A Molecular Analysis of Mutations at the Complex dumpy Locus in Drosophila melanogaster

Amber Carmon¹, Michael J. Guertin¹, Olga Grushko², Brad Marshall¹, Ross MacIntyre¹*

¹ Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York, United States of America, ² Life Sciences Institute, University of Michigan, Ann Arbor, Michigan, United States of America

Abstract

The Drosophila dumpy gene consists of seventy-eight coding exons and encodes a huge extracellular matrix protein containing large numbers of epidermal growth factor-like (EGF) modules and a novel module called dumpy (DPY). A molecular analysis of forty-five mutations in the dumpy gene of Drosophila melanogaster was carried out. Mutations in this gene affect three phenotypes: wing shape, thoracic cuticular defects, and lethality. Most of the mutations were chemically induced in a single dumpy allele and were analyzed using a nuclease that cleaves single base pair mismatches in reannealed duplexes followed by DHPLC. Additionally, several spontaneous mutations were analyzed. Virtually all of the chemically induced mutations except for several in a single exon, either generate nonsense codons or lesions that result in downstream stop codons in the reading frame. The remaining chemically induced mutations remove splice sites in the nascent dumpy message. We propose that the vast majority of nonsense mutations that affect all three basic dumpy phenotypes are in constitutive exons, whereas nonsense mutants that remove only one or two of the basic functions are in alternatively spliced exons. Evolutionary comparisons of the dumpy gene from seven Drosophila species show strong conservation of the 5′ ends of exons where mutants with partial dumpy function are found. In addition, reverse transcription PCR analyses reveal transcripts in which exons marked by nonsense mutations with partial dumpy function are absent.

Introduction

The history of the dumpy gene in Drosophila melanogaster encompasses virtually the entire history of Drosophila genetics itself. Early last century, several mutants which initially seemed to have different phenotypes were recovered by the Morgan lab at Columbia University. Morgan himself noticed a fly in August of 1910 with shortened wings which he called Truncate [1–3]. A fly with pits on the thorax and whorls of the bristles was found in 1916 and termed vortex-II due to its location on the second chromosome [4]. In 1918 [5] a fly was discovered with both shortened wings and with whorls of bristles and hairs on the thorax. This mutant was called dumpy, the first time this term was used. These mutants, along with a second vortex mutant and another mutant named thoraxate showing thoracic vortices and homozygous lethality, were eventually combined by Bridges and Muller as a series of recessive allelomorphs possibly occurring in different parts of a single gene [6–8]. In the 1950s, Elso Carlson, then at UCLA, and his students generated a large number of dumpy mutant alleles, primarily with chemical mutagens [9–13]. A genetic fine structure map with discrete subloci was developed culminating in the map published by Dale Grace in 1980 [14–17].

Beginning in the middle of the last century, Drosophila geneticists defined and analyzed a number of complex loci. Like dumpy, these genes were characterized primarily by mutations with different and sometimes overlapping phenotypes, complex patterns of complementation, and genetic fine structure maps exhibiting separable clusters of mutant sites called subloci. With the advent of molecular cloning and sequencing, the underlying basis for the phenotypic complexity and the complementation patterns of many, if not most, of these loci could be explained. In addition, different functions could be assigned to groups of mutant alleles mapping at discrete subloci in fine structure maps of the genes. Two genes where cloning and sequencing provided explanations for their complexity are rudimentary, where complementing mutants affect distinct domains in the protein [18] and cut, where complementing mutants map either in the regulatory region or in the coding exons of the gene [19]. In contrast, the complexity of the dumpy gene in Drosophila melanogaster, despite being cloned and sequenced [20], has remained unexplained.

Recessive mutant alleles of dumpy have three primary effects: oblique (dpv) that affect the shape of the wing, vortex (dpv) that disrupt the attachments of indirect flight muscle to the dorsal thoracic cuticle causing pits and protrusions, and lethal (dpl) acting mostly at larval moults. The oblique and vortex phenotypes are shown in Figure 1, b and d respectively from Wilkin et al. [20].
Pleiotropic individual alleles of dumpy, shown in Table 1, can exhibit any combination of the three mutant phenes, and heteroallelic heterozygotes will show the phenotype of the homozygous allele, e.g. dp<sup>ov</sup> / dp<sup>l</sup> flies will be viable with normal wings but mutant for vortex.

Importantly, there are also cases of intragenic or interallelic complementation between some dp<sup>ol</sup>, dp<sup>lv</sup>, dp<sup>dumpy</sup>, and dp<sup>olv</sup> alleles—marked with an asterisk in Table 1—revealing additional genetic complexity, presumably reflecting different biological roles for dumpy at different developmental stages.

The large size of the dumpy gene (the largest euchromatic gene in Drosophila) has made the construction of fine structure maps of the locus feasible. A detailed map—adapted from Grace’s paper [17] to include just the mutants analyzed in this study—is shown in Figure 1. Note that dp<sup>ol</sup>, dp<sup>lv</sup>, dp<sup>dumpy</sup>, and dp<sup>olv</sup> alleles occupy recombinationally distinct subloci, whereas dp<sup>olv</sup> alleles are found throughout the locus. In Grace’s original genetic map, dp<sup>ol</sup> alleles also mapped at several places in the gene.

As shown in Figure 2, dumpy encodes a large protein comprised of more than 300 epidermal growth factor (EGF) repeats, a class of modules found in many extracellular matrix (ECM) proteins. Most of the EGF modules are interspersed with a novel repeat of 21 amino acids, which we call “PIGSFEAST” (PF) (adapted from Jenkins [13]). The EGF modules are interrupted by an insert of a repetitive, proline-rich sequence three-module EGF-DPY-EGF unit. The EGF-DPY-EGF repeats of the Dumpy protein is composed of contiguous repeats of a single Zona Pellucida (ZP) domain, found in a number of important ECM proteins where they mediate homotypic and heterotypic covalent crosslinking to other ZP domains.

Dumpy, along with two other ZP domain proteins, Piopio and Papilotte, functions in the adhesion of the apical surface of the Drosophila wing epithelia to the overlying cuticle, and loss of function of each of these three genes results in a blistering phenotype in the wing [23,24]. A further role of Dumpy in cuticle adhesion is revealed by certain larval lethal dumpy mutations that fail to molt due to a failure of cuticle detachment rather than a failure of adhesion. Dumpy also plays an important role in the epithelial cells that mediate the attachment of the muscles to the overlying cuticle. As mentioned above, dumpy vortex (dp<sup>v</sup>) mutations result in depressions or pits in the cuticle where it overlies the muscle attachment sites. During embryogenesis, dumpy is expressed in many tube-forming structures that form an apical ECM that lines their internal lumens. These include the salivary gland, fore and hind-gut, and developing trachea [20]. Certain embryonic lethal dumpy mutations result in failure of tracheal cells in the small vessels to connect to form tubes [25]. The effect of dumpy mutations on the trachea may be responsible for the lethal phenotype of dp<sup>v</sup>, dp<sup>olv</sup>, dp<sup>lv</sup>, and dp<sup>olv</sup> mutations. Hence, Dumpy has functionally diverse roles including cell adhesion, ECM assembly and mechanical properties, morphogenesis and tube formation, and as a ZP domain containing protein may interact with and modulate developmental signaling pathways [25,26].

In this paper, we identify the molecular lesions responsible for some 45 dumpy mutants including examples of each kind of mutant allele shown in Table 1, and those that either complement or fail to complement other alleles. We report in this paper that most dumpy mutants are directly due to or lead to downstream nonsense codons, even when the mutation disrupts only one or two of the three basic mutant phenotypes. We propose that such mutants mark alternatively spliced exons whereas mutants which affect all three phenotypes (dp<sup>olv</sup>) are located in constitutive exons. We provide some experimental evidence for this hypothesis using RT-PCR analyses. We also discuss the possibility that the complementation of certain dp<sup>olv</sup> mutations results from trans-splicing. Hence, alternative cis and trans-splicing events generating different and perhaps tissue specific Dumpy isoforms can provide a rationale for the complexity of this long studied Drosophila gene.

**Results**

Properties of chemically induced dumpy mutations

The crosses employed in the screens for EMS induced dumpy mutations in defined isolelic backgrounds are outlined in Table 2. The distribution of mutations from crosses 1, 2A, and 2B is as follows: 60 dp<sup>olv</sup>, 32 dp<sup>lv</sup>, 7 dp<sup>dumpy</sup>, 2 dp<sup>ol</sup>, and 2 are dp<sup>v</sup>. Like Jenkins [13], the majority of our mutants were dp<sup>olv</sup>. However, in our case dp<sup>ol</sup> mutants outnumbered dp<sup>v</sup> mutants. From the screens depicted in crosses 2A and 2B in Table 2, we recovered 90 transmitted

---

**Table 1. Phenotypes and complementation patterns of the classes of dumpy mutant alleles.**

| Allele | Phenotype | Phenotypes of heterozygotes |
|-------|-----------|-----------------------------|
|       |           | dp<sup>o</sup> | dp<sup>l</sup> | dp<sup>olv</sup> | dp<sup>dlv</sup> | dp<sup>d</sup> | dp<sup>olv</sup> |
| dp<sup>o</sup> | oblique | O | + | O | O | + | O |
| dp<sup>v</sup> | vortex | V | + | V | + | V |
| dp<sup>d</sup> | lethal | L | + | L | L* |
| dp<sup>olv</sup> | oblique, vortex | OV | O | V | OV |
| dp<sup>d</sup> | oblique, lethal | L | L* |
| dp<sup>olv</sup> | vortex, lethal | L | L* |
| dp<sup>ol</sup> | oblique, lethal, vortex | L |

+ = wild type, O = oblique wings, V = vortex, L = lethal.

*may show interallelic complementation for lethality.

doi:10.1371/journal.pone.0012319.t001
dumpy mutants, 46 from cross 2A in the net chromosome and 44 from cross 2B in the clot chromosome. All of these mutations are in an identical dumpy allele derived from an isofemale line from Australia. The mutants, along with the flanking visible marker, dumpy phenotypes, and the balancer chromosomes are listed in Table S1. The cross schemes followed in Table 2 also allowed us to detect dumpy lethal alleles which complement the dp lvI mutation in the CyO balancer chromosome.

We initiated the screens designated as crosses 3 and 4 in Table 2 to enrich for dumpy vortex mutants, since none was recovered from crosses 1 and 2. Cross 3, in which F1 males carrying the dp lvI mutant from the Bloomington stock collection in the presence of the e(dpv) mutation on the 3rd chromosome were scored, produced eleven dp olv mutations, one which complements dp lvI, and two new dp lv alleles. These are listed in Table S1 as dp olvRX or dp lvRX respectively. Since alleles with oblique phenotypes also came through this screen, we set up cross 4 in Table 2, this time examining F1 males carrying a previously generated dp lv allele and homozygous for e(dpv). 24,000 F1s were scored and four complementing dp olv mutants were obtained, along with a single dp ov mutant allele. Hence this screen appears to enrich for complementing dp olv mutations. We will discuss below how such mutations may help identify putative trans-splicing events in the dumpy gene.

Molecular basis of dumpy mutations

We outline in the Materials and Methods section the approaches we’ve taken to characterize preexisting dumpy mutations, those we generated in strains isozallelic for a wild type dumpy allele from Australia, and those recovered in a screen for spontaneous mutations. Most of these analyses relied on the generation of overlapping amplicons across the entire gene and the use of the WAVE dHPLC machine from Transgenomic, Inc. to detect cleavage fragments generated by the Surveyor nuclease at the sites of base pair mismatches. A typical dHPLC chromatogram is shown in Figure 3, where two mutations dp olv48a and dp olv104A are located in an amplicon from exon 11.

Table 3 and Figure 4 show our results to date. Clearly this approach is very effective in detecting and identifying mutations in the dumpy gene. The data are remarkable in that most of the mutations, including dp ol and dp lv mutants, result in stop codons either at the site of the mutation, are generated from a deletion, or cause the removal of a splice site. It is interesting that all missense

Table 2. Crosses used to produce dumpy mutants in defined chromosomal backgrounds.

| Cross | Mutagenized males | Females | F1 phenotype |
|-------|------------------|---------|--------------|
| 1     | cn bw- 2nd chromosome isogenic | dp lvI cl | oblique, vortex |
| 2A    | net, dp+ isozallele | net dp lvI cl | oblique, vortex |
| 2B    | cl, dp+ isozallele | dp lvI cn bw | oblique, vortex |
| 3     | cl; e(dpv) dp+ isozallele | dp lvI; e(dpv) | vortex |
| 4     | net; dp+ isozallele | In(2LR) Gla/dp lvI cl; e(dpv) | vortex |

*A single dp+ allele isolated from a wild type strain collected in Australia. See text for details.

Figure 2. Modular structure of the dumpy gene product. Adapted from Wilkin et al. [20]. Modules are designated as shown in the key. Note that a large part of the protein is composed of EGF-DPY-EGF triads, with two repeated regions containing PIGSFEAST (PF) and Proline rich repeats respectively. The N-terminus is enriched in calcium binding EGF repeats and the C-terminus contains a transmembrane domain and a Zona Pellucida (ZP) domain. The arrowheads denote the positions of introns in the gene.

doi:10.1371/journal.pone.0012319.g002
mutations identified to date change cysteines in the protein. Given the repetitive nature of the Dumpy protein, i.e. all the EGF-DPY-EGF motifs, perhaps most missense mutations don’t produce a visible phenotype.

The dp<sup>olv</sup> mutants are found scattered throughout the locus. Again, most identified so far introduce a stop codon or otherwise lead to a truncation of the protein via a frameshift, or remove a splice junction. It should be noted that this leads to the same severe recessive, loss of function alleles. This implies that, to retain any end of the gene.

other mutants are found clustered in discrete regions. We will discuss the identity of the two olv mutants which are located at many different places in the gene, an observation that is consistent with their many sites in Grace’s fine structure map [17], most of the other dumpy mutants are found clustered in discrete regions. We will discuss each of these clusters, proceeding from the 5’ to the 3’ end of the gene.

dp<sup>ol</sup> mutants

To date, we’ve not recovered any EMS induced vortex mutants in the dumpy allele in 2<sup>nd</sup> chromosomes marked with the <i>net</i> and <i>clot</i> mutants. We did, however, examine pre-existing dp<sup>ol</sup> alleles, dp<sup>ol1</sup> and dp<sup>ol2</sup> [4,7,8]. Both were originally recovered in the Morgan laboratory at Columbia University early last century. Using the primers listed in Table S2, we generated amplicons across the entire gene from each homozygous vortex mutant except with primers 5’19F and 5’19R. We then used long range PCR (see Materials and Methods) and recovered amplicons approximately 10kb in length from each mutant. Sequences from the ends of these amplicons indicated the presence of a <i>roo</i> element at bp 15448 upstream of the start codon, in the 5’ region of the gene in both alleles. If the <i>roo</i> element is responsible for the vortex mutant phenotype, its position in the gene is consistent with the position of the <i>ol</i> sublocus in Grace’s map. Except for <i>dp<sup>ol36a</sup></i>, these mutations produce or result in a stop codon. <i>dp<sup>ol36a</sup></i> is a missense mutation in which a cysteine residue in a Ca<sup>2+</sup> binding EGF motif is replaced by a tyrosine. The cysteine residues in these short motifs (ca 35 amino acids) are essential for their correct tertiary structure due to their participation in disulphide bonds [20].

The position of this missense mutation is in exon 5 and is the <i>dp<sup>ol</sup></i> mutant closest to the 5’ end of the gene (see Figure 4).

dp<sup>ov</sup> mutants

All of these mutations are in exon 11 or in the adjacent intron. The mutant <i>dp<sup>ov1</sup></i>, discovered by Morgan in 1918 [5], has been the canonical dumpy mutant used by many investigators in the last century. It is characterized by full penetrance and intermediate expressivity, e.g. virtually all females in <i>dp<sup>ov1</sup></i> containing stocks have oblique wings with an intermediate score of 3 on Dale Grace’s scale [17,27]. Males exhibit a lower penetrance and lower oblique scores on the Grace scale. We examined <i>dp<sup>ov1</sup></i> after finding the EMS induced <i>dp<sup>ov</sup></i> mutations in the ampiclon from exon 11 by sequencing. No nonsynonymous changes were found nor any changes affecting the canonical splice site.

mutants affecting exon 11 were then sequenced. The primers used to amplify a region of intron 11 failed to produce a product from <i>dp<sup>ov1</sup></i> DNA. Long range PCR, however, did produce a product containing a blood transposon. This insertion is in the intron just preceding the large exon encoding the PIGSFEAST repeats that are undergoing concerted evolution [21,22]. The EMS induced <i>dp<sup>ov</sup></i> mutants, including <i>dp<sup>ov4DG2</sup></i> from Dale Grace, are missense mutants affecting cysteine residues in the EGF-DPY diads that characterize exon 11. Curiously, except for <i>dp<sup>ov39a</sup></i> which affects exon 5 and both <i>dp<sup>ov</sup></i> mutants affecting exon 72, exon 11 is the only other exon in which missense mutations have been recovered.

dp<sup>sv</sup> mutants

We have characterized 11 <i>dp<sup>sv</sup></i> mutations and most either directly generate a stop codon or are out of frame deletions (see Table 3). The EMS generated mutant, <i>dp<sup>svR2</sup></i>, is a G to A transition removing a splice site between exons 43 and 44, and none of the <i>dp<sup>sv</sup></i> mutants is due to an amino acid substitution, although the region of the protein affected, viz exons 40–49 consists of consecutive repeats of EGF-DPY-EGF triad domains whose tertiary structures are surely stabilized by cysteines participating in disulphide bonds (see Figure 3c in Wilkin et al. [20]). Note also
Table 3. Results of molecular analyses of selected dumpy mutations.

| Mutant | Origin | Allele Class | Exon | Mutation or Deletion | Effect |
|--------|--------|--------------|------|-----------------------|--------|
| v2     | SC     | v            | 5’ region | roo transposon | unknown |
| 1C5    | BM     | olv          | 3    | G->A                 | removes splice site |
| 36a    | MG     | ol           | 5    | TGC->TAC            | Cys->Tyr |
| 38a    | MG     | ol           | 6    | 2bp deletion        | frameshift and stops |
| 67b    | MG     | ol           | 7    | TGT->TGA            | Cys->STOP |
| 2P1    | BM     | ol           | 7    | CAA->TAA            | Gln->STOP |
| 71a    | MG     | ol           | 7    | TGG->TAG            | Trp->STOP |
| 18b    | MG     | ol           | 7    | TCG->TAG            | Ser->STOP |
| D1311A | OG     | ol           | 7    | 15bp inversion      | creates a STOP |
| 2G1A   | BM     | olv          | 9    | CAA->TAA            | Gln->STOP |
| 105A   | MG     | olv          | 11   | CAA->TAA            | Gln->STOP |
| 56a    | MG     | ov           | 11   | TGG->TAT            | Cys->Tyr |
| 7b     | MG     | ov           | 11   | TGC->CGC            | Cys->Arg |
| 104A   | MG     | olv          | 11   | TGC->AGC            | Cys->Ser |
| 61B    | MG     | olv          | 11   | TGT->CGT            | Cys->Arg |
| A12    | RM     | ov           | 11   | TGC->TAC            | Cys->Tyr |
| 27B    | MG     | olv          | 11   | TGC->TAC            | Cys->Tyr |
| ovgDG2 | SC     | ov           | 11   | TGT->TAT            | Cys->Tyr |
| 48a    | MG     | olv          | 11   | TGC->TCC            | Cys->Ser |
| ov1    | SC     | ov           | Intron 11 | blood transposon  | unknown |
| 6      | MG     | olv          | 15   | 4bp deletion        | frameshift->STOP |
| R11    | RM     | olv          | 19   | CGA->TGA            | Arg->STOP |
| D2011A | OG     | olv          | 19   | 16bp deletion       | frameshift->STOP |
| G8202B | OG     | olv          | Intron 21 | 6bp deletion       | unknown |
| R4     | RM     | olv          | 33   | AGA->TGA            | Arg->STOP |
| 2C1    | BM     | olv          | 34   | TGT->TGA            | Cys->STOP |
| 89a    | MG     | olv          | 34   | CAG->TAG            | Gln->STOP |
| G3030B | OG     | lv           | 40   | 89bp deletion       | frameshift ->STOP |
| L2311B | OG     | lv           | 40   | TAC->TAA            | Tyr->STOP |
| 23b    | MG     | lv           | 43   | 368bp deletion      | frameshift |
| H1230B | OG     | lv           | 43 to 45 | 1140bp deletion   | frameshift |
| lvR2   | RM     | lv           | 43   | G->A                | removes splice site |
| 16     | MP     | lv           | 45   | 10bp deletion       | frameshift and stops |
| 12     | MP     | lv           | 46   | 139bp deletion      | removes splice site |
| 7a     | MG     | lv           | 46   | CAG->TAG            | Gln->STOP |
| D1119A | OG     | lv           | 47   | CAA->TAA            | Gln->STOP |
| 65f    | MG     | lv           | 48   | CAG->TAG            | Gln->STOP |
| P1129B | OG     | lv           | 49   | 1bp deletion        | frameshift |
| GD82   | SC     | l            | 58   | GAG->TAG            | Gln->STOP |
| o2     | SC     | o            | 72   | TGT->TAT            | Cys->Tyr |
| 014b   | MG     | o            | 72   | TGT->TAT            | Cys->Tyr |
| 5B1    | BM     | olv          | 73   | CAA->TAA            | Gln->STOP |
| 1B     | BM     | olv          | 76   | 1482bp deletion     | unknown |
| 21C2   | BM     | olv          | 76   | CAG->TAG            | Gln->STOP |
| R3     | RM     | olv          | 76   | CAG->TAG            | Gln->STOP |

doi:10.1371/journal.pone.0012319.t003
that there are no dp<sup>olv</sup> mutations located in the dp<sup>l</sup> region, nor did Grace map any dp<sup>olv</sup> mutations in the lv sublocus. Once again the positions of the dp<sup>olv</sup> mutations in the protein are colinear with the position of the lv sublocus in Grace's map.

**dp<sup>l</sup> mutant**

We did not recover any dp<sup>l</sup> mutants in our screen for EMS induced mutants, nor did we expect to given the design of the cross scheme. We screened F1s for oblique and/or vortex phenotypes over dp<sup>olv</sup> or dp<sup>olvR1</sup> alleles. Indeed, it is difficult to envision an F1 screen for mosaics which would allow for the recovery of dp<sup>l</sup> mutations. A search of the literature including Masters and PhD theses did not reveal how such mutants were recovered. Nevertheless, we were able to obtain two mutants, dp<sup>olvG12</sup> and dp<sup>olvR13</sup>, from the Kyoto stock center induced by Dale Grace [15,17]. Crosses with these mutually non-complementing mutants do indeed confirm their status as dp<sup>l</sup> mutants, i.e. they produce wild type F1 adults when crossed to dp<sup>l</sup>, dp<sup>l</sup>, or dp<sup>olv</sup> mutants and F1s from crosses to dp<sup>l</sup>, dp<sup>l</sup>, and most dp<sup>olv</sup> flies do not survive to adulthood (see Table 1). We determined that dp<sup>olvG12</sup> is due to a nonsense mutation in exon 58. The dp<sup>l</sup> mutant, identified as distinct recombinationally from the lv sublocus, is also molecularly discrete from the exons marked by dp<sup>olv</sup> mutations.

**dp<sup>olv</sup> mutants**

Grace mapped dp<sup>olv</sup> mutations at many different sites in the gene, and we also find these mutants at many different places in the Dumpy protein. For example, dp<sup>olv1C5</sup>, which affects the 3<sup>rd</sup> exon, dp<sup>olv2C1</sup> and dp<sup>olv2C1C21</sup>, are both stop codons in exon 34, whereas dp<sup>olv21C2</sup> and dp<sup>olvR3</sup> are due to nonsense codons in exon 76 which encodes the ZP domain very near the C-terminus. Except for those mutations in exon 11 and dp<sup>olv1C5</sup>, which results in the removal of a splice site, dp<sup>olv</sup> mutants result from either stop codons or deletions, which generate frameshifts and downstream stop codons. Dp<sup>olv12B1</sup> is a very large in frame deletion which removes a large portion of exon 76. Again, in agreement with Grace's genetic map, no dp<sup>olv</sup> mutations are found in the exons of the ol, lv, or l subloci. The lv sublocus, presumably encompassing only exon 11 and an adjacent intron, is another story. Here Grace mapped dp<sup>l</sup>, dp<sup>olv</sup>, and dp<sup>l</sup> mutations at the same site, given the limited resolving power of recombination in a higher eukaryote such as Drosophila. We too find both dp<sup>olv</sup> and dp<sup>l</sup> mutations in exon 11, but curiously, and except for dp<sup>olv105A</sup>, the dp<sup>olv</sup> mutations in this exon and only this exon are missense mutations, all four of which substitute a different amino acid for a cysteine residue. We will discuss below how dumpy mutations with several different phenotypes could be found in a single exon. We have also observed that certain dp<sup>olv</sup> mutations will complement other dumpy lethal alleles, particularly other dp<sup>l</sup> and dp<sup>olv</sup> mutations. In these cases, the surviving F1s show good viability but will exhibit vortices or have oblique wings respectively. In these cases the complementing mutations result from a stop codon. We also find cases of complementation between different dp<sup>olv</sup> mutants for example, dp<sup>olv105A/dp<sup>olv</sup>66</sup> F1s are fully viable but have vortices and oblique wings. Note in Figure 4 that complementing dp<sup>olv</sup> mutants we have analyzed appear to closely flank the highly repeated PIGSFEAST region, and indeed, all but dp<sup>olv</sup>66 are located in exon 11. Three of the complementing dp<sup>olv</sup> mutants in exon 11 are missense mutations but the complementing mutant dp<sup>olv105A</sup> is due to a nonsense mutation, and dp<sup>olv</sup>66 in exon 15 on the other side of the PIGSFEAST exon is a frameshift mutation which generates a stop codon.

**dp<sup>o</sup> mutants**

Our sample of sequenced mutants is deficient for oblique or dp<sup>o</sup> alleles. These unfortunately are only rarely recovered in EMS screens, although Grace found that, like dp<sup>olv</sup> mutants, they map at many places in the gene. We did analyze two dp<sup>o</sup> mutations. Both are missense mutations that, remarkably, are due to G to A transitions of the same nucleotide resulting in cysteine to tyrosine...
substitutions. We are certain these are different mutations since the SNP patterns and synonymous substitutions in the chromosomes surrounding the site are very different.

Alternative splicing in Dumpy: Evolutionary evidence

Our molecular analyses of the dumpy mutants indicates most are due to nonsense mutations. One might predict, if dumpy encodes a single transcript and translated message, that most, if not all, of these would affect all three basic functions, i.e. wing shape, tendon cell-cuticle attachment and ultimately viability. Hence they should have a dpolv phenotype. How then do we explain the observations that the dpolv, dploc, and dpmutants, i.e. those that have only partial dumpy function, are also due to the presence of stop codons in the dumpy message? We propose that these mutations producing partial functions will be found in alternatively spliced exons. For example, exons tagged by dpolv nonsense mutations will be expressed in certain tissues, e.g. in tendon cells and in the trachea, the latter presumably necessary for viability, but not present in dumpy messages in the developing wing. Dpolv nonsense mutants would be found in so called constitutive exons expressed in most, if not all, tissues at all developmental stages. Other explanations of our results are discussed below.

The hypothesis that the dumpy gene encodes both alternative and constitutive exons makes several predictions. First, there are distinct differences between alternative and constitutive exons in other systems. Xing and Lee [28,29] noted that RNA sequences from alternatively spliced exon/intron boundaries leads to selection pressure for nucleotide sequence conservation in these regions while there is significantly less conservation in constitutive exons. Thus, they noted that Ks, the number of synonymous substitutions per synonymous site, is much lower in human-mouse comparisons of alternatively spliced exons than in constitutive exons. To assess sequence divergence in dumpy’s exons, we compared the first 30 nucleotides of the exons from seven Drosophila species (D. melanogaster, D. ananassae, D. pseudoobscura, D. willistoni, D. mojavensis, D. virilis, and D. grimshawi), as shown in Figure 5.

It should be noted that the seven species compared in a pairwise fashion (the averages are shown in Figure 5) are, in each case, from different subgenera or from different species groups. The conservation of the 5’ ends of exons marked by dpolv (blue stars) mutations is striking. Exon 11 (green star) and exon 15 (red star), both of which are affected by complementing dpolv mutations, are also striking in their conservation of 5’ nucleotides. The nearly complete sequence identity of the 5’ ends of these exons over 60 million years of evolutionary time indicates there is a highly conserved interaction between the dumpy message from these regions and proteins involved in the splicing process, possibly in a tissue specific manner [30]. The pattern of conservation in the lv and l regions is also very interesting (bracketed yellow and orange areas). There may be several different mechanisms creating various alternative transcripts in the lv and l region such as competing intronic RNA secondary structures [31], steric hinderance of multiple splicing factor binding sites, or major and minor splicesome usage [32]. In general, exons with non-complementing dpolv mutations do not show marked 5’ end conservation. For example, exons 19 and 34 show higher levels of nucleotide divergence viz. 0.178 and 0.512 respectively and each is

Figure 5. Graph of average dnucleotide differences between the first 30 sites at the 5’ ends of dumpy exons from seven Drosophila species. There is strong conservation of the sequences at the 5’ ends of of exons (blue stars), exon 11 (green star), and exon 15 containing the complementing mutation dpolv (red star). In general, exons containing other dpolv mutations show higher levels of divergence. See Materials and Methods for details regarding alignment and computational procedures.

doi:10.1371/journal.pone.0012319.g005
marked by two non complementing dpolv nonsense mutations, dpolvR11 and dpolvD2011A in exon 19, and dpolv2C1 and dpolv89a in exon 34. As in the case of the human-mouse comparisons [28,29], there is marked conservation in some alternatively spliced exons.

Alternative splicing in Dumpy: Evidence from RT-PCR

We have evidence that alternatively spliced dumpy mRNAs can be detected by RT-PCR. We extracted mRNA from wild type 3rd instar larvae and from Drosophila S2 cell lines and used primers spanning the set of ol and lv exons. Primers spanning the two exons of the ZP domain were used as a positive control. We also chose primers located in exons marked by dpolv nonsense mutations that we believe are constitutive.

Figure 6 is a gel image of the RT-PCR products obtained from the two sources of mRNA. There are a number of shorter products, some of which were sequenced (identified by asterisks in Figure 6). Sequence data clearly indicate the presence of alternatively spliced mRNAs in both 3rd instar larvae and S2 cells. The primers spanning the ol region detected a mRNA lacking exons 6 and 7 diagrammed in Figure 7. Recall that 6 of the 7 dpolv nonsense mutations are in these two exons. At least two differently spliced messages were obtained using the primers spanning the lv exons, one is missing exons 35 to 50 where all of the dpolv mutations are located, whereas the other skips a larger number of exons, 35 to 69. The largest band in each is consistent with inclusion of all exons between the selected primers and illustrates we can amplify at least 10kb by RT-PCR. The intermediately sized bands most likely represent different alternatively spliced products with various exons included in the transcripts.

The Flybase Drosophila EST database (http://flybase.org) for dumpy is highly enriched for clones with sequences from the 3’ end of the gene and essentially is non informative with regard to alternative splicing patterns. There is one EST from the database, however, that excludes exons 53 to 69 (see Figure 7). Recall that the dpolvDG82 nonsense mutation is in exon 58, and according to our hypothesis this exon should be alternatively spliced.

![Figure 6. Gel showing RT-PCR products from the dumpy gene in 3rd instar larvae and S2 cells. Primers flanking the ol and lv regions and the ZP domain that were used to generate these products are shown in Table S2. Bands marked with asterisks were excised and sequenced. doi:10.1371/journal.pone.0012319.g006](http://www.plosone.org/figure6)

![Figure 7. Diagram of RT-PCR products exhibiting alternative splicing in the dumpy gene. The middle line shows the intron-exon structure of the wild type dumpy gene. The ol region is shown above the line and the RT-PCR product which is missing exons 6 and 7. Below the line depicting the wild type gene is the intron-exon structure of the 3’ end of the gene showing the lv and l regions and the ZP domain. The RT-PCR products are missing a number of exons from each region. In each case, arrowheads mark the positions of primers used to obtain the RT-PCR products. These primer sequences are shown in Table S2. doi:10.1371/journal.pone.0012319.g007](http://www.plosone.org/figure7)
Discussion

Our attempts to generate new dumpy mutants in an isoallelic 2nd chromosome with ethyl methanesulfonate produced a distribution of dumpy alleles similar to that of Jenkins [13] Table 5. Thus the majority of our mutants generated in crosses 1 and 2 (see Table S1) were dpy\textsuperscript{olv}, dpy\textsuperscript{v1}, and dpy\textsuperscript{v2} in decreasing order. Other chemical mutagens – see Table 6 in Jenkins [13] – produce similar distributions. All of these cross schemes involved screening F1s heterozygous for dpy\textsuperscript{u} for oblique wing phenotypes and/or thoracic vortices. It’s possible that the greater ease with which oblique versus vortex phenotypes are detected could bias such screens utilizing dpy\textsuperscript{u} toward the recovery of dpy\textsuperscript{olv}, dpy\textsuperscript{v1}, and dpy\textsuperscript{v2} mutants, yet the latter two types are recovered very infrequently. Hence we feel the distribution of mutants accurately reflects the sizes of “targets” within the dumpy gene which, when mutant, affect one or more combinations of phenotypes. We are collaborating with the laboratory of Olga Grushko and Alexey Kondrashov at the University of Michigan where spontaneous dumpy mutants are being isolated as non-fliers at 28°C. To date, eight such mutations have been analyzed molecularly (see Table 3).

As with the EMS induced mutants, lethal classes predominate and in this case they are either dpy\textsuperscript{u} (1), dpy\textsuperscript{olv} (2) or dpy\textsuperscript{v} (5) mutants. Except for two mutants, dpy\textsuperscript{L2511B} and dpy\textsuperscript{V1D191A}, all of the others are deletions or small inversions which create frame shifts and/or stop codons. There is also a deletion in an intron with an unknown effect.

Due to our failure to recover dpy\textsuperscript{u} mutants in the screens utilizing crosses 1 and 2 in Table 2 and because only two such alleles currently exist in stock centers, dpy\textsuperscript{olv} and dpy\textsuperscript{v2}, we set up crosses 3 and 4 to enrich for new vortex mutants in the isoallelic 2nd chromosome from the Australia line. Following Jenkins [13], we estimated that we would obtain one transmitted mutant vortex in 17,500 F1 flies screened (frequency of F1 dumpy mutants, ca 0.8% in his Table 4, times 1/20 vortex mosaics or completes in his Table 5 times 15% of transmitted vortex mutations in his Table 6).

The number of F1s screened in cross 3 was not estimated, but in cross 4, in two separate mutageneses, we examined an estimated 24,000 flies. Five F1 flies with mosaic or complete vortices transmitted the mutation but none was a dpy\textsuperscript{olv} mutant. Of the five mutants, four were complementing dpy\textsuperscript{olv} alleles and one was a homozygous viable dpy\textsuperscript{v2} allele. As mentioned above, there are no stocks of the dpy\textsuperscript{u} mutants obtained by Jenkins, or Grace [17]. Our failure to obtain such mutants with EMS makes it more likely, in our view, that the roo element in the 5’ end of the dumpy gene in mutants dpy\textsuperscript{olv} and dpy\textsuperscript{v2} is responsible for the mutant phenotype. It may also be that these alleles are, in fact, the same—the roo element is inserted between the same two base pairs in each case and were perhaps inadvertently isolated at different times in the Morgan laboratory and named as separate alleles. Clearly the vortex sublocus, so elegantly mapped by Grace [17] at the 5’ end of the dumpy gene itself, needs to be better defined mutationally. The dumpy lethal sublocus, currently defined by only two alleles dpy\textsuperscript{D2701} and dpy\textsuperscript{D2705}, also needs to be further analyzed. It is not clear, however, how additional dpy\textsuperscript{u} alleles can be obtained, since F1 screens cannot be used. F2 screens would be tedious, although it might be possible to screen the progeny of F2 individuals crossed with dpy\textsuperscript{olv} flies for homozygous lethality.

The approach we have taken to define the mutations generated with EMS in isoallelic dumpy wild type alleles, i.e. producing overlapping amplicons from the entire locus and screening for base pair mismatches in reannealed duplexes with Surveyor nuclease and dHPLC as shown in Figure 3, has been very effective. The data are remarkable in that most of the mutations, including dpy\textsuperscript{olv} and dpy\textsuperscript{v2} mutants, result in stop codons either at the site of the mutations or are generated from a deletion, or cause the removal of a splice site. It is also clear that all eleven missense mutations change cysteines in EGF or DPy motifs, presumably altering or destabilizing their tertiary structures. Given the repetitive nature of the Dumpy protein, i.e. all the EGF-DPY-EGF motifs, perhaps most other kinds of missense mutations don’t produce a viable phenotype. As shown in Table 3, eight of the eleven missense mutations are found in exon 11 which defines the ov sublocus in Grace’s map. This is in stark contrast to the mutational spectrum in probably all other dumpy exons where the mutations are almost exclusively deletions or nonsense mutations. Grace mapped three kinds of oblique mutations viz. dpy\textsuperscript{olv}, dpy\textsuperscript{v2}, and dpy\textsuperscript{olv} at the ov sublocus, and we also find both dpy\textsuperscript{olv} and dpy\textsuperscript{v2} mutations in exon 11. The exon is remarkable in that it encodes 4 simple EGF-DPY diads, which, although most of the protein consists of EGF-DPY-EGF triads (see Figure 2), are also found at other positions in the protein. Six of the missense mutants are in EGF modules and two are in the DPY members of the diads. Interestingly, in the EGF module in the second diad, there are four mutations affecting four of the six cysteines. Two of the mutations exhibit a dpy\textsuperscript{olv} phenotype and two are dpy\textsuperscript{v2} mutants, one of which, dpy\textsuperscript{101A}, is a complementing mutant (see below).

At this point we don’t know if the dumpy exon 11 is alternative or constitutive since our RT-PCR experiments did not utilize primers flanking this exon. The extreme conservation found at its 5’ end (see Figure 5) indicates it is alternatively spliced, but the presence of a dpy\textsuperscript{olv} nonsense mutation (dpy\textsuperscript{olv101A}) according to our hypothesis, would make exon 11 constitutive. We are currently analyzing the splice variants in dumpy RNA by RNA-seq. [33,34], the results of which should clarify the status of exon 11.

As mentioned above, we believe the long standing and hitherto unexplained complexity of the dumpy gene can best be explained by extensive alternative splicing where dpy\textsuperscript{olv} nonsense mutations tag constitutive exons presumably located at several different places in the gene. Nonsense mutations with partial dumpy function, e.g. dpy\textsuperscript{olv105A} and dpy\textsuperscript{olv106A}, will be found in alternatively spliced exons and should be more localized in the gene.

In this regard, typical characteristics of alternatively spliced exons are small size and divisibility by 3 so as not to affect the reading frame depending on their inclusion or exclusion. The dumpy gene contains 78 coding exons, many are very small, i.e. under 80bp, and the number of nucleotides in all but 1 internal exon is divisible by 3.

We predict and, indeed, have found that alternative splicing produces tissue specific isoforms of Dumpy encoded by at least several kinds of mRNAs. The dpy\textsuperscript{olv}, dpy\textsuperscript{v2}, and dpy\textsuperscript{olv} nonsense mutations either could result in truncated isoforms only in the affected tissues or, if they are located in splicing protein binding sites in the exon, prevent the formation of the alternatively spliced transcript (see [35]). These two possibilities make different predictions about whether alternatively spliced transcripts would be found in dpy\textsuperscript{olv}, dpy\textsuperscript{v2}, and dpy\textsuperscript{olv} mutants, viz. if the tissue specific mutation results in truncated isoforms, the alternatively spliced transcripts would be present in mutant flies but if the mutations interfere with the splicing process, the transcripts would be aberrant or absent. Thus, a comparison of RT-PCR products in tissues from wild type and mutant flies should distinguish between the two mechanisms.

How else can we explain the observation that most of the lethal classes of dumpy alleles, viz. dpy\textsuperscript{olv}, dpy\textsuperscript{v2}, and dpy\textsuperscript{olv}, are due to nonsense mutations? One explanation might be that some premature stop codons in the dpy\textsuperscript{olv} mutants result in nonsense mediated decay of the mRNA [36] and this completely removes all functions in the olv class. However, a paradox in the results is that
many of the less severe $dp^{oa}$, $dp^{ov}$, and $dp^{b}$ mutations that retain some dumpy function also introduce premature stop codons that truncate the protein within its extracellular domain. Indeed, these data do not appear compatible with the hypothesis that the dumpy locus generates only one molecular product. Indeed, the RT-PCR results reported here (see Figures 6 and 7) clearly show that certain exons are excluded from some dumpy mRNAs by alternative splicing. Hence, we propose that dumpy generates multiple products by alternative splicing which are specialized to particular biological functions.

As mentioned above the mutant screens outlined in cross 4, Table 2, failed to generate new $dp^{v}$ (vortex) mutants but did, however, produce four mutant alleles, $dp^{olv105A}$, $dp^{olv104A}$, $dp^{olv105B}$ and $dp^{olv106B}$. These came through the screen because they complemented the lethal phenotype associated with a $dp^{b}$ allele. Earlier, crosses 1 and 2 produced additional complementing $dp^{ov}$ mutants, e.g. $dp^{olv105A}$, $dp^{olv104A}$, $dp^{olv27B}$, $dp^{olv106A}$, and $dp^{olv6}$. These were also recovered as complementers of the $dp^{b}$ allele carried on the In(2LR) CyO balancer. Also several inter se crosses have revealed additional cases of complementation between these individual $dp^{ov}$ mutant alleles. For example, $dp^{olv105A}$ and $dp^{olv6}$, mutants which flank the PIGSEAFE st exon (number 12) fully complement for viability (i.e. 1/3rd of the F1s from a cross between flies from balanced lethal parents survive to adulthood, but still show oblique wings and vortices. Such interallelic or intragenic complementation is generally explained by association and functional complementation between mutant protein subunits [37], but when one (e.g. in $dp^{olv6}$/$dp^{olv105A}$ heterozygotes or both alleles (e.g. in $dp^{olv6}$/$dp^{olv105B}$ heterozygotes) are nonsense mutations, a different explanation for the complementation must be found.

Complementing $dp^{ov}$ nonsense mutants, are also difficult to explain by cis alternative splicing. Alternative trans-splicing, however, could be operating in the processing of dumpy messages, perhaps only in certain tissues, and provide an explanation for the viability of some $dp^{ov}$ heteroallelic heterozygotes. Paradigms for putative trans-splicing events have been documented [38-42].

Our current RNA-seq approach to detect exon-exon junctions in the dumpy “transcriptome,” when coupled with the identification of the mutant vs. wild type codons or SNP associations in the exons of the dumpy gene. If trans-splicing turns out to be responsible for the interallelic complementation between lethal dumpy nonsense mutants, we can begin to identify the mechanism and the gene products that are responsible for the splicing events. This can be accomplished with screens for suppressors and enhancers of the complementation patterns of different dumpy mutants.

In summary, the molecular analysis of 45 preexisting, spontaneous, and/or EMS induced dumpy mutations revealed most missense mutations were found in exon 11. All other mutations except two transposon insertions generated stop codons, were deletions, an inversion, or frameshift generated nonsense codons, even those which exhibited only one or two of the three dumpy mutant phenotypes. We present evolutionary and experimental evidence for cis alternative splicing of dumpy transcripts and argue that these observations, along with the distribution within the gene of nonsense mutations with different dumpy mutant phenotypes, makes it likely that alternative splicing underlies the genetic and phenotypic complexity of this long studied, paradigmatic Drosophila complex gene. In addition, complementation between certain dumpy nonsense mutant alleles can be explained by trans-splicing.

### Materials and Methods

**Drosophila strains**

Dumpy mutant alleles with undefined genetic backgrounds were ordered from Bloomington or Kyoto stock centers. These are identified in Table 3 as SC. Mutants 12 and 16 were provided by Jim Fristrom and Mary Prout, and identified as MP in Table 3. They were generated by gamma ray mutagenesis and are $dp^{olv}$ or $dp^{ov}$ alleles which, when homozygous in somatic clones, give rise to wing blisters (see Prout et al. [25]). Mutants designated with an OG in Table 3 were recovered by Olga Grushko and Alexey Kondrashov in a screen for spontaneous dumpy mutants. They are present in two 2nd chromosomes designated as A and B, extracted from natural populations near Ann Arbor, MI, and made homozygous for chromosome 2. Each has a different pattern of single nucleotide polymorphisms (SNPs) in the dumpy gene. The screen selects for non fliers and, hence, the mutant alleles have an oblique wing or vortex phenotype. Most of the dumpy mutants we analyzed were generated by EMS following Jenkins [13] in stocks isolealic for dumpy. Those labeled with a BM in Table 3 are in an isolealic 2nd chromosome carrying cn and bse mutant alleles and generated by Brad Marshall, then an undergraduate researcher in the laboratory. Crosses 2A and 2B were performed and mutants were recovered by Michael Guertin are labeled MG in Table 3. Ross Macintyre produced and recovered dumpy mutants from crosses 3 and 4 and are labeled RM in Table 3. In crosses 2, 3, and 4, the 2nd chromosome was marked with either net (II, O.O) or clot (II, 16.O). These chromosomes were recovered as recombinants from net $dp^{+}$ clot / net $dp^{+}$ clot females where the net $dp^{+}$ clot chromosome had been isolated from a wild population from Australia, provided by Chip Aquadro, and made isolealic with crosses to appropriate 2nd chromosome balancer stocks. We confirmed the identity of the dumpy alleles in the net and clot stocks by a Southern blot which allowed us to analyze the PIGSEAFE repeat number, thus confirming the $dp^{b}$ allele came from the chromosome from the Australia line (see Carmon et al. [21]). Second, as described below, when ampiclons spanning the entire dumpy gene from the two strains, net $dp^{b}$ and $dp^{clot}$, were denatured, reannealed, treated with the Surveyor nuclease, no base pair mismatches were detected, confirming the sequence identity of the wild type dumpy alleles.

**Mutagenesis and mutant allele recovery**

For the dumpy mutant alleles induced with EMS in the cn bal, net or clot chromosomes, we fed males 0.024M EMS for 24 hours following Lewis & Bacher [43] and Jenkins [13]. F1 males were then scored for mosaic or complete dumpy mutant phenotypes. Several different crosses were carried out as shown in Table 2.

In crosses 2A and 2B, mutant F1 males were crossed back to $dp^{olv1}$ cl; e or net $dp^{olv1}$ cl females respectively to determine which mutants transmitted a mutant allele, i.e. had a partially or completely mutant gonad. The mating of the phenotypically dumpy F1 flies indicated that only 29% transmitted the new mutation. Previous studies [13] found that 35% of the F1 dumpy mutants transmitted the mutant allele to their offspring. Mutant F2 males were then mated to In(2LR) CyO, $dp^{b}$ clot -4 [44]/ In(2LR) Gla females, and the curly winged progeny assessed for their eye color, i.e. either clot or wild type. In cross 2A, when net $cl^{+}$ males were mutagenized, surviving curly sibs with wild type eyes were mated to establish a stock of the new dumpy mutant. In most cases, the newly induced mutant was lethal towards the CyO, $dp^{b}$ clot -4 chromosome. In these instances, 5–10 single Gla/dumpy females were then backcrossed to CyO, $dp^{b}$ clot -4/Gla females. Glazed eyed flies from vials with no curly winged clot eyed progeny were then
mated to establish the new mutant strain. In cross 2B, when dp\textsuperscript{u} clot males were mutagenized, F1 males were crossed to dp\textsuperscript{ov1} cn bw females. If the new mutant transmitted, mutant detected, it was cloned using a TOPO Zero Blunt Cloning Kit. An example of a dHPLC analysis of a cleaved amplicon is shown which were detected as smaller, discrete peaks following dHPLC. If one of the reannealed DNAs in the mutant pool did not give any mismatched base pairs in any of the initial 85 sizing gradient. DNA from the original unmutated isoallelic stocks reannealed, treated with the Surveyor Mutation Detection Kit followed by 74 cycles of 30 cycles of 94°C for 2 min and 35 cycles of 98°C for 5 sec, 63°C for 15 sec, 72°C for 7.5 min, followed by 72°C for 10 min. After visualization on a gel showing the presence of a long insertion, its ends were sequenced to identify the transposon.

**v2 and ov1**

Since the progenitor chromosome for these mutants was not known, DNA from homozygous individuals was extracted. Primer pairs not giving a product were then used for amplification with iProof Polymerase (Bio-Rad) and GC buffer. Cycling conditions were 98°C for 2 min and 35 cycles of 98°C for 5 sec, 63°C for 15 sec, 72°C for 7.5 min, followed by 72°C for 10 min. After visualization on a gel showing the presence of a long insertion, its ends were sequenced to identify the transposon.

**o2 and ovDG2**

DNA from homozygous individuals for dp\textsuperscript{o2} and dp\textsuperscript{ovDG2} F1s since both mutants were made by Dale Grace [15,17]. Preliminary data indicated the two mutants were induced in the same progenitor chromosome and dp\textsuperscript{ovDG2} was subsequently analyzed as above using Surveyor and WAVE analysis.

**RT-PCR**

Total RNA was extracted from 3rd instar larvae and S2 cells. RT-PCR was performed using the SuperScript III One-Step RT-PCR System with Platinum Taq High Fidelity (Invitrogen). Cycling conditions were 55°C for 30 min, 94°C for 2 min, 40 cycles of 94°C for 15 sec, 55°C for 30 sec, 72°C for 8 min, followed by 72°C for 10 min. Primer pairs spanning putative alternatively spliced exons in various dumpy subloci were used as well as a pair spanning the ZP domain. Amplified products were separated in gels and bands excised and sequenced.

**Computer based analyses**

Evolutionary comparisons of the 30 nucleotides at the 5’ ends of dumpy exons were made in the dumpy genes from 7 Drosophila species – D. melanogaster, D. ananassae, D. pseudoobscura, D. willistoni, D. mojavensis, D. virilis, and D. grimshawi whose genomes have been sequenced. Analysis of 30 nucleotides was chosen due to the extremely variable length of dumpy exons from 54bp to over 13kb. To do this, we aligned the sequences in MEGA 4 and used the nucleotide model to calculate the average pairwise distance with the Kimura 2-parameter correction [46].

**Supporting Information**

Table S1 Properties of dumpy mutants derived from crosses 1, 2A, 2B, 3, and 4 shown in Table 2 - *indicates complementation with at least one other lethal allele. **oblique score according to Grace [17] in parentheses. The oblique phenotype was tested over

Table S2 Primers used in this study - *Primers used for RT-PCR products shown in Figures 6 and 7

Found at: doi:10.1371/journal.pone.0012319.s001 (0.20 MB DOC)

Found at: doi:10.1371/journal.pone.0012319.s002 (0.26 MB DOC)
Acknowledgments

We thank Chip Aquadro, Jim Fristrom, and Mary Prout for fly stocks. We thank Alexey Kondrashov for designing the flightless mutant screen. We also thank former undergraduate student Alex Helkin for help with defining the vortex sublocus. Lauren MacIntyre Ampel provided excellent assistance on the preparation of this manuscript.

Author Contributions

Conceived and designed the experiments: AC OG RM. Performed the experiments: AC MJG OG BM RM. Analyzed the data: AC MJG OG. Contributed reagents/materials/analysis tools: AC. Wrote the paper: AC MJG OG RM.

References

1. Morgan TH (1911) The origin of nine wing mutations in drosophila. Science 33: 496–499.
2. Morgan TH, Sturtevant AH, Muller HJ, Bridges CB (1915) The Mechanism of Mendelian Heredity. New York: Henry Holt and Company. 262 p.
3. Altenburg E (1920) The genetic basis of truncate wing. - An inconstant and modifiable character in Drosophila. Genetics 5: 1–59.
4. Bridges CB, Mohr OL (1919) The inheritance of the mutant character “vortex”. Genetics 4: 283–306.
5. Morgan TH (1929) Data relating to six mutants of Drosophila. In: Sturtevant AH, Bridges CB, Morgan TH, Li JC, eds. Contributions to the genetics of Drosophila simulans and Drosophila melanogaster. Washington: Carnegie Institution of Washington. pp 171–199.
6. Muller HJ (1922) Variation due to change in the individual gene. Am Nat 56: 32–36.
7. Mohr OL (1923) Modifications of the sex-ratio through a sex-linked semi-lethal in Drosophila melanogaster (Besides notes on an autosomal section deficiency). Studia Mendeliana, ad centesimum diem natalem Gregorii Mendelii a gratia patria celebrandum, adhucvante ministerio Pragenti edita Brunae. pp 266–267.
8. Mohr OL (1929) Exaggeration and Inhibition Phenomena encountered in the analysis of an autosomal dominant Zeitschrift für induktive Abstammungs- und Vererbungslehre 50: 113–200.
9. Carlson EA (1959) Allelism, Complementation, and Pseudallelism at the dumpy locus in Drosophila melanogaster. Genetics 44: 347–373.
10. Carlson EA (1959) Comparative Genetics of Complex Loci. Q Rev Biol 34: 23–50.
11. Carlson EA, Southin JL (1962) Comparative Mutagenesis of Dumpy Locus in Drosophila melanogaster. 1. X-Ray Treatment of Mature Sperm - Frequency and Distribution. Genetics 47: 321–336.
12. Southin JL (1966) An analysis of eight classes of somatic and gonadal mutation at dumpy locus in Drosophila melanogaster. Mutat Res 3: 54–65.
13. Jenkins JB (1967) Mutagenesis at a complex locus in Drosophila with the monofunctional alkylating agent, ethyl methanesulfonate. Genetics 57: 783–789.
14. Schede R (1967) Rare Pseudallelic Crossover between 2 Phenotypically Identical Alleles at a Restricted Sublocus of Dumpy in Drosophila Melanogaster. Nature 216: 1348–1349.
15. Grace D (1970) Genetic analysis of dumpy Region in Drosophila - Its multigenic composition. Mutat Res 10: 489–496.
16. Morgenstern DW (1974) Recombination and mutation analysis of lethals at the dumpy locus in Drosophila melanogaster. Montreal, Canada: McGill University. 168 p.
17. Grace D (1980) Genetic analysis of the dumpy complex locus in Drosophila melanogaster - Complementation, Fine-Structure and Function. Genetics 94: 637–662.
18. Segraves WA, Louis C, Tsutoba S, Schell P, Ravds JM, et al. (1984) The rudimentary locus of Drosophila melanogaster. J Mol Biol 173: 1–17.
19. Jack J, Delosco Y (1995) Structure and regulation of a complex locus - the cut gene of Drosophila. Genetics 139: 1689–1700.
20. Wilkin MB, Becker MN, Mulvey D, Phan I, Chao A, et al. (2000) Drosophila Dumpy is a gigantic extracellular protein required to maintain tension at epidermal-cuticle attachment sites. Curr Biol 10: 559–567.
21. Carmon A, Wilkin M, Hassan J, Baron M, MacIntyre R (2007) Concerted evolution within the Drosophila dumpy gene. Genetics 176: 309–325.
22. Carmon A, Larson M, Wayne M, MacIntyre R (2010) The rate of unequal crossing over in the dumpy gene from Drosophila melanogaster. J Mol Evol 70: 269–283.
23. Prout M, Diamania Z, Soong J, Fristrom D, Fristrom JW (1997) Autosomal mutations affecting adhesion between wing surfaces in Drosophila melanogaster. Genetics 146: 273–285.
24. Bohel C, Prokop A, Brown NH (2005) Papillate and Papio: Drosophila ZP-domain proteins required for cell adhesion to the apical extracellular matrix and microtubule organization. J Cell Sci 118: 633–642.
25. Jazwinska A, Kibara C, Affolter M (2003) Epithelial tube morphogenesis during Drosophila tracheal development requires Papio, a luminal ZP protein. Nat Cell Biol 5: 895–901.
26. Mahoney MB, Parks AL, Rudhy GA, Tiong SYK, Easrig H, et al. (2006) Presenilin-based genetic screens in Drosophila melanogaster identify novel notch pathway modifiers. Genetics 172: 2309–2324.
27. Carmon A, Topbar F, Baron M, MacIntyre R (2010) dumpy interacts with a large number of genes in the developing wing of Drosophila melanogaster. Fly 4: 117–127.
28. Xing Y, Lee C (2005) Evidence of functional selection pressure for alternative splicing events that accelerate evolution of protein substructures. P Natl Acad Sci USA 102: 13526–13531.
29. Xing Y, Lee C (2006) Alternative splicing and RNA selection pressure - evolutionary consequences for eukaryotic genomes. Nat Rev Genet 7: 499–509.
30. Lu H, Liu L, Sato S, Xing Y, Lee C (2009) Predicting Functional Alternative Splicing by Measuring RNA Selection Pressure from Multigencode Alignments. Plos Comput Biol 5: 21–26.
31. Graveley BR (2005) Mutually exclusive splicing of the insect Drosa Pre-mRNA directed by competing intronic RNA secondary structures. Cell 123: 65–73.
32. Olson S, Blanchette M, Park J, Savva Y, Yeo GW, et al. (2007) A regulator of Drosa mutually exclusive splicing fidelity. Nat Struct Mol Biol 14: 1134–1140.
33. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18: 1509–1517.
34. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57–63.
35. Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nat Rev Genet 3: 285–298.
36. Behm-Ansandt I, Izaurralde E (2006) Quality control of gene expression: a stepwise assembly pathway for the surveillance comp ex that triggers nonsense-mediating mRNA decay. Gene Dev 20: 391–398.
37. Fuchsan JRS (1966) Splicing events that accelerate evolution of protein subsequences. P Natl Acad Sci USA 102: 13526–13531.
38. Krauss V, Dorn R (2004) Evolution of the trans-splicing Drosophila locus mod(mdg4) in several species of Diptera and Lepidoptera. Gene 331: 165–176.
39. Horlchi T, Gniéger E, Aigaki T (2003) Alternative trans-splicing of constant and variable exons of a Drosophila axon guidance gene, lola. J Dev Biol 17: 2496–2501.
40. Rowley JD, Blumenthal T (2008) The cart before the horse. Science 321: 1302–1304.
41. Li H, Wang JL, Mor G, Sklar J (2008) A neoplastic gene fusion mimics trans-splicing of RNAs in normal human cells. Science 321: 1357–1361.
42. Lewis E, Bacher F (1969) Method of feeding ethylmethanesulfonate (EMS) to Drosophila males. IRS 43: 191.
43. Davis MB, Macintyre RJ (1988) A genetic analysis of the alpha-Glycerophosphate Oxidase locus in Drosophila melanogaster. Genetics 120: 755–766.
44. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biostatisticians. In: S K, S M, eds. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa NJ, USA: Humana Press.
45. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.