively via self-crossing by the hermaphrodites. Instead a two-step process is required: first, a wild-type male is crossed with a hermaphrodite that is homozygous for a multivulval mutation. The heterozygous males that result from the cross are then crossed with a female that is homozygous for a different multivulval mutation, and also for a “dumpy” mutation that results in a distinct slug-like phenotype. This recessive “dumpy” dpy- mutation is used to identify any offspring that are the

result of hermaphroditic self-crossing. Offspring fathered by the male will have one wild-type allele dpy+ and will not exhibit the dumpy phenotype.

Figure 4

Of these offspring from the second cross, 50% will be heterozygous for multivulval mutations, meaning they have one wild-type allele from the male parent (muv+). These animals will have the wild-type phenotype. The other 50% will have two different multivulval mutations (muv1- and muv2-). If these occur within two different genes, the animal will have a wild-type phenotype. But if the mutations are in the same gene, the animal is effectively homozygous for the mutation, and will display the multivulval phenotype. This is what we mean by a failure to complement.

Mapping the Mutations

Having used the complementation tests to narrow down the number of genes in question, we next hoped to determine the locus of each mutation. *C. briggsae* has five somatic chromosomes and one X chromosome. Sex is determined by an XO system, males having only the five somatic chromosomes plus one X while hermaphrodites have two X chromosomes. None of the multivulval mutations behave as if they are sex-linked. In order to establish in what

general region of the chromosome each mutation is located, a series of mapping experiments was performed.

The mutations were initially induced in wild-type animals of the AF16 strain. HK104 is another wild-type strain, with over 30,000 identified regions that are polymorphic with respect to AF16 (Koboldt et al). Animals of the two strains are morphologically and developmentally identical. HK104 males were crossed with AF16 hermaphrodites that were homozygous for a multivulval mutation. These offspring were then allowed to self-cross, and from the F2 generation a set of multivulval and a set of non-multivulval animals were isolated. For the chromosome that contains the mutation, the multivulval set will be enriched for the AF16 background, while the non-multivulval set will be slightly enriched for HK104. Because there are many small polymorphic regions between the HK104 and AF16 strains, it is possible to amplify short regions of DNA harvested from the multivulval and non-multivulval sets, and by comparing the two determine whether there is enrichment for the AF16 background in the multivulval animals. A large set of primers have been developed by Dr. Daniel Koboldt for this purpose. Figure 5 shows the PCR products of a mapping experiment showing X3a to be linked to a region near the center of chromosome II.

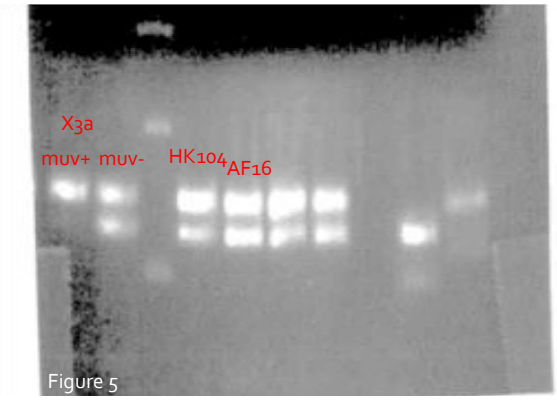


Figure 5

Sequencing

Following the tentative identification of the loci responsible for each strain's mutation, candidate genes in the area in question can be amplified using a PCR reaction and sequenced. Genes are amplified with nested primers specifically designed for each gene. A fraction of the PCR sample is run on a 1% agarose gel to check for amplification of a product of the desired size. If the product failed to amplify, a second internal or "nested" set of primers can be used to amplify the products present in the first PCR reaction. This technique can sometimes enhance the concentration of the desired product beyond what is possible using a single reaction.

After the product is visually confirmed on a gel the remainder of the sample is purified by running it on an agarose gel and excising the band of DNA, which is then purified on a column using a standard Qiagen kit. Amounts of DNA in the final sample are quantified using a spectrophotometer and appropriate dilutions of the sample and accompanying primers are delivered to Ohio State's Plant-Microbe Genomics Facility for sequencing. The sequencing data

are later downloaded and saved as .ape files to be compared to the known *C. briggsae* sequences available at wormbase.org. Sequences are visually compared using the program Sequencher to search for mutations such as point mutations in the form of transitions or transversions, or single-base-pair deletions. We predict the presence of such a mutation within an exon of the gene identified by the mapping experiments to be associated with the multivulval strain in question.

Results

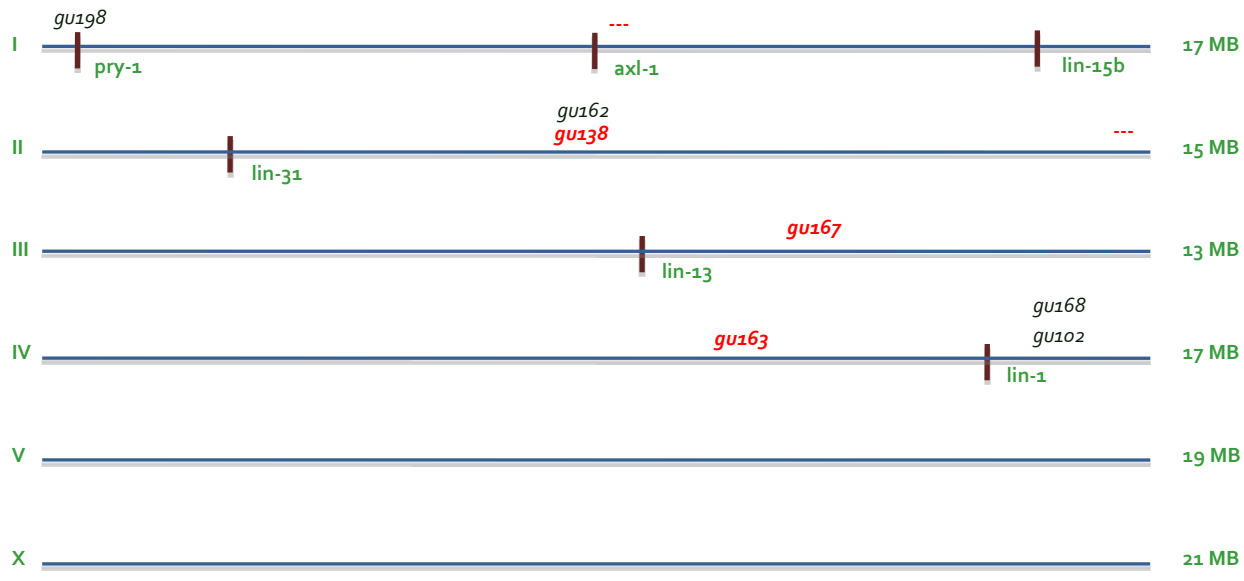
	hermaphrodite	**muv/muv-**		**dpy/dpy**				
male								
muv-/muv+	gu102 (B16b)	gu167 (Q2b)	gu138 (S3a)	gu137 (S8a)	gu162 (X3a)	gu163 (X13a)	gu168 (AA7b)	gu198 (AC18a)
gu102 (B16b)	0.38	0.00	0.00	nd	0.00	0.00	**0.24**	0.00
gu167 (Q2b)	0.00	0.27	0.05	nd	nd	0.00	0.00	0.00
gu138 (S3a)	0.00	0.00	nd	nd	**0.52**	0.00	0.00	nd
gu137 (S8a)	nd	nd	nd	nd	0.00	nd	0.00	0.00
gu162 (X3a)	0.22	0.00	**0.47**	nd	0.30	0.29	nd	nd
gu163 (X13a)	nd	0.00	0.00	nd	0.00	nd	0.00	nd
gu168 (AA7b)	nd	nd	0.00	nd	0.00	nd	0.38	nd
gu198 (AC18a)	0.00	0.02	0.01	nd	0.00	nd	0.00	nd

Figure 6

Figure 6 shows a table of the results from the complementation tests. Some are still in progress, and marked 'nd' for 'not determined'. Note that no crosses into *gu137* hermaphrodites have been performed because it has not been possible to create a strain that is homozygous both for the *gu137* mutation and for the dumpy mutation, evidence that suggests the two may be linked. The numbers represent the proportion of the F2 animals that displayed the multivulval phenotype, where N>=25. On average, 50% of the F2 animals are expected to have inherited a wild-type allele from the male and be phenotypically wild-type. The two cells marked in red indicate crosses that failed to complement, strongly suggesting that *gu162* and *gu138* are allelic mutations in the same gene. The cell in bold indicates another failure to complement, *gu102* into *gu168*, that has not been contradicted. Only 24% of the F2 animals exhibited the multivulval phenotype, less than the expected 50%. This may be the result of differential growth and survival rates between wild-type and multivulval animals. Results of the *gu168* into *gu102* cross are not yet complete. Crosses of *gu162* into *gu102* and *gu163* also produced low levels of multivulval offspring, at 22% and 29% respectively. However we do not believe these results to be indicative of an actual failure to complement, as neither reverse cross *gu102* into *gu162* or *gu163* into *gu162* produced any multivulval offspring.

The most conclusive preliminary results from the mapping experiments are shown in figure 7 below. (See appendix for details of the complete mapping data). In *C. elegans*, single point mutations in genes *pry-1*, *lin-31*, *lin-13*, and *lin-1* are all sufficient to produce the multivulval phenotype. Mutations in *axl-1* and *lin-15b* produce the multivulval phenotype only in conjunction with mutations in other genes. Figure 7 shows in green the locations of the homologues of genes associated with multivulval phenotypes in *C. elegans*. We have high confidence of the location of the mutant alleles listed in red- i.e. the results have proven to be repeatable. The mutations in black are suspected, but the evidence for their position is less strong. Currently we are in the early stages of sequencing some of the candidate loci in hopes of pinpointing the exact location and nature of the mutations.

Figure 7



Discussion

Based on the results of the complementation tests, we believe that at least six and possibly seven of the eight multivulval *C. briggsae* strains resulted from mutations in genes that are independent and distinct from one another. At least two of the multivulval alleles, *gu138* and *gu162*, exhibit a failure to complement and are likely the products of mutations within the same gene. It is yet to be determined whether *gu102* and *gu168* consistently fail to complement, but preliminary data suggests that this may be the case. As single mutations in only four genes in the *C. briggsae* sister species *C. elegans* are sufficient to yield a multivulval phenotype it is possible that this mapping experiment may result in the discovery of additional genes in *C. briggsae* previously unknown to influence vulval organogenesis.

Tentative locations of seven of the eight multivulval alleles have been determined using standard gene mapping methods. It has not been possible to map the *gu137* strain as we have been unable to construct a muv-/muv-;dpy/dpy line according to the usual protocol. This is likely the result of linkage between the two, suggesting the *gu137* muv- allele and the dpy allele are located in close proximity to one another on a chromosome. Rather than map the *gu137* mutation directly, we are attempting to map the dpy mutation itself in order to narrow down the approximate location of both mutations. Currently candidate genes in six of the eight multivulval strains are being amplified and sequenced for analysis: *lin-13* in *gu167*; *lin-31* in *gu138* and *gu162*; *lin-1* in *gu102* and *gu168*; and *pry-1* in *gu198*. Pending complete sequence data we will analyze the DNA to determine if and where the mutations of interest have occurred.

Understanding the genetic basis for the development of the multivulval phenotype is the first step in understanding the evolution of the complex molecular networks involved in vulval organogenesis. The phenotype results from inappropriate excess cell division, a process that parallels the development of malignant tumors. Some cancerous cells are known to be influenced by mutations to EGF pathways, especially those activating EGFR to cause overexpression, just as VPCs can begin to divide improperly in an environment containing high levels of transcription factor EGF (Zhang et al). Research of this type is critical to aid our understanding of how genetic changes act to regulate molecular pathways and development.

Acknowledgments

This project was advised by Associate Professor Dr. Helen Chamberlin, of the Ohio State University Molecular Genetics Department.

The primers and methods for mapping mutations using polymorphic sites between HK104 and AF16 was developed by Dr. Daniel Koboldt et al.

This research was made possible by an award from the Pelotonia Fellowship Program, a grant from the Ohio State University College of Arts and Sciences, and an NSF grant.

Many thanks to the members of the Chamberlin lab for their help and support.

References

Andersen, E.C. 2008. *The synthetic multivulva genes and their suppressors regulate opposing cell fates through chromatin remodeling*. Massachusetts Institute of Technology. http://hdl.handle.net/1721.1/42072

Cho et al. 2004. A phylogeny of Caenorhabditis reveals frequent loss of introns during nematode evolution. Genome Research 14:1207-1220.

Ferguson, E. L. and H. R. Horvitz. 1989. The multivulva phenotype of certain *Caenorhabditis elegans* mutants results from defects in two functionally redundant pathways. Genetics Society of America 123:109-121.

Ferguson et al. 1987. *A genetic pathway for the specification of vulval cell lineages in* Caenorhabditis elegans. Nature 326:259-267.

Gupta et al. 2007. Genomics and biology of the nematode *Caenorhabditis briggsae.* WormBook, ed. The *C. elegans* Research Community. doi/10.1895/wormbook.1.136.1. http://www.wormbook.org

Gupta et al. 2003. *Genetics analysis of the* C. briggsae *multivulva mutants*. International Worm Meeting 1008.

Jacobs et al. 1998. *Gain-of-function mutations in the* Caenorhabditis elegans lin-1 *ETS gene identify a C-terminal regulatory domain phosphorylated by ERK MAP Kinase*. Genetics Society of America 149:1809-1822.

Koboldt et al. 2010. A toolkit for rapid gene mapping in the nematode *Caenorhabditis briggsae.* BMC Genomics 11:236.

Sulston, J.E. and J. Hodgkin. 1988. The nematode *Caenorhabditis elegans*, W. Wood ed. Cold Spring Harbor, New York Cold Spring Harbor Laboratory. 87-606

wormbase.org

Zhang et al. 2007. *ErbB receptors: from oncogenes to targeted cancer therapies*. The Journal of Clinical Investigation 117(8):2051-2058.

Appendix

Mapping Data

Table 1: *gu102* (B16b)*; gu167* (Q2b)*; gu163* (X13a)*; gu138* (S3a)

DNA prep 1 (2.3.11)

Chr. #	Position (cM)	#	Linkage?	No Linkage	PCR	Gel
1	0.7	bhp-19	B16b	Q2b;X13a;S3a	2.28.11	3.1.11
		bhp-19	B16b	Q2b;X13a;S3a	3.30.11	4.1.11
1	29.2	bhp-1	none	B16b;Q2b;X13a;S3a	2.28.11	3.1.11
1	~52	bhp-29	Q2B	B16B;Q2B;X13A	3.1.11	3.7.11
2	~10	bhp-2	___	Q2b(?);X13a(?);S3a(?)	2.22.11	3.1.11
2	23.3-28.6	bhp-21	S3a	B16b;Q2b;X13a	2.8.11	2.9.11
2	49.9	bhp-8	none	B16b;Q2b;X13a;S3a	2.15.11	2.17.11
3	12.5	bhp-14	X13a	B16b;Q2b;S3a	2.4.11	2.7.11
3	30	bhp-40	Q2b	B16b;X13a;S3a	2.8.11	2.9.11
4	1.9-5.1	bhp-13	Q2b;X13a	B16b;S3a	2.17.11	2.18.11
4	20.6	bhp-11				
4	31	bhp-9	X13a	B16b;Q2b;S3a	2.15.11	2.17.11
4	43.5	bhp-16				
4	57.8	bhp-30	___	?	2.17.11	2.18.11
5	2.5-3.2	bhp-31	X13a	B16b;Q2b;S3a	2.18.11	2.21.11
5	26.7	bhp-5	___	?	2.18.11	2.21.11
5	56.9	bhp-24	none	B16b;Q2b;X13a;S3a	2.21.11	2.22.11
X	8.4	bhp-25	___	?	2.21.11	2.22.11
X	21-21.7	bhp-26	___	S3a	2.22.11	3.1.11

Table 2: *gu102* (B16b)*; gu167* (Q2b)*; gu163* (X13a)*; gu138* (S3a)

DNA prep 2 (2.8.11)

Chr. #	Position (cM)	#	Linkage?	No Linkage	PCR	Gel
1	0.7	bhp-19	X13a	B16b;Q2b;S3a	3.7.11	3.8.11
1	.7	bhp-19	X13a	B16b;Q2b;S3a	3.30.11	4.1.11
1	29.2	bhp-1	none	B16b;Q2b;X13a;S3a	3.14.11	3.16.11
1	~52	bhp-29	none	B16b;Q2b;X13a;S3a	3.8.11	3.10.11
2	~10	bhp-2	___		3.11.11	3.14.11
2	~25	bhp-21	S3a	B16b;Q2b;X13a	3.11.11	3.14.11
2	49.9	bhp-8	none	B16b;Q2b;X13a	3.28.11	3.30.11
3	12.5	bhp-14	none	B16b;Q2b;X13a;S3a	3.16.11	3.28.11
3	30	bhp-40	Q2b?	B16b;X13a;Q2b	3.7.11	3.8.11
4	1.9-5.1	bhp-13	Q2b? X13a?	B16b;S3a	3.8.11	3.10.11
4	20.6	bhp-11	___	____	3.10.11	3.11.11
4	20.6	bhp-11	___	____	3.29.11	3.31.11
4	31	bhp-9	Q2b;X13a;B16b?	S3a	3.16.11	3.28.11
4	43.5	bhp-16	___	____	3.10.11	3.11.11
4	43.5	bhp-16	none	B16b;Q2b;X13a;S3a	3.29.11	3.31.11
4	57.8	bhp-30				
5	2.5-3.2	bhp-31	none	B16b;Q2b;X13a;S3a	3.14.11	3.16.11
5	26.7	bhp-5				
5	56.9	bhp-24	X13a	B16b;Q2b;S3a	3.28.11	3.30.11
X	8.4	bhp-25				
X	21-21.7	bhp-26				

Table 3: *gu162* (X3a); *gu168* (AA7b); *gu198* (AC18a)

DNA prep 1 (2.8.11)

Chr. #	Position (cM)	#	Linkage?	No Linkage	PCR	Gel
1	0.7	bhp-19	AC18a	X3a;AA7b	4.4.11	4.6.11
1	29.2	bhp-1	none	X3a;AA7b;AC18a	3.31.11	4.4.11
1	~52	bhp-29	none	X3a;AA7b;AC18a	3.31.11	4.4.11
2	~10	bhp-2				
2	23.3-28.6	bhp-21	X3a	AA7b;AC18a	4.4.11	4.6.11
2	49.9	bhp-8	none	X3a;AA7b;AC18a	4.7.11	4.8.11
3	12.5	bhp-14	none	X3a;AA7b;AC18a	4.7.11	4.8.11
3	30	bhp-40	none	X3a;AA7b;AC18a	4.5.11	4.7.11
4	1.9-5.1	bhp-13				
4	20.6	bhp-11				
4	31	bhp-9				
4	43.5	bhp-16	none	X3a;AA7b;AC18a	4.5.11	4.7.11
4	57.8	bhp-30				
5	2.5-3.2	bhp-31	none	X3a;AA7b;AC18a	4.11.11	4.12.11
5	26.7	bhp-5				
5	56.9	bhp-24	none	X3a;AA7b;AC18a	4.11.11	4.12.11
X	8.4	bhp-25				
X	21-21.7	bhp-26				

Table 4: *e9a* (dumpy mutation)

DNA prep 1 (4.11.11)

Chr. #	Position (cM)	#	Linkage?	PCR	Gel	Notes
1	0.7	bhp-19	no	4.13.11	4.14.11	
1	29.2	bhp-1	no	4.26.11	4.27.11	
1	~52	bhp-29	no	4.13.11	4.14.11	
2	~10	bhp-2	no	4.28.11	5.4.11	
2	~10	bhp-33		5.13.11		
2	23.3-28.6	bhp-21	no	4.13.11	4.14.11	
2	49.9	bhp-8	no	4.13.11	4.14.11	
3	0	bhp-18		5.13.11		
3	12.5	bhp-14	no	4.13.11	4.14.11	
3	16-17	bhp-38		5.13.11		
3	21.2	bhp-12		5.13.11		
3	21.8	bhp-34		5.13.11		
3	30	bhp-40	?	4.15.11	4.18.11	fail to amplify
4	1.9-5.1	bhp-13	?	4.15.11	4.18.11	fail to amplify
4	7.9	bhp-15		5.13.11		
4	20.6	bhp-11	?	4.28.11	5.4.11	fail to amplify
4	31	bhp-9	no	4.15.11	4.18.11	
4	43.5	bhp-16	?	4.26.11	4.27.11	fail to amplify
4	57.8	bhp-30				
5	2.5-3.2	bhp-31	no	4.20.11	4.25.11	pos controls fail
5	26.7	bhp-5	no	4.20.11	4.25.11	pos controls fail
5	56.9	bhp-24	no	4.20.11	4.25.11	pos controls fail
X	8.4	bhp-25				
X	21-21.7	bhp-26				