bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1

2

3

4

5 # Transcriptomic and functional analysis of the
6 # oosome, a unique form of germ plasm in the
7 # wasp *Nasonia vitripennis*

8

9 Honghu Quan and Jeremy Lynch*

10 University of Illinois at Chicago, Department of Biological Sciences, Chicago IL 60607

11 USA

12 *Author for correspondence

13 Honghu Quan email: hquan2@uic.edu

14 Jeremy Lynch email: jlynch42@uic.edu

15

16

17

18

19

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

20 # Abstract

21 *Background*

22 The oosome is the germline determinant in the wasp *Nasonia vitripennis* and is homologous to
23 the polar granules of *Drosophila*. Despite a common evolutionary origin and developmental role,
24 the oosome is morphologically quite distinct from polar granules. It is a solid sphere that
25 migrates within the cytoplasm before budding out and forming pole cells.

26 *Results*

27 To gain an understanding of both the molecular basis of the novel form of the oosome, and the
28 conserved essential features of germ plasm, we quantified and compared transcript levels
29 between embryo fragments that contained the oosome, and those that did not. The identity of the
30 localized transcripts indicated that *Nasonia* uses different molecules to carry out conserved germ
31 plasm functions. In addition, functional testing of a sample of localized transcripts revealed
32 potentially novel mechanisms of ribonucleoprotein assembly and pole cell cellularization in the
33 wasp.

34 *Conclusions*

35 Our results demonstrate that numerous novel and unexpected molecules have been recruited in
36 order produce the unique characteristics of the oosome and pole cell formation in *Nasonia*. This
37 work will serve as the basis for further investigation into the patterns of germline determinant
38 evolution among insects, the molecular basis of extreme morphology of ribonucleoproteins, and
39 the incorporation of novel components into developmental networks.

40

41 ***Keywords***

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

42 Oosome, Germ plasm, Primordial Germ Cells, Germline, Nasonia, Drosophila, RNAseq, RNA
43 Interference

44

45 Background

46 Germline establishment is a crucial event for sexually reproducing organisms. Germline cells are
47 special in that they are able to generate all of the cell fates of the soma and to regenerate
48 themselves. There are two major strategies to specify the germline among animals: zygotic
49 induction and maternal provision. In zygotic induction, inductive signals from surrounding
50 tissues drive the establishment of germline fate, usually relatively late in embryogenesis, after
51 the transition from maternal to zygotic control [1]. In contrast, in the maternal provision mode,
52 the germ cells are specified by determinants called germ plasm that are synthesized and localized
53 during oogenesis and are the first group cells formed during embryogenesis. Germline
54 specification in this mode occurs very early in development, usually prior to the activation of the
55 zygotic genome [1, 2]. Classical experiments have shown that germ plasm is both necessary and
56 sufficient to establish the germline fate [3-5].

57 It is likely that the maternal provision mode of germ plasm evolved multiple times among the
58 animals, and this reflected in the molecular basis of germ plasm determinants [6-27]. In
59 vertebrates, germ plasm (where it exists) is dependent on the maternal localization of *bucky ball*
60 [28]. In insects, the gene products of *oskar* (*osk*) are both necessary and sufficient to induce germ
61 plasm, and thus, primordial germ cells (PGCs) [29-32]. Downstream of these nucleators is a suite
62 of highly conserved germline-associated molecules (i.e., Vasa (Vas), Nanos (Nos), Tudor (Tud),

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

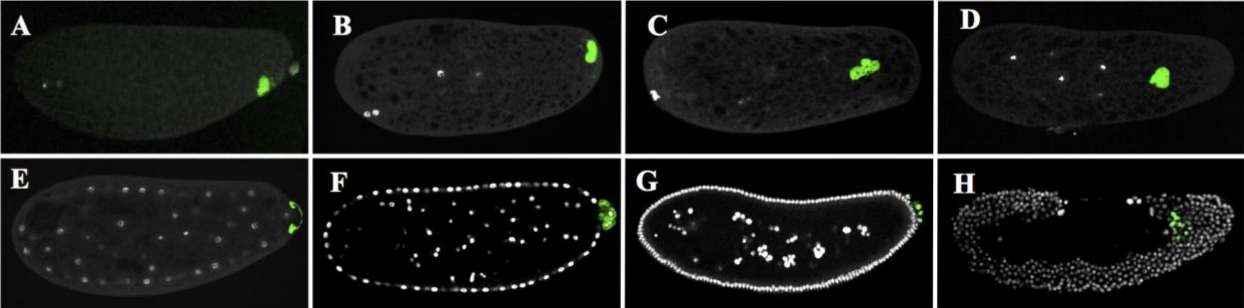
63 etc.) that are recruited to the germ plasm where they presumably carry out functions ancestral for
64 the animal germline [30, 33-35].

65 There are several conserved properties of PGCs that must be conferred to naive cells by
66 germ plasm. One of these is a period of transcriptional quiescence germ cells undergo after
67 being specified. This feature makes the composition of a germ plasm even more important since
68 it means that mRNAs and proteins critical for at least the early germ cell functions need to be
69 provided in the germ plasm itself. Germ cells also usually become highly migratory later in
70 development as they seek to colonize the developing gonad. They are often highly enriched for
71 mitochondria and have specific metabolic needs [36]. Since they carry the genome that will be
72 passed to future generations, germ cells have enhanced mechanisms to prevent DNA damage and
73 to reduce the activity of transposable elements [37, 38]. Finally, they must have the capability to
74 induce pluripotency to their genome, as germline cells will be the source for all cell fates in the
75 eventual progeny.

76 Like all developmental processes, there is likely to be variation in the details of these
77 conserved features of germline determination, whether due to selective or neutral forces.
78 Pressures that could impact the composition of the germ plasm could be differential activity of
79 transposable elements in the genome, or a novel path for migration of the PGCs to the gonad, for
80 example. Novelties in embryogenesis may also drive germ plasm composition. For example, in
81 the parthogenetic, paedomorphic embryos of the midge *Miastor,* somatic nuclei undergo
82 significant chromosomal diminution in the early cleavages while the cell that takes up germ
83 plasm maintains the full complement of chromosomes, indicating that the germ plasm contains a
84 component that prevents the loss of chromosomes during the early cleavages [39].

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

85 Holometabolous insects are an ideal system with which to study how germ plasm evolves,
86 given that it is ancestral in this clade [35], the unparalleled levels of diversity, and the strong
87 baseline understanding germ plasm function derived from *Drosophila melanogaster*. Here we
88 focus on the parasitic wasp *Nasonia vitripennis* as a model to compare to the fruit fly. Like
89 *Drosophila*, *Nasonia* depends on Osk, Vas, and Tud to assemble germ plasm [35, 40]. However,
90 in contrast to the collection of small granules stably associated with the posterior pole that make
91 up the *Drosophila* germ plasm, the *Nasonia* germ plasm forms a very large, dense organelle
92 called the oosome (Figure 1). This highly divergent morphology strongly indicates that the
93 composition of the *Nasonia* oosome may diverge significantly from the polar granules of
94 *Drosophila*.
95



96

97 **Fig. 1. The complete process of *Nasonia* germline development marked by the novel gene *Nv-bark*.** A-D: the green shows the oosome shapes and localizations in different nuclear divisions during pre-blastoderm stage. E: the oosome flattenes into the bud (green) at the beginning of blastoderm stage. F: the first batch of pole cells (green) is formed from the bud during early syncytial blastoderm stage. G: the
98 pole cells divide several times to make more and smaller cells (green) during cellularized blastoderm stage. H: the pole cells (green) has entered into the embryo and migrate to the gonads during gastrulation stage. The embryos are arranged according to their embryogenesis stages with the posterior side to right and dorsal side on the top. DAPI (white) marks the nuclei.
99

100 The behaviors of the oosome and the PGCs in *Nasonia* further imply a divergent
101 composition of the oosome. In freshly laid eggs, the oosome is tightly bound to the ventral-
102 posterior cortex of the embryo (Fig. 1A). When the zygotic nucleus forms and moves into the
103 interior of the embryo, the oosome detaches from the cortex and coalesces into a dense,
104 extremely large structure in the same central column as the nuclei. It migrates anteriorly,

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

105 sometimes being found near to 50% egg length, before migrating back to the posterior pole (Fig.
106 1B-D). As the cleavage nuclei migrate toward the cortex, the oosome flattens into a crescent on
107 the posterior pole of the embryo while a large bud protrudes from the pole (Fig. 1E). Typically,
108 two or three nuclei become associated with the bud and the oosome material. The bud pinches
109 off, and the nuclei rapidly individuate into pole cells (Fig. 1F,G).

110 This is distinct from pole cell formation in *Drosophila,* where each pole cell forms
111 individually upon association with a certain critical amount of pole plasm [41]. After dividing a
112 few times, the pole cells remain stable at the posterior pole of the egg until gastrulation, when
113 they migrate through the posterior epithelium and coalesce into to two masses, presumably
114 where the primordial gonads are developing (Fig. 1H). This migration is distinct from
115 *Drosophila* pole cells which remain external on the posterior pole as the germ band undergoes
116 massive extension well after gastrulation and migrate internally through the gut wall [41].

117 Thus, it is clear that *Nasonia* and *Drosophila* share some fundamental aspects of germline
118 establishment, but they also have their own diverged features that fit in their own embryogenesis
119 program. This raises the question of which genes are the core components for the maternal
120 provision mode and which genes contribute to their own distinct features in germline
121 development.

122 To address these questions, we compared the mRNA content of anterior and posterior
123 halves of the pre-blastoderm stage *Nasonia* embryos in an effort to identify the components
124 specifically localized to the oosome. We found only a few of mRNAs conserved in both fly polar
125 granules and the *Nasonia* oosome, such as *osk*, *nos*, and *ovo* (see Supplemental File 1). The rest
126 are all novel components germ plasm that either lack *Drosophila* homologs or have homologs in
127 *Drosophila* that do not play any roles in germline development. We performed functional studies

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

128 for a set of localized transcripts, all of which showed roles in the unique features of *Nasonia*
129 germline development, demonstrating the value of our approach to identify the molecular
130 sources of novelty among various insect lineages.
131

132 Results

133 RNA-seq analyses of the anterior and posterior poles of the wasp *Nasonia* early embryos

134 To identify the maternal transcripts in the oosome, we isolated the total RNA separately from
135 anterior and posterior poles of the pre-blastoderm stage *Nasonia vitripennis* embryos. Six
136 samples that comprised three sets of each pole were sequenced and subjected to differential
137 expression analysis, which are described in more detail in the Methods.

138 Our first attempt yielded encouraging, but mixed results. Several novel posteriorly
139 enriched factors were identified, and we did recover some known posterior factors (*Nv-osk* and
140 *Nv-nos*). However, other known localized factors such as *Nv-cad* and *Nv-dpp* were not found to
141 be statistically significant. In addition, one of the posterior samples was of poor quality, and
142 could not be used in the differential expression analyses, reducing the power of our approach.
143 Since we wanted results as comprehensive as possible, we repeated the experiment using a few
144 adjustments (see Methods) and obtained much more robust results. While this second experiment
145 was being prepared, we moved forward with expression and functional analysis of novel
146 candidates identified in the first. Since a handful of transcripts with confirmed oosome
147 localization were found only in the first experiment, the remainder of the results will include
148 these transcripts along with those identified in the second experiment.

149 All known maternally localized molecules (*Nv-osk*, *nos*, *dpp*, *cad*, *otd1*, and *gt*) were
150 found with high significance, giving confidence that our experimental design and analysis

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

151 approaches were appropriate for our goal of discovering the mRNAs localized to the oosome.
152 Overall, we found 92 transcripts with apparent significant enrichment at the posterior pole. These
153 ranged in levels of enrichment from 1.4 to 55 times higher in the posterior fragments compared
154 to anterior fragments. Our analyses also uncovered anteriorly enriched mRNAs, of which there
155 were 89, with a range of fold enrichment from 1.4 to 10 times higher at the anterior.
156

157 **Novel transcripts localized in the anterior half of the *Nasonia* early embryos**

158 While the anterior factors are not the focus of this manuscript, some interesting observations
159 were made in examining a small sample of candidate mRNAs. Most transcripts are localized in
160 small domains at the anterior cortex and seem to extend toward the posterior in variable tendrils,
161 rather than being uniform or graded caps of anterior localization (Fig. 2A1-K1). A notable
162 exception is the transcript of *Nasonia* homolog of *mex-3,* an RNA binding protein known for
163 controlling translation of orthologs of the posterior patterning factor *caudal* in the nematode *C.*
164 *elegans* and the beetle *Tribolium* [24, 42]. *Nv-mex3* mRNA is localized in a broad domain
165 extending far toward the posterior of the embryo (Fig. 2L1). *Nv-mex3* expression is highly
166 dynamic and variable in both the blastoderm and post-gastrular stages (Fig. 2L2-L3, additional
167 images not shown).

168 Transient localization is the most common feature of the anteriorly localized transcripts.
169 Most are ubiquitous or absent by the time the early syncytial blastoderm forms (Fig. 2 A2-H2,
170 J1-K3, M1-M3), except for *LOC100313502* which persists into the blastoderm stage where it
171 forms a anterior cap (Fig. 2 I1-I3). Among the anteriorly enriched transcripts, many have
172 predicted functions that may be relevant to egg activation (e.g. four encoding ion channels), and
173 anterior-posterior polarity and patterning (e.g. transcription and translation factors, see

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	Pre-blastoderm	**Blastoderm**	**Gastrulation**
LOC103 315431	A1	A2	A3
LOC100 124035	B1	B2	B3
LOC1001 21321	C1	C2	C3
Nv-exdl2	D1	D2	D3
LOC100 680515	E1	E2	E3
Nasvi2EG0 06979	F1	F2	F3
LOC1001 19882	G1	G2	G3
LOC1079 82211	H1	H2	H3
LOC100 313502	I1	I2	I3
LOC100 120674	J1	J2	J3
LOC100 115729	K1	K2	K3
Nv-mex3	L1	L2	L3
Nv-ucth	M1	M2	M3

174

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

175
Fig. 2. Genes expressed anteriorly in the pre-blastoderm stage. All embryos are aligned and grouped into three columns (Pre-blastoderm, Blastoderm and Gastrulation) according to their embryogenesis stages, with posterior side to the right and dorsal side on the top.
176

177 Supplemental file 1). There are also a large number of transcripts with no clear homologs outside
178 of *Nasonia* see (Additional File 1). One of these (*LOC100119982*, Fig. 2G1-G3) is a member of
179 a novel family of ankyrin domain containing molecules that are specific to Chalcid wasps and
180 appear to have obtained a broad diversity of expression and potential function during *Nasonia*
181 development [43].

182 Finally, a handful of molecules localized at both poles were detected. These include the
183 known factors *Nv-otd1* [44] and *Nv-TollC* (JAL in preparation), along with *Nv-endoglucanase*
184 (Fig. 3E), *Nv-insulin-like growth factor (Nv-igf,* Fig. 3C*), Nv-ucth* (a ubiquitin carboxy terminus
185 hydrolase, Fig. 2M)). The fact that these were found despite lower apparent fold differences
186 between the two embryonic halves, further gave confidence that our analysis was robust enough
187 to detect even subtle germ plasm localization of the vast majority of mRNAs.

188

189 **General description of the novel transcripts localized in the posterior pole of the *Nasonia***
190 **early embryos**

191 From the two analyses, we identified 92 candidate transcripts that were statistically enriched at
192 the posterior half of the *Nasonia* embryos. We then isolated PCR products for 54 of these genes
193 and made probes to determine their expression patterns during early embryogenesis. We were
194 able to confirm 47 transcripts that are expressed posteriorly during pre-blastoderm stage. Of the
195 remaining transcripts, one of them was not successfully cloned, and the rest were successfully
196 cloned but did not have any specific, localized expression. Interpretation of these potential false
197 positives (as well as some false negatives) will be discussed later. Among the 28 transcripts with

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

198 expression in the posterior half of the embryo, we grouped them into three categories based on
199 their expression patterns: 1) genes expressed in oosome and with expression maintained in the
200 PGCs (Fig. 3), 2) mRNAs localized in the oosome but then degraded in the PGCs (Fig. 4), and 3)
201 mRNAs strongly enriched in the posterior region of the embryo, but not incorporated in the
202 oosome (Fig. 5).

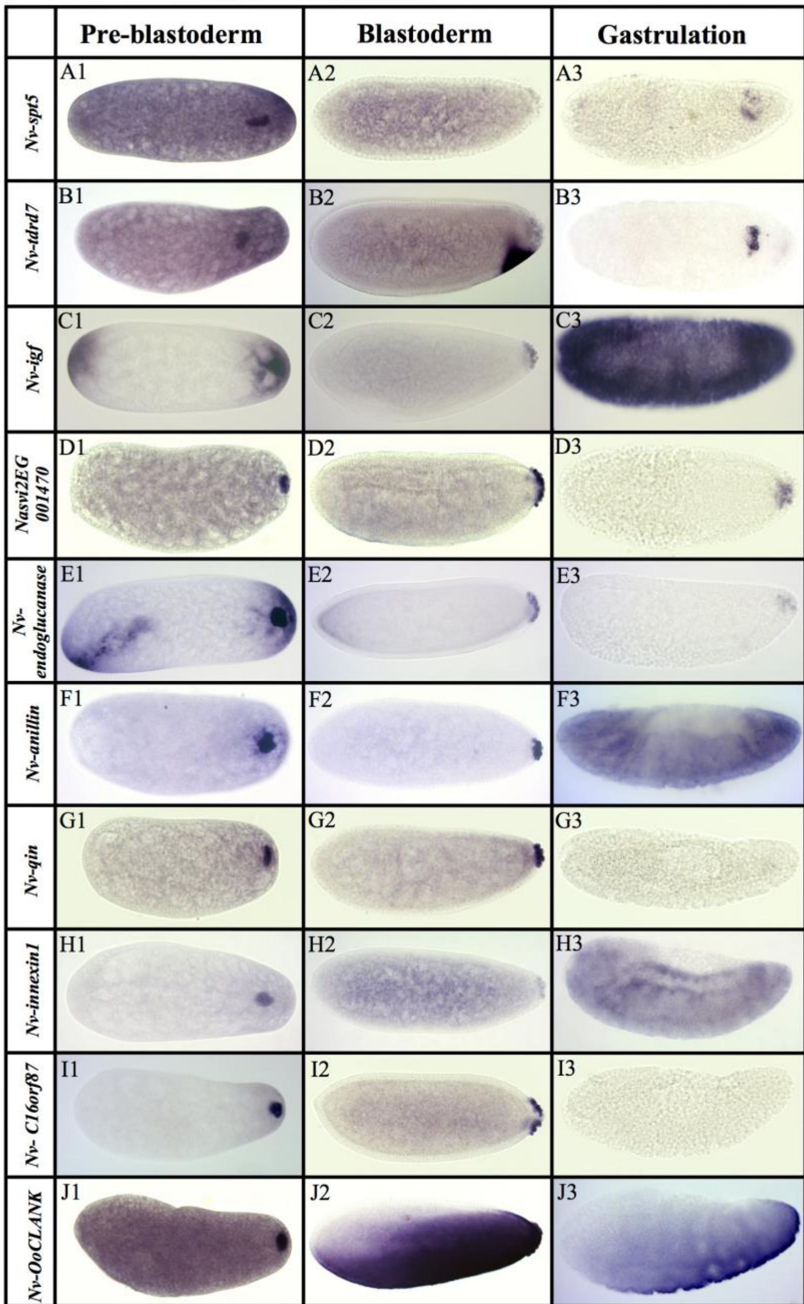
203 **Highly germline-associated transcripts**

204 We identified 11 transcripts that are localized to the oosome and maintained in the PGCs at pole
205 cell formation and beyond. We consider these to be the strongest candidates for having important
206 roles in the specification and function of the PGCs but cannot exclude *a priori* that they have
207 other (or no) important functions. Six transcripts are maintained throughout the stages of
208 germline development followed in this manuscript: oosome, pole cells, and migrating germ cells.
209 They include *Nasonia* homologs of the fly genes *bark beetle* (*Nv-bark*) (Fig. 1), *spt5* (*Nv-spt5*),
210 *tejas* (*Nv-tdrd7*), *insulin-like growth factor* (*Nv-igf*), and two transcripts without fly homologs
211 (*Nasvi2EG001470* and *Nv-endoglucanase*) (Fig. 3A1-E3).

212 Among these, *Nv-bark* is the best germline marker, bearing strong and consistent
213 germline association throughout early development, including expression in the late embryonic
214 gonads (Fig. 1). In this respect, it is better than our previously favored marker, *Nv-nos*, which is
215 downregulated significantly toward the end of PGC migration [40]. Bark is a large
216 transmembrane protein and is not expressed in the *Drosophila* germline. Its only known role is in
217 stabilizing tricellular junction in epithelial cells during embryogenesis [45, 46]. It is not clear
218 how this function is relevant to germline function, and indicates a novel recruitment of this factor
219 in the wasp.

220

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



221

222 **Fig. 3. Transcripts localized to the oosome that are subsequently maintained in pole cells.** All embryos are aligned and grouped into three columns (Pre-blastoderm, Blastoderm and Gastrulation) according to their embryogenesis stages, with posterior side to the right and dorsal side on the top.

223

224 *Nv-spt5* is significantly enriched in the oosome with low levels of ubiquitous expression
225 in the rest of the embryos in pre-blastoderm stage (Fig. 3A1). In the blastoderm and migrating
226 germ cell stages expression is reduced but still enriched in the pole cells while the ubiquitous

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

227 expression in the embryo persists (Fig. 3A2-A3). Spt5 homologs are involved in regulating RNA
228 polymerase progression during transcription [47], which might indicate that Nv-Spt5 is involved
229 in repressing or otherwise regulating the onset of transcription in the germ cells.

230 *Nv-tdrd7* is present at appreciable levels throughout the bulk cytoplasm and is also
231 strongly localized in the oosome (Fig. 3B1). This pattern is well reflected in the quantification of
232 mRNA levels in the two halves of the embryo, which show significant numbers of reads coming
233 from the anterior half of the embryo. At the blastoderm stage, *Nv-tdrd7* is moderately enriched in
234 the pole cells and is zygotically expressed in a ventral-posterior patch (Fig. 3B2), which was
235 detected in our earlier analysis of dorsal-ventral patterning [48]. After gastrulation, *Nv-tdrd7* is
236 strongly upregulated in a group of cells that are near to where the germ cells migrate, but it is not
237 clear if they are germ cells (Fig. 3B3).

238 *Nv-igf* is initially expressed in a bipolar pattern (similar to *Nv-otd1*[44]), before becoming
239 specific to the pole cells during the blastoderm stage and the migrating germ cells after
240 gastrulation (Fig. 3C1-C3).

241 *Nasvi2EG001470* encodes a short peptide of 80 amino acids and was not included in the
242 most recent annotation of the *Nasonia* genome at NCBI, but was present in OGS 2.0 [49]. A very
243 similar sequence is annotated in the close relative *Trichomalopsis* (JAL personal observation),
244 indicating that it is a *bona fide* transcript that is either novel, or very rapidly evolving.
245 *Nasvi2EG001470* is strongly expressed in the oosome and pole cells, while levels markedly
246 decrease in migrating germ cells (Fig. 3D1-D3).

247 Besides the expression in the oosome and the pole cells, *Nv-endoglucanase* is initially
248 expressed at both poles during pre-blastoderm stage and early blastoderm stage (Fig. 3E1-E2).
249 Later in blastoderm stage, the expression is down-regulated at the anterior pole and become

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

250 specific to the pole cells (Fig. 3E3). Proteins of this type are found extensively in Hemimetabola,
251 Hymenoptera, and Coleoptera, and appear to have been lost at the origin of Lepidoptera and
252 Diptera clades. It is intriguing that what roles this protein plays during *Nasonia* embryogenesis.

253 Five transcripts localized to the oosome enter and are maintained in the pole cells but are
254 then downregulated in the migrating germ cells (Fig. 3F1-J3). This set includes homologs of
255 *Drosophila anillin* (*scraps*), *qin, and innexin1 (ogre),* (*Nv-anillin*, *Nv-qin*, and *Nv-innexin1*,
256 respectively). Anillin is an actin binding protein that localizes to the contractile ring during
257 cytokinesis [50]. In *Drosophila* Anillin protein is localized in the cleavage furrows when forming
258 the PGCs [51], but its mRNA is ubiquitous. The early oosome localization of *Nv-anillin* mRNA
259 suggests that the protein might also play roles in oosome outside of a potential conserved role in
260 pole cell formation.

261 *Nv-qin* encodes a protein containing tudor domains along with an E3 ubiquitin ligase
262 domain. *qin* is important in processing germline piRNAs, repressing retroelements assembling
263 the nuage, and proper completion of oogenesis in the fly [52-54]. While *qin* has an important
264 late role in germline cells, it is only weakly and diffusely expressed during embryogenesis in
265 *Drosophila* [55]. Its mammalian homolog *RNF17* is required for production of specific particles
266 in the germline nuage and for sperm development, but not for early germline specification [56].

267 *Nv-innexin1* encodes a putative gap junction protein whose fly homolog is most well-
268 known for its role in proper development and function of the nervous system [57]. Other
269 unexpected roles for innexin proteins have been described and proposed in insects [58], but at the
270 moment the potential functional significance of the germline localization in *Nasonia* is unclear.

271 Two transcripts localized to the oosome and preserved in the pole cells have not clear
272 *Drosophila* homologs (Fig. 3I1-J3). One of these (*Nv-C16orf87*) encodes a homolog of the

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

273 human protein *C16orf87* and is expressed in the posterior region as well as specifically in the
274 oosome in pre-blastoderm stage (Fig. 3I1-I3). The protein of *Nv-C16orf87* belongs to the
275 uncharacterized protein family UPF0547, which contains the zinc-ribbon motif, and functions of
276 this protein and it homologs are not known. It appears that this gene has been lost specifically in
277 the Brachyceran fly lineage as it is found in beetles, moths, and some mosquitos, but not
278 *Drosophila* (JAL personal observation).

279 Finally, another ankyrin domain encoding transcript is strongly localized to the oosome
280 and is taken up into pole cells (Fig. 3J1-J2). It later has a complex and dynamic pattern in the
281 blastoderm stages (Fig. 3J2-J3, additional images not shown). This transcript is a member of the
282 newly described CLANK (Chalcid Lineage-specific ANKyrin-domain gene) family, of which
283 there are nearly 200 in the *Nasonia* genome [43]. To differentiate from the others, we name it
284 *Nasonia vitiripennis Oosome CLANK* (*Nv-OoCLANK*).
285

286 **Transcripts enriched in the oosome but excluded from pole cells**

287 A set of 12 genes is expressed in the oosome but not transported to the pole cells. We predicted
288 that these would have germline roles primarily in the oosome itself, or in the early stages of pole
289 cell formation. We also expect many transcripts in this set will have roles outside of germline
290 production, such as in embryonic patterning (as already known for *Nv-dpp* and *Nv-cad*) [59, 60].
291 Examples of what we consider potential embryonic patterning factors include: a CLIP protease
292 encoding message (*Nv-mp1*) related to fly Melanization Protease and Easter (Fig. 4A1-A3), *Nv-*
293 *kayak* (Fig. 4 B1-B3) encoding at transcription factor downstream of JNK signaling[61], and *Nv-*
294 *elbow* (Fig. 4C1-C3) encoding a single zinc-finger transcription factor [62]. *Nv-kayak* is later
295 expressed in a dorsal domain, indicating a conserved role in extraembryonic patterning.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

296 Interestingly, *Drosophila elbow* interacts with *orthodenticle* in specifying the ocelli and in
297 photoreceptor cell fate determination [63], and an intriguing possibility is that these proteins
298 work together in posterior patterning, with *Nv-elbow* possibly playing an important role in
299 differentiating the posterior targets of *otd* from the anterior ones.

300 Several oosome resident transcripts have suggestive functional annotations. For example,
301 the *Nasonia coronin* gene (*Nv-coronin,* (Fig. 4D1-D3)) encodes a protein whose homologs are
302 known to bind and modulate actin, provide links between the actin and microtubule
303 cytoskeletons, and regulate endo- and exocytosis in several developmental contexts [64, 65]. A
304 germline role for the *Drosophila* Coronin ortholog has not been observed. *Nasonia milton* (*Nv-*
305 *milton*) is another exciting transcript (Fig. 4E1-E3). *Drosophila* Milton is an adaptor protein that
306 allows mitochondria to be loaded onto, and transported by, microtubule motors [66].

307 An oosome resident mRNA encodes a protein whose fly ortholog is uncharacterized, but
308 whose function may be relevant to oosome function. This is the *Nasonia* homolog of *CG4552*
309 (*Nv-CG4552*), which encodes a protein with a TBC25 domain (Fig. 4F1-F3). Proteins with this
310 domain interact with Rabs to regulate membrane trafficking and dynamics. Such activities have
311 been shown to be crucial for Osk function in the fly [67], and *Nv-CG4552* may play a supporting
312 role in regulating membrane dynamics in the wasp. Another suggestive localized factor does not
313 have clear orthologs outside of the hymenoptera, but it does have two predicted RNA
314 Recognition Motifs, therefore we name it *Nv-rrm* (Fig. 4G1-G3). RRM domains bind RNA and
315 are components of proteins that regulate RNA localization and translation. This novel lineage
316 specific protein could therefore be involved in the localization of specific RNAs in the oosome,
317 or the regulation of translation of specific RNAs within it.

318

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

319

320 **Fig. 4. mRNAs localized in the oosome (Pre-blastoderm) but are excluded from the pole cells.** All embryos are aligned and grouped into three columns (Pre-blastoderm, Blastoderm and Gastrulation) according to their embryogenesis stages, with posterior side to the right and dorsal side on the top.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

321 Many of the oosome localized transcripts do not have annotations that lead to simple
322 hypotheses about their roles in specifying the germline. One of these is *Nv-waterproof*, which
323 encodes a fatty acyl-CoA reductase. *Drosophila waterproof* produces the hydrophobic molecules
324 that coat the trachaeal tubes during *Drosophila* embryogenesis and is essential for gas filling of
325 the trachea [68]. The protein's novel role in *Nasonia* germline is not clear and worth to
326 investigate, as is its early maternal expression in the oosome and later zygotic expression as the
327 dorsal strip and in the extraembryonic tissue. (Fig. 4H1-H3). *Nv-CG42269* encodes a predicted
328 organic ion transporter protein whose *Drosophila* homolog (*CG42269*) has no described function
329 (Fig. 4 I1-I3).

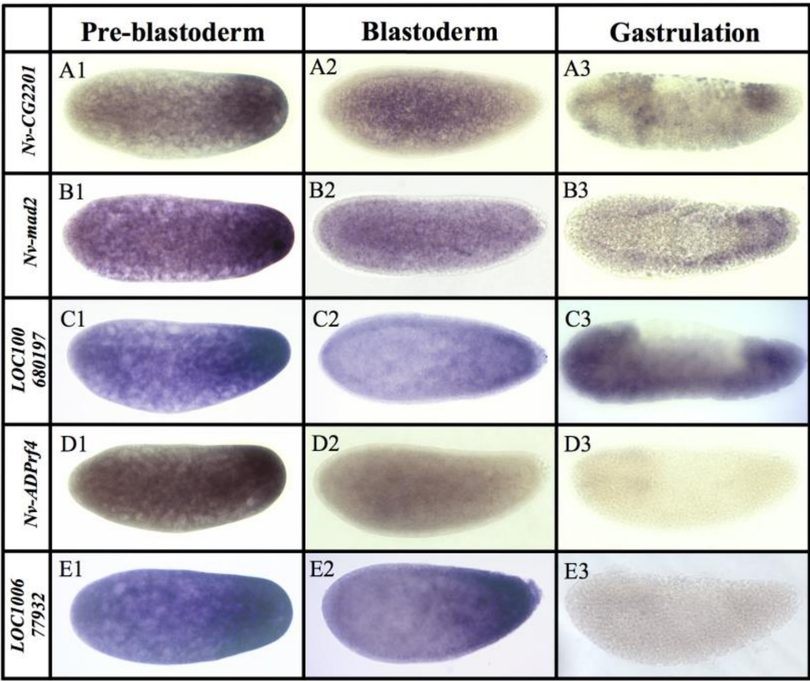
330 Three oosome localized transcripts have no clear homologs in *Drosophila* or in other
331 model organisms. *LOC100123551* has a sterile alpha motif (SAM) domain, which might indicate
332 protein-protein or -RNA interactions (Fig. 4J1-J3). *LOC103315681* contains weak similarity to
333 the N-terminal domain of Folded-gastrulation proteins (but is not a folded gastrulation ortholog)
334 (Fig. 4 K1-K3), while *Nasvi2EG022821* has no discernible conserved domains (Fig. 4L1-L3).
335 The functions of these factors will be the object of future investigation.

336 **Transcripts enriched in the posterior pole but not specifically the oosome**

337 There are five transcripts that are significantly enriched in the posterior region of the embryo, but
338 do not show significant association with the oosome (Fig. 5). The early embryonic expression for
339 these transcripts appears as a cap or broad posterior to anterior gradient of mRNA. The
340 significance of such transcripts to oosome assembly or to germ cell formation is not clear. Two
341 of this class are known developmental transcription factors: orthologs of Zerknuellt [69], and
342 Mothers against dpp [70] (*Nv-zen* and *Nv-mad2,* respectively (there are two mad paralogs in
343 *Nasonia*). Two transcripts are predicted to encode catalytic enzymes: a choline kinase

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

344 homologous to the *Drosophila CG2201* (*Nv-CG2201*), and a homolog of the ADP ribosylation
345 factor-like 4 protein (*Nv-ARL4*). Finally, *LOC100680197* and *LOC100677932* have no
346 identifiable homologs outside of hymenoptera. *LOC100680197* encodes a protein with MYND-
347 type zinc-fingers and a p27-like domain, while *LOC100677932* has no clear conserved or
348 functional domains.



349

350 **Fig. 5. mRNAs strongly enriched in the posterior region of the embryos, but not enriched in the oosome.** All embryos are aligned and grouped into three columns (Pre-blastoderm, Blastoderm and Gastrulation) according to their embryogenesis stages, with posterior side to the right and dorsal side on the top.

351

352 Functional analysis by parental RNA interference showed low phenotypic penetrance

353 While localization of an mRNA to the oosome and pole cells may strongly suggest a function
354 related to PGC specification, demonstration of any such function is required. We chose a
355 sampling of five promising molecules for in depth functional analysis (*Nv-bark*, *Nv-anillin*, *Nv-*
356 *rrm*, *Nv-coronin*, and *Nv-innexin1*).

357 We initially tried to apply our parental RNAi (pRNAi) approach [44], but quickly found
358 that this was not the ideal approach. Most dsRNAs caused partial sterility, with most of the

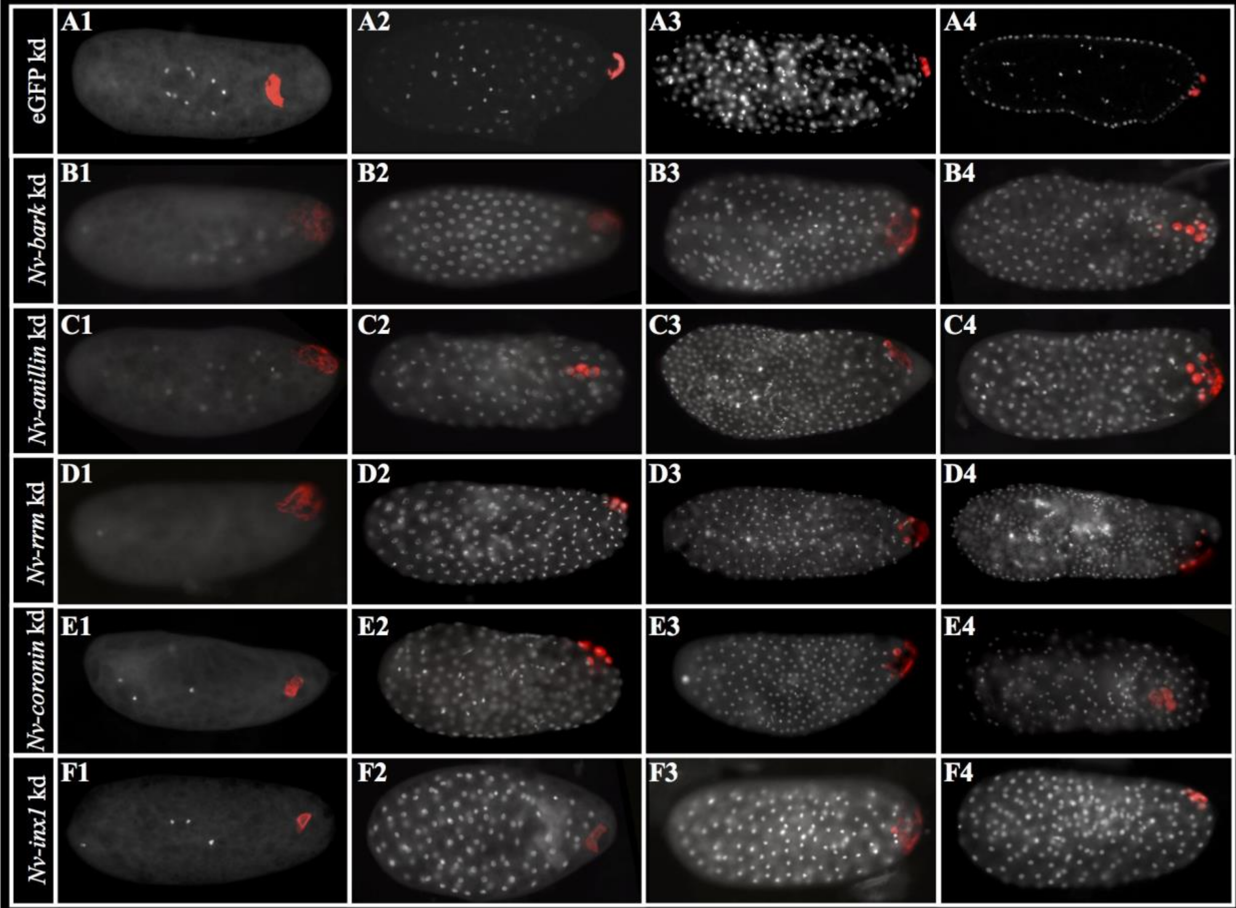
bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

359 obtainable eggs being apparently normal escapers. Eventually, we managed to collect embryos
360 with phenotypes in a very low penetrance (2%-6%) from three of the five genes that we studied
361 with the dsRNA concentration of 1ug/uL to 2.5ug/uL. These genes are *Nv-rrm*, *Nv-coronin*, and
362 *Nv-innexin1*. They all either have no pole cells formed or less pole cells formed with and
363 disorganized germ plasm residue at the posterior pole of the embryos in blastoderm stage.
364 Infrequently, *Nv-coronin* knockdown embryos were characterized by pole cells that did not
365 migrate to the gonad, but instead remained at the pole after gastrulation. (data not shown).
366 This issue was more serious for *Nv-bark* and *Nv-anillin*. At concentrations from
367 1.5ug/uL-2.5ug/uL), there were no eggs laid. At lower concentrations (250ng/uL, 500ng/uL and
368 750ng/uL), the eggs we collected all showed normal development. Since phenotypic embryos
369 were either completely absent or extremely rare for our genes of interest, it became necessary to
370 develop a new technique to assess the functions of the novel oosome genes we discovered.
371 **Development of an embryonic injection protocol for *Nasonia* RNAi**
372 To circumvent the low penetrance problem from the pRNAi, we developed a protocol for
373 embryonic injection of dsRNA followed by fixation and *in situ* hybridization (see details in
374 Methods and Materials). As a negative control, we injected dsRNA against eGFP to test whether
375 the physical injection and the dsRNA itself would affect the structure of the oosome and
376 formation of the pole cells non-specifically. We were happy to find that even in embryos with
377 obvious physical damage, the oosome and pole cells could form normally (Fig. 6A1-A4).
378 Of course, at a certain point, damage becomes too severe, leading to the death of the
379 embryo. We set a stringent criterion for collecting embryos for later analysis by removing those
380 where yolk leakage exceeded more than 10% of the embryo size (although all embryos with this
381 amount of damage showed normal germ plasm and germline development) (Fig. 6A2, A4). We

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

382 also excluded embryos that showed major morphological changes as compared to uninjected

383 embryos.



384

385 **Fig. 6. Functional analysis of a subset of germplasm localized factors by eRNAi.** A1-A4: eGFP dsRNA injected embryos of increasing as the negative control. B1-B4: *Nv-bark* dsRNA injected embryos, C1-C4: *Nv-anillin* dsRNA injected embryos, D1-D4: *Nv-rrm* dsRNA injected embryos, E1-E4: *Nv-coronin* dsRNA injected embryos, F1-F4: *Nv-innexin1* dsRNA injected embryos. All the embryos are positioned with posterior side to the right and dorsal side on the top. The red in B1-B4 are marked by *Nv-nos*, and the rest are marked by *Nv-bark*. DAPI (white) marks the nuclei.

386

387 After removing the embryos with obvious major damage both after injections, and when

388 performing eggshell dissection after the fixation, we were left with about 85% embryos with

389 viable embryogenesis by the time of imaging after *in situ* hybridization. The same criteria were

390 also applied to the experimental knockdowns where the percentages were roughly the same as in

391 the negative control When determining the phenotypes for the five genes, we considered the

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

392 disruptions of developmental events to be potential effects of a knockdown when they were only
393 specific to the knockdown and were never observed in the negative control.

394 We performed the same procedures when injecting and collecting the embryos as
395 described for the negative control. The penetrance given by the embryonic RNAi (eRNAi)
396 knockdowns is higher than the pRNAi knockdown, ranging from 20% to 39% across
397 experiments. Phenotypes are evident in both the pre-blastoderm and pole cell stages, for *Nv-bark*,
398 *Nv-anillin* and *Nv-rrm*. In contrast, the oosome is not affected in pre-blastoderm embryos after
399 *Nv-coronin* and *Nv-innexin1* dsRNA injection, with phenotypes becoming evident only after the
400 pole cells should have formed.

401

402 **RNAi against three novel germ plasm components unexpectedly disrupts the oosome at**
403 **from an early stage**

404 One transcript that particularly captured our attention was *Nv-bark*, as it was the transcript most
405 strongly and consistently associated with PGC specification over embryonic development.
406 However, its potential function in the germline is not clear. Since it encodes a transmembrane
407 protein involved in epithelial junctions in *Drosophila*, we speculated that it might have a role in
408 mediating adhesion or migration of the pole cells once they were formed. Surprisingly, the
409 phenotypes produced by knocking down this transcript showed a much earlier requirement for
410 this transcript. In early embryos (before migration of nuclei to periphery), the oosome has lost
411 its integrity as a single unit (Fig. 6B1). Instead there are scattered particles of what appears to be
412 oosome-like material, concentrated around the lateral cortex of the embryo. Later, at the time the
413 pole cells would normally form, no budding is observed, and some loose aggregates of germ
414 plasm like material remains attached to the cortex (Fig. 6B2-B3). In some cases, germ plasm

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

415 surrounds nuclei in a way similar to what is seen in pole cells (Fig. 6B3, B4). However, these
416 nuclei remain part of the embryonic syncytium.

417 Like *bark,* the mRNA of *Drosophila anillin* is not localized to the polar granules [71].
418 Anillin protein, however, accumulates at the base of pole cells when they are budding in
419 *Drosophia* [51]. Since Anillin is a major component of the contractile ring during mitotic
420 cytokinesis [50] and is enriched at the bud furrow during pole cell formation in *Drosophila* [51],
421 we predicted that the enrichment of *Nv-anillin* mRNA in the oosome and pole cells would be
422 related to *Nasonia's* unique way of forming a single large pole bud instead of several small ones,
423 as occurs in *Drosophila*. Surprisingly, the phenotype of *Nv-anillin* is indistinguishable from that
424 of *Nv-bark*: The oosome does not form properly, and germ plasm material remains bound to the
425 posterior cortex of the embryo, leading to the association of the syncytial nuclei at random
426 posterior locations (Fig. 6C1-C4). The same set of phenotypes is observed when we knock down
427 *Nv-rrm* with eRNAi (Fig. 6D1-D4). Thus, three genes with very different predicted functions all
428 result in the same phenotype.

429 It is important to note at this point that the common phenotype of the above three
430 knockdowns is not the same as the complete loss of germ plasm activity. Such phenotypes are
431 seen for *Nv-osk*, *Nv-vas*, and *Nv-aubergine (Nv-aub).* When these genes are knocked down,
432 posterior mRNAs such as *Nv-nos* take on a uniformly graded posterior cap and no enriched
433 accumulation of germ plasm markers is ever observed at the posterior [35, 40]. This indicates
434 that these transcripts are involved in the specific form of the oosome, but may not be essential for
435 the production of germline-like cells.

436 Finally, despite the fact that the oosome does not form or migrate through the embryo as
437 in wildtype, a bud like protrusion similar to the one that initiates pole cell formation is regularly

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

438 observed (Fig. 6C3, Fig. 1E), and this bud is not associated with the bulk of the germ plasm like
439 material. This indicates that neither the oosome, nor its remnants, induce the bud, and that bud
440 formation may be autonomous to the embryo. This has some precedent in *Drosophila*, where the
441 autonomous ability of the fly embryo to produce pole-cell like structures at both poles of the
442 embryo is revealed when Arf guanine exchange factor Steppke is reduced [72]. However, in the
443 fly, normal global repression of pole-cell formation is overcome by germ plasm components
444 (primarily *germ-cell-less*), while in *Nasonia,* at least the initial budding appears to be germ plasm
445 independent.

446

447 **RNAi against *Nv-coronin* and *Nv-innexin1* does not affect the oosome, but disrupts pole cell**
448 **formation**

449 We tested the functions of two other transcripts with eRNAi: *Nv-coronin* and *Nv-innexin1*.
450 Knockdown of both genes left the oosome intact, and able to migrate through the embryonic
451 cytoplasm normally (Fig. 6E1, F1). However, in both cases, pole cell formation fails, indicating
452 that these genes have downstream functions that are specific to pole cell formation (Fig. 6E2-E4,
453 F2-F4). *Nv-coronin*'s functional annotation is consistent with a role in cellularization, as it is
454 predicted to interact with both the microtubule and actin cytoskeletons, both of which are crucial
455 for mitosis and cell formation. While Innexin1 is most well known as a crucial component of
456 gap junctions and a role in the nervous system [57], a role for at least one Innexin in cellurization
457 has been demonstrated in the beetle *Tribolium* [58], suggesting that the full potential of these
458 proteins in regulating cellular processes has not been fully explored.

459

460 # Discussion

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

461 **RNA-seq analyses**

462 Our results have uncovered an unexpectedly large amount novelty in the mRNA content of the
463 germline determinant of the wasp *Nasonia vitripennis*. This was achieved using RNAseq
464 followed by statistical detection of differential expression of mRNAs between the anterior and
465 posterior poles of the early embryo. The statistical predictions were then extended by *in situ*
466 hybridization and RNAi of candidate genes.

467 Our goal was to identify all transcripts specifically localized to the oosome, and it is
468 worth contemplating how close we came to achieve this goal. We used three different
469 approaches that varied in the strategy for mapping, quantifying transcripts and assessing
470 differential expression. These analyses agreed on the vast majority of the genes with putative
471 significantly different enrichment at the two poles, and most of the disagreement was at the
472 margin of differential expression that would be likely biologically significant. In addition, all
473 analyses found genes previously identified to be localized anteriorly or posteriorly. We even
474 detected all known genes with both strong oosome localization and strong localization at the
475 anterior pole. We thought these could be missed, because if the anteriorly localized mRNA
476 population approached the same levels as that in the oosome, the enrichment might be obscured.

477 For these reasons, we believe that we have uncovered the vast majority of mRNAs
478 localized to the *Nasonia* oosome. Of course, no approach can guarantee comprehensiveness, and
479 some molecules may have been missed for several reasons. For example, while the *Nasonia*
480 genome is well annotated, it is possible that a very few genes are not represented in the genome
481 assembly or the transcriptome annotations, and thus would not have been assessed. Indeed, we
482 did identify slightly different sets of genes when using NCBI annotation 102 versus the OGS 2.0
483 annotation [49], which were created using different approaches to identify and predict genes.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

484 Most of the differences are at the margins of significance and low levels of differential
485 expression. However, there are a handful of confirmed localized transcripts detected using
486 annotation 102 that were missed in analyses with OGS 2.0. In addition, at least one transcript
487 was not annotated in annotation 2.1, but was found in annotation 2.0 and was confirmed to be
488 localized to the oosome (*Nasvi2EG001470*, Fig 3D). This is likely a false negative annotation in
489 annotation 102, as a very similar sequence is found in the transcriptome of the very closely
490 related wasp *Trichomalopsis sarcophagae* [73].

491 Another issue that can cause false negatives is the large number of *Nasonia* transcripts
492 that overlap, which can lead to the concatenation of transcripts. If a localized transcript is fused
493 to a ubiquitous, highly expressed transcript, the signal of localization can be lost. This problem
494 is largely solved by preventing novel junctions when mapping with tophat2. However, this also
495 seem to change the calculation of significance for some transcripts, so performing the analysis
496 with and without novel junctions gives more complete results.

497 Finally, there are some unknown artifacts which may cause significantly enriched genes
498 to be missed. A prime example of this is our discovery of *Nv-coronin*. A preliminary attempt at
499 this experiment resulted in generally poor sequencing results, and a completely unusable
500 replicate. In general, this analysis gave predictably worse results, where several known localized
501 genes were not found to be significant. However, all of the molecules whose functions were
502 analyzed in this manuscript were found to be statistically significantly enriched at the posterior,
503 including *Nv-coronin*, and we proceeded to clone them to test their localization and function. In
504 the subsequent analyses based on high quality sequencing results (used as the basis for this
505 manuscript), *Nv-coronin* was excluded by cuffdiff for statistical testing for unknown reasons,
506 despite showing similar posterior enrichment that had previously been deemed statistically

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

507 significant. Thus, there is a potential for false negatives if this artifact affects several genes.
508 Very few transcripts in the "not tested" category show a similarly strong posterior enrichment, so
509 we believe that this effect is also small.
510

511 **Comparison of the oosome of *Nasonia* to the polar granules of *Drosophila***

512 Germ plasm in insect embryos must perform several different functions. First, it must be able to
513 concentrate and arrange the set of proteins and RNAs (e.g. Tud, Vas, *nos* mRNA, Piwi/Aub) that
514 are associated with germline fate in the posterior of the egg. Second, it must ensure its own
515 incorporation into the PGCs as they form by interacting with the specialized cytoskeletal
516 structures that mediate pole cell formation. Germ plasm must also include molecules involved
517 specialized features of PGCs. Such functions include the concentration and selection of
518 mitochondria, repression of transcription, repression of transposons, and guidance of the germ
519 cell migration to the gonad primordia.

520 On the other hand, we had already observed that the oosome's morphology and
521 interaction with the cytoskeleton (during its migration and formation of the pole cells) was quite
522 distinct from *Drosophila*, as is the way the pole cells migrate into the interior of the embryo.
523 Therefore, we expected to find a mixture of conservation and novelty when we examined the
524 mRNA content of the oosome. Indeed, this is what was found, but with a surprisingly strong bias
525 toward novelty.

526 In terms of conservation, our analyses found *Nv-nos, Nv-osk* (already known factors
527 found in both oosome and polar granules) as well as *Nv-ovo*. We also found *Nv-aub* whose
528 ortholog is localized as protein, but not mRNA, in *Drosophila*. While there are a large number of

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

529 polar granule localized mRNAs that are not found in the oosome, we will only discuss a few that
530 are significant.
531 mRNA for *polar granule component (pgc)* encodes a small peptide [74], and is strongly
532 localized to the posterior pole [75]. Polar-granule-component protein has a crucial role in the
533 global repression of transcription that occurs in pole cells upon their formation, through an
534 interaction with the transcription elongation factor TEF-b [74]. This repression is a widely
535 conserved feature of PGCs across animals, which makes it somewhat surprising that *pgc* appears
536 to be a novelty in the *Drosophila* lineage [75].
537 Fascinatingly, TEF-b may be a unifying factor underlying germline quiescence, as it
538 interacts with PIE-1 in the worm *C. elegans* to repress transcription [76]. In line with this, we
539 have found a highly conserved transcription elongation factor, *Nv-spt5* (Fig. 3 A1-A3) localized
540 strongly to the oosome and pole cells. In human cells, Spt5 acts as an inhibitor or transcriptional
541 elongation, until its C-terminal domain is phosphorylated in a TEF-b dependent manner [77]. If
542 an interaction between *Nasonia* Spt5 and TEF-b does have a role in regulating the cessation of
543 transcription in *Nasonia* pole cells, it would be strong evidence for TEF-b being a core conserved
544 component of the germline fate, whose interaction partners and regulators are labile across
545 lineages.
546 Another crucial *Drosophila* germ cell factor that is not present in the *Nasonia* oosome is
547 *germ cell less* (*gcl*). The Germ-cell-less protein itself is very highly conserved at the sequence
548 level in *Nasonia,* but the mRNA showed no enrichment in our RNAseq experiments, and we
549 independently confirmed by *in situ* hybridization that it is expressed uniformly throughout the
550 early embryo (not shown). Gcl is important for the proper production of pole cells, apparently by
551 regulating the orientation centrosome separation at the posterior pole, which is required for

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

552 efficient pole cell formation and uptake polar granules by the pole cells [78]. At the molecular
553 level, Gcl seems to act by downregulating torso signaling, to allow the proper conditions for pole
554 cells to form. The lack of Gcl function in the germline of *Nasonia* is consistent with the lack of
555 Torso signaling at the termini in the wasp [79], making the need for Gcl redundant. At the
556 moment, it is not clear whether the use of Gcl in pole cell formation is a recent novelty in
557 *Drosophila*, or whether it was present ancestrally, but lost in the Hymenopteran lineage.

558 A number of transcripts found in the *Nasonia* oosome are good candidates for generating
559 essential PGC features. For example, a high concentration of mitochondria is a strongly
560 conserved feature of germ plasm and PGCs across animals [80, 81]. In *Drosophila*, the long Osk
561 isoform plays an important role in concentrating mitochondria in the pole plasm [82]. But, since
562 the long form of Osk appears to be a novelty of *Drosophila* and its close relatives, other
563 molecules should be expected to perform this role in other species. Suggestively, mRNA
564 encoding a Milton ortholog was found strongly localized to the *Nasonia* oosome. Milton acts an
565 adaptor that loads mitochondria onto microtubule motors for transport and localization within
566 and between cells in *Drosophila* [83], and we propose that *Nasonia* Milton plays a role in
567 enriching mitochondria around the oosome and in the pole cells in the wasp, and perhaps other
568 insect species that lack the specialization of long Osk isoform.

569 Another critical function for germ cells is the control of transposable elements, which is
570 often dependent on Tudor domain containing proteins. mRNAs for two Tudor domain proteins
571 are present in the oosome, including *Nv-qin* and *Nv-tdrd7*. Neither of them is enriched in the
572 polar granules or pole cells in *Drosophila*, but both have crucial roles during oogenesis to reduce
573 the activity of transposable elements [52, 53, 84]. The presence of these additional Tudor
574 domain encoding transcripts may indicate that either there is an increased activity of transposable

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

575 elements in *Nasonia* that requires an earlier response, or perhaps other mechanisms are employed
576 in *Drosophila* to combat transposon activity in the early PGCs. Further sampling of germ plasm
577 of other insects should help to resolve these questions.

578 Germ cells are known to have a distinct metabolic profile from somatic cells, and this
579 difference is related to their pluripotent stem cell-like properties, and to the requirements of their
580 migratory properties [85, 86]. Potentially related to this we have found that a *Nasonia* insulin-
581 like growth factor I mRNA is localized to the oosome. Interestingly, this mRNA encodes a short
582 protein that shows much stronger similarity to insulin proteins of vertebrates, than it does to any
583 of the insulin-like molecules of *Drosophila* (DILPs) (JAL personal observation). In addition, an
584 mRNA encoding a putative organic cation transporter (*Nv-CG42269)* containing a Major
585 Facilitator Superfamily (MFS) domain is strongly localized to the oosome. Such molecules are
586 crucial for regulating cellular metabolism and signaling at multiple levels, by controlling the
587 trafficking of many small organic molecules (including sugars) within and between cells [87, 88].
588 Lipid metabolism is also uniquely regulated in germ cells, and the identification of the *Nasonia*
589 homolog of the Acyl-CoA reductase *waterproof* may reflect this [89]. Finally, we surprisingly
590 found a transcript encoding a protein similar to endoglucanases found in several insect lineages
591 (but absent from Diptera). Enzymes of this type are broadly defined by their ability to break
592 down polymers glucose. The substrate and potential role for this enzyme in the germline is not
593 yet known.

594 In addition to providing insight into the conserved functions of germ cell components, we
595 also found several molecules that do not have clear homology outside of the Hymenoptera, or in
596 some cases outside of *Nasonia* and its closest relatives. This includes a novel RNA recognition
597 domain containing protein whose function we analyzed in depth (*Nv-rrm* discussed in the

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

598 following section). We also found that an mRNA encoding an ankyrin domain protein (*Nv-*
599 *OoCLANK* (*LOC100679945*)) that belongs to a family of proteins that underwent a massive
600 amplification within chalcid wasp lineage, and which appears to have entered the ancestral
601 chalcid wasp by horizontal transfer [43]. Finally, a handful of transcripts have no identifiable
602 domains or homologs. The functional relevance of these molecules will be an area of intense
603 interest in the future.
604
605 **Unexpected functions of novel oosome components**
606 Early in our analysis we chose a handful of transcripts for functional analysis. These were
607 chosen based on a combination of criteria that included high enrichment in the oosome, novelty,
608 and the potential to give phenotypes that we could characterize with the current set of functional
609 tools available to us in *Nasonia*. Three of these (*Nv-bark*, *Nv-anillin* and *Nv-rrm*) gave an
610 unexpected phenotype, where oosome-like material was not coalesced into the typical spherical
611 oosome structure, but rather was scattered in clumps attached to the plasma membrane near the
612 posterior pole. Eventually, these clumps of germ plasm-like material come into contact with
613 syncytial nuclei when they migrate to the cortex. However, this remnant material is unable to
614 induce the pole cell fate. It is important to reiterate that this phenotype is quite distinct from that
615 seen for genes that have a core role in oosome assembly (*Nv-osk*, *Nv-tud*, *Nv-vas* and *Nv-aub*
616 (not shown). In these cases, there is no hint of the oosome, and mRNAs normally localized in
617 the oosome are distributed in homogenous caps at the posterior pole, rather than as discrete
618 clumps of material [35].
619 These phenotypes indicate these genes are involved in the coalescence of the oosome into
620 single entity within the central column of cytoplasm in the embryo, and/or maintenance of

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

621 oosome integrity. The molecular bases of such functions are not completely clear at the moment
622 for any of these three genes. *Nasonia* Bark is a putative transmembrane protein and would be
623 predicted to be targeted to the membrane. One potential hypothesis is that *Nasonia* Bark is
624 targeted to the membrane at the posterior pole of the embryo, releases and/or repels oosome
625 material from adhering to the plasma membrane, forcing the oosome to remain in the bulk
626 cytoplasm where it concentrated into a large sphere and moved around by strong cytoplasmic
627 flows that occur during the earliest cleavages. This model would also imply that interaction with
628 the cortex prevents oosome material from coalescing, leading to the scattered clumps we observe
629 in *Nv-bark* knockdowns. Mechanistically, this could be related to the ability of Bark to induce
630 endocytosis [90], a process that is associated with proper anchoring of the germ plasm and the
631 recruitment of specialized actin binding proteins to the posterior pole in *Drosophila* [91].

632 An alternative hypothesis is that *Nasonia* Bark produced in the oosome is not secreted but
633 is instead incorporated as an important structural component of the large, solid form of the
634 oosome. This is consistent with the structure of the protein, which contains main protein-protein
635 interaction domains of different types [45, 46]. Testing these hypotheses will require in depth
636 analysis of the subcellular localization of *Nasonia* Bark during early embryogenesis, and
637 proteomic analysis of binding partners of *Nasonia* Bark.

638 The role of *Nv-anillin* in maintaining the stability of the oosome was also surprising and
639 the molecular basis of the phenotype will require further investigation. Anillin orthologs are well
640 known as actin binding proteins involved in assembling the contractile ring required to separate
641 cells in cytokinesis [50]. Anillin also plays a crucial and novel role in the specialized cytokinesis
642 of the *Drosophila* pole cells [51]. While these known functions might have indicated that
643 oosome localization of *Nv-anillin* was related to an important role in the specialized polar bud

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

644 formed in *Nasonia*, RNAi showed that this protein has an earlier role in oosome
645 assembly/maintenance (Fig. 6 C1). Similar to *Nasonia* Bark, one possible function of *Nasonia*
646 Anillin is as a structural component of the oosome, which may or may not be related to its ability
647 to bind actin and associated proteins. Alternatively, *Nv-anillin* may act to release and/or repel the
648 oosome from the cortex, as proposed above for *Nv-bark*.
649
650 Anillin homologs have known functions which are directly related to the processes of release of
651 germ plasm from the embryonic plasma membrane. In *Drosophila*, the germ plasm is tightly
652 bound to the plasma membrane until the nuclei reach the posterior pole of the embryo. The
653 centrosomes associated with these nuclei mediate detachment of the pole plasm from the cortex
654 through interactions of the astral microtubules emanating from the centrosomes [92]. This is in
655 contrast to *Nasonia,* where the oosome detaches from the cortex at about the same time as the
656 zygotic nucleus begins its first division at the anterior pole. While nuclei are lacking at this time,
657 numerous centrosomes are present, as they are provided maternally in a process characteristic of
658 many Hymenopteran embryos [93]. *Nasonia* Anillin could be relevant to a model where astral
659 microtubules emanating from maternally provided centrosomes detach oosome material from the
660 cortex, because Anillin homologs have been shown to mediate interactions between the actin
661 cytoskeleton and cortical and subcortical microtubule arrays in multiple model systems [94, 95].
662 Again, in depth examination of the subcellular localization and interactions of *Nasonia* Anillin
663 will be required to completely understand its role in maintaining the oosome.
664 Based on the presence of only RNA recognition motifs in the protein, we predict that *Nv-*
665 *rrm* will have one of two likely roles. One possibility is that it is involved in translational
666 regulation of key regulators of oosome structure, presumably including *Nasonia* Anillin and

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

667 *Nasonia* Bark. Alternatively, *Nv-rrm* may be important in binding RNA and protein in order to
668 maintain the structural integrity of the oosome. In addition, neither of these possibilities is
669 mutually exclusive.
670 The knockdowns of *Nv-coronin* and *Nv-innexin1* had specific effects only on the
671 formation of pole cells, while the oosome appeared to remain intact. Coronin is an actin binding
672 protein associated with the formation of highly concentrated networks of F-actin [65]. It seems
673 likely that *Nasonia* Coronin has an important role in organizing an actin cytoskeleton
674 arrangement specialized for the formation of the large polar bud that initiates pole cell formation.
675 The potential role of *Nv-innexin1* is somewhat more mysterious. Innexins are typically
676 known as components of the gap junctions that are found in some tightly integrated epithelial
677 tissues. Such junctions would not be expected of the motile pole cells. Interestingly, Innexin-7 (a
678 paralogous protein with a similar structure to *Nasonia* Innexin1) in the beetle *Tribolium* has a
679 novel role in cellularization of the syncytial blastoderm. Such a function for *Nasonia* Innexin1
680 would explain the failure in pole cell formation we see after RNAi.
681
682 **Conclusion**
683 This work has revealed numerous unexpected mRNAs that are localized to the germ plasm of the
684 wasp *Nasonia.* The results have given insights on the potentially ancestral mechanisms used by
685 germ plasm to accomplish conserved required functions, such as a possible ancestral role for
686 Milton orthologs to bring mitochondria to the germ plasm, a function replaced by the long Osk
687 isoform in *Drosophila.* On the other hand, our results have identified numerous components that
688 are likely to be specific to *Nasonia* and its relatives in the parasitic wasp lineage (e.g., *Nv-*
689 *OoCLANK, Nv-rrm*, and the use of Nv-Bark and Nv-Anillin in assembling the oosome). Deeper

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

690 analyses of the functions of these molecules in *Nasonia*, and broader sampling of germ plasm in
691 other holometabolous insects will be required to determine the patterns of evolutionary change in
692 the germ plasm. Such analyses will be important because the germ plasm is a uniquely powerful
693 organelle that can rapidly drive naive nuclei into a highly specialized, yet functionally totipotent
694 state. Understanding how and why such a fundamental substance changes and is even lost in the
695 course of evolution will provide foundational insights into the mechanisms of cell fate
696 determination and the interaction of subcellular organelles and their cellular milieu.
697

698 # Methods

699 **Sample preparation**

700 In order to identify the components of the maternally deposited mRNAs in the oosome located in
701 the posterior half of the parasitoid wasp *Nasonia vitripennis* embryos, we collected and bisected
702 the pre-blastoderm stage embryos (0-2 hours after egg lay at 25 °C) to detect the differential
703 expression levels between the anterior and posterior halves of the embryos.

704 The embryos were aligned with the anterior pole to the right on the ice-prechilled and
705 heptane glue-coated slide. Then a thin layer of halocarbon oil 700 (Sigma) was applied to cover
706 the embryos. The slide was transferred on the dry ice-prechilled "guillotine" and then was
707 anchored by tightening the screws on each end. After the halocarbon oil 700 was solidified on
708 dry ice, put the guillotine on the dry ice-prechilled stabilizer and transferred it under the
709 dissection microscope. The embryos were positioned to match the slot in the guillotine where the
710 dry ice-prechilled razor blade will be inserted into. After the embryos were cut, the anterior and
711 posterior halves of the embryos were immediately collected and transferred into the two 1.5mL

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

712 non-stick RNase-Free microfuge tubes (Ambion) with the dry ice-prechilled probe, separately.
713 Three biological replicates were created.

714 Total RNA was isolated from these six samples for library preparation. In the library
715 preparation upon which this manuscript is based, we used around one microgram of total RNA
716 from each sample, in which 100ng of total RNA was from the *Nasonia* other 900ug was from a
717 distantly related parasitic wasp (*Melittobia digitata* [43]). Libraries were prepared using the
718 NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB #E7420) in conjunction
719 with NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB #E7490). Libraries were
720 validated and quantified before being pooled and sequenced on an Illumina HiSeq 2000
721 sequencer with a 100 bp paired-end protocol. Sequence files are available in the NCBI SRA
722 database under accession SRP156232.
723
724 **RNA sequencing data analysis**
725 The quality of the sequencing data was determined using FastQC software. The sequences were
726 processed by Cufflinks package for differential expression detection, using multiple variations on
727 the default parameters (job files in Supplemental File 2). Briefly, raw reads were aligned to
728 annotation 102 of assembly 2.1 of the *N. vitripennis* genome
729 (https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Nasonia_vitripennis/102/) using
730 TopHat2[96]. These results were either used directly in cuffdiff [97], or were further processed
731 using stringtie [98] to generate new transcriptome predictions and quantification. Various
732 normalization parameters were used, and each permutation gave slightly different results. In
733 addition, we mapped the reads using assembly 1.0 and annotation 2.0 (OGS 2.0) [49]. All of the
734 commands jobs in this process of analysis are included in the Supplemental File 2. The cuffdiff

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

735 results of all permutations of the analysis are provided in Supplemental File 3. The computing
736 work was done by the High Performance Computing Cluster located at University of Illinois at
737 Chicago.

738 Probes and dsRNAs for the chosen genes were generated by the protocol described in
739 [40]. Primers for generating these templates are provided in Supplemental File 4. Alkaline
740 phosphatase *in situ* hybridization was performed by the protocol described in [99].

741

742 **Embryonic RNA interference (eRNAi)**

743 In order to perform eRNAi on the early *Nasonia* embryos to study the germline candidate genes'
744 functions, we created the following workflow:

745 Around 30 pre-blastoderm stage embryos (0-1 hour after egg lay at 25 °C) were collected
746 and quickly aligned vertically on the heptane glue-coated 18mm×18mm coverslip. This coverslip
747 was transferred and anchored on the ice-prechilled slide by applying a thin layer of water. The
748 slide was then put in an air tight petri dish with proper amount of desiccant (Drierite with
749 indicator, 8 mesh, ACROS Organics) pre-chilled at 4 °C. The embryos were dehydrated in the
750 desiccant at 4 °C for 45 minutes. After dehydration, the embryos were covered with a layer of
751 halocarbon oil 700 and were ready for microinjection.

752 The dsRNAs were dissolved in Nuclease-Free Water (Ambion) at the concentration of
753 1mg/mL and loaded into the Femtotips II Microinjection Capillary (Cat. No. 930000043,
754 Eppendorf). The constant pressure was set at 500hpa and the injection pressure was set initially
755 at 250hpa with periodic adjustment as the needle changed over the course of injection. The
756 process of injection was performed at room temperature and needed to be done as soon as
757 possible. After injection, the slide was transferred into a paper towel-moisturized petri dish pre-

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

758 warmed at 28 °C to incubate the injected embryos for specific developmental stages. The
759 embryogenesis of these embryos was stopped at pre-blastoderm stage (before the budding),
760 beginning of blastoderm stage (during budding), and later in blastoderm stage (pole cells formed
761 and/or after pole cell divisions). To stop the development, the coverslip was put into the heptane
762 to wash off the halocarbon oil 700 for three minutes, and then transferred into the 37%
763 formaldehyde-saturated heptane for 2-5 hours fixation in the dark with the embryos facing up.
764 After fixation, the coverslip was carefully taken out of the fixative, and flipped upside
765 down to gently press the embryos on a double-sided tape that was taped on a petri dish, so that
766 all the embryos can be anchored on the tape for dissection. Add about 15mL PBS with 1%
767 Tween, use the needle (BD PrecisionGlide Needle, 30G × 1) to carefully remove the eggshells
768 from the embryos. The dissected embryos were then transferred by pipette into the 1.5mL non-
769 stick RNase-Free microfuge tubes. The embryos were immediately dehydrated by 100%
770 Methanol and stored at -20 °C.
771 Before performing fluorescent *in situ* hybridization (FISH) on those eRNAi knocked out
772 embryos, they need to be rehydrated with a series of Methanol/PBT washes (75%, 50%, 25%).
773 The protocol for FISH was adapted from [40]. A detailed protocol is available on request.
774

775 ## Declarations

776 **Ethics approval and consent to participate**

777 Not applicable

778 **Consent for publication**

779 Not applicable

780 **Availability of data and material**

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

781 Sequencing results can be found in the NCBI SRA database at:
782 https://www.ncbi.nlm.nih.gov/sra/SRP156232.

783 **Funding**

784 This work was supported by startup funds from UIC, and NIH grants 1R03HD087476 and
785 1R03HD078578.

786 **Authors' contributions**

787 JAL and HHQ collaborated on experimental design, analyzed the data and wrote the paper.
788 HHQ performed the experiments and collected the data.

789 **Acknowledgements**

790 We thank Daniel Pers for performing the combined *Nasonia/ Melittobia* RNAseq libraries, and
791 the UIC RRC Sequencing core for arranging the sequencing of our samples.

792 **Competing Interests:**

793 The authors declare no competing interests.

794

795

796 References

797 1. Extavour CG, Akam M: **Mechanisms of germ cell specification across the metazoans:
798 epigenesis and preformation**. *Development* 2003, **130**(24):5869-5884.
799 2. Extavour CG: **Evolution of the bilaterian germ line: lineage origin and modulation of
800 specification mechanisms**. *Integr Comp Biol* 2007, **47**(5):770-785.
801 3. Hegner RW: **Experiments with Chrysomelid Beetles. III. The Effects of Killing Parts of
802 the Eggs of Leptinotarsa decemlineata**. *Biological Bulletin* 1911, **20**:237-251.
803 4. Illmensee K, and A. P. Mahowald: **Transplantation of posterior polar plasm in
804 Drosophila. Induction of germ cells at the anterior pole of the egg**. *Proceedings of the
805 National Academy of Sciences* 1974, **71**(4):1016-1020.
806 5. Illmensee K, and A. P. Mahowald: **The autonomous function of germ plasm in a somatic
807 region of the Drosophila egg**. *Experimental cell research* 1976, **97**(1):127-140.
808 6. Xia QY, Zhou, Z. Y., Lu, C., Cheng, D. J., Dai, F. Y., et al.: **A draft sequence for the genome
809 of the domesticated silkworm (Bombyx mori)**. *Science* 2004, **306**:1937-1940.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

810 7. Nagy L, Riddiford, L., Kiguchi, K. : **Morphogenesis in the Early Embryo of the**
811 **Lepidopteran Bombyx-Mori**. *Developmental Biology* 1994, **165**:137-151.
812 8. Juhn J, Marinotti, O., Calvo, E., James, A. A.: **Gene structure and expression of nanos**
813 **(nos) and oskar (osk) orthologues of the vector mosquito, Culex quinquefasciatus**.
814 *Insect Molecular Biology* 2008, **17**:545-552.
815 9. Suzuki N, Shimizu, S., Ando, H. : **Early Embryology of the Alderfly, Sialis-Mitsuhashii**
816 **Okamoto (Megaloptera, Sialidae)**. *International Journal of Insect*
817 *Morphology & Embryology* 1981, **10**:409-418.
818 10. Werren JH, Richards, S., Desjardins, C. A., Niehuis, O., Gadau, J.: **Functional and**
819 **Evolutionary Insights from the Genomes of Three Parasitoid Nasonia Species**. *Science*,
820 **327**:343-348.
821 11. Berg GJ, Gassner, G.: **Fine structure of the blastoderm embryo of the pink bollworm,**
822 **Pectinophora Gossypiella (Saunders) (lepidoptera: Gelechiidae)**. *International Journal*
823 *of Insect Morphology and Embryology* 1978, **7**:81–105.
824 12. Miura T, Braendle, C., Shingleton, A., Sisk, G., Kambhampati, S., Stern, D. L.: **A**
825 **comparison of parthenogenetic and sexual embryogenesis of the pea aphid**
826 **Acyrthosiphon pisum (Hemiptera : Aphidoidea)**. *Journal of Experimental Zoology Part*
827 *B-Molecular and Developmental Evolution* 2003, **295B**:59-81.
828 13. Bull AL: **Stages of Living Embryos in the Jewel Wasp Mormoniella-**
829 **(Nasonia)-Vitripennis-(Walker)(Hymenoptera, Pteromalidae).** *Journal of Insect Morphology &*
830 *Embryology* 1982, **11**:1-23.
831 14. Klag J: **Differentiation of primordial germ cells in the embryonic development of**
832 **Thermobia domestica. Pack. (Thysanura): an ultrastructural study**. *Journal of*
833 *Embryology and Experimental Morphology* 1977, **38**:93-114.
834 15. Chang C-c, Dearden P, Akam M: **Germ Line Development in the Grasshopper**
835 **Schistocerca gregaria: vasa As a Marker**. *Developmental Biology* 2002, **252**(1):100-118.
836 16. Consortium THGS: **Insights into social insects from the genome of the honeybee Apis**
837 **mellifera**. *Nature* 2006, **443**(7114):931-949.
838 17. Dearden PK: **Germ cell development in the Honeybee (Apis mellifera); vasa and nanos**
839 **expression**. *BMC Dev Biol* 2006, **6**:6.
840 18. Dearden PK, Wilson MJ, Sablan L, Osborne PW, Havler M, McNaughton E, Kimura K,
841 Milshina NV, Hasselmann M, Gempe T *et al*: **Patterns of conservation and change in**
842 **honey bee developmental genes**. *Genome Res* 2006, **16**(11):1376-1384.
843 19. Juhn J, James AA: **oskar gene expression in the vector mosquitoes, Anopheles gambiae**
844 **and Aedes aegypti**. *Insect Mol Biol* 2006, **15**(3):363-372.
845 20. Nakao H, Hatakeyama M, Lee JM, Shimoda M, Kanda T: **Expression pattern of Bombyx**
846 **vasa-like (BmVLG) protein and its implications in germ cell development**. *Dev Genes*
847 *Evol* 2006, **216**(2):94-99.
848 21. Schroder R: **vasa mRNA accumulates at the posterior pole during blastoderm**
849 **formation in the flour beetle Tribolium castaneum**. *Dev Genes Evol* 2006, **216**(5):277-
850 283.
851 22. Khila A, Abouheif E: **Reproductive constraint is a developmental mechanism that**
852 **maintains social harmony in advanced ant societies**. *Proc Natl Acad Sci U S A* 2008,
853 **105**(46):17884-17889.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

854 23. Mito T, Nakamura T, Sarashina I, Chang CC, Ogawa S, Ohuchi H, Noji S: **Dynamic**
855 **expression patterns of vasa during embryogenesis in the cricket Gryllus bimaculatus**.
856 *Dev Genes Evol* 2008, **218**(7):381-387.
857 24. Tribolium Genome Sequencing C, Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell
858 R, Beeman RW, Gibbs R, Beeman RW, Brown SJ *et al*: **The genome of the model beetle**
859 **and pest Tribolium castaneum**. *Nature* 2008, **452**(7190):949-955.
860 25. Consortium TIAG: **Genome Sequence of the Pea Aphid Acyrthosiphon pisum**. *PLoS Biol*
861 2010.
862 26. Khila A, Abouheif E: **Evaluating the role of reproductive constraints in ant social**
863 **evolution**. *Philos Trans R Soc Lond B Biol Sci* 2010, **365**(1540):617-630.
864 27. Shigenobu S, Bickel RD, Brisson JA, Butts T, Chang CC, Christiaens O, Davis GK, Duncan EJ,
865 Ferrier DE, Iga M *et al*: **Comprehensive survey of developmental genes in the pea aphid,**
866 **Acyrthosiphon pisum: frequent lineage-specific duplications and losses of**
867 **developmental genes**. *Insect Mol Biol* 2010, **19 Suppl 2**:47-62.
868 28. Bontems F, Stein A, Marlow F, Lyautey J, Gupta T, Mullins MC, Dosch R: **Bucky ball**
869 **organizes germ plasm assembly in zebrafish**. *Curr Biol* 2009, **19**(5):414-422.
870 29. Lehmann R, Nussleinvolhard C: **Abdominal Segmentation, Pole Cell-Formation, and**
871 **Embryonic Polarity Require the Localized Activity of Oskar, a Maternal Gene in**
872 **Drosophila**. *Cell* 1986, **47**(1):141-152.
873 30. Ephrussi ALKD, L. K.; Lehmann, R.: **oskar organizes the germ plasm and directs**
874 **localization of the posterior determinant nanos**. *Cell* 1991, **66**(1):14.
875 31. Ephrussi A, Lehmann R: **Induction of Germ-Cell Formation by Oskar**. *Nature* 1992,
876 **358**(6385):387-392.
877 32. Rongo CG, E. R.; Lehmann, R.: **Localization of oskar RNA regulates oskar translation and**
878 **requires Oskar protein**. *Development* 1995, **121**(9):2737-2746.
879 33. Breitwieser W, Markussen, F. H., Horstmann, H., Ephrussi, A.: **Oskar protein interaction**
880 **with Vasa represents an essential step in polar granule assembly**. *Genes &*
881 *Development* 1996, **10**:2179-2188.
882 34. Boswell RE, Mahowald AP: **Tudor, a Gene Required for Assembly of the Germ Plasm in**
883 **Drosophila-Melanogaster**. *Cell* 1985, **43**(1):97-104.
884 35. Lynch JA, Ozuak O, Khila A, Abouheif E, Desplan C, Roth S: **The phylogenetic origin of**
885 **oskar coincided with the origin of maternally provisioned germ plasm and pole cells at**
886 **the base of the Holometabola**. *PLoS genetics* 2011, **7**(4):e1002029.
887 36. Jaglarz MK, Howard, K. R.: **The active migration of Drosophila primordial germ cells**.
888 *Development* 1995, **121**:3495-3503.
889 37. Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, Rakheja D, Chen KS, Hammer RE,
890 Comerford SA *et al*: **p53 genes function to restrain mobile elements**. *Genes Dev* 2016,
891 **30**(1):64-77.
892 38. Tiwari B, Kurtz P, Jones AE, Wylie A, Amatruda JF, Boggupalli DP, Gonsalvez GB, Abrams
893 JM: **Retrotransposons Mimic Germ Plasm Determinants to Promote Transgenerational**
894 **Inheritance**. *Curr Biol* 2017, **27**(19):3010-3016 e3013.
895 39. Bantock CR: **Experiments on chromosome elimination in the gall midge, Mayetiola**
896 **destructor**. *Journal of Embryology and Experimental Morphology* 1970, **24**:257-289.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

897 40. Lynch JA, Desplan C: **Novel modes of localization and function of nanos in the wasp**
898 **Nasonia**. *Development* 2010, **137**(22):3813-3821.
899 41. Santos AC, Lehmann R: **Germ cell specification and migration in Drosophila and beyond**.
900 *Curr Biol* 2004, **14**(14):R578-589.
901 42. Draper BWM, C. C.; Bowerman, B.; Hardin, J.; Priess, J. R. : **MEX-3 Is a KH Domain**
902 **Protein That Regulates Blastomere Identity in Early C. elegans Embryos**. *Cell* 1996,
903 **87**:205-216.
904 43. Lynch J, Pers D: **Ankyrin domain encoding genes resulting from an ancient horizontal**
905 **transfer are functionally integrated into developmental gene regulatory networks in**
906 **the wasp Nasonia**. *bioRxiv* 2018.
907 44. Lynch JA, Brent AE, Leaf DS, Pultz MA, Desplan C: **Localized maternal orthodenticle**
908 **patterns anterior and posterior in the long germ wasp Nasonia**. *Nature* 2006,
909 **439**(7077):728-732.
910 45. Byri S, Misra T, Syed ZA, Batz T, Shah J, Boril L, Glashauser J, Aegerter-Wilmsen T, Matzat
911 T, Moussian B *et al*: **The Triple-Repeat Protein Anakonda Controls Epithelial Tricellular**
912 **Junction Formation in Drosophila**. *Dev Cell* 2015, **33**(5):535-548.
913 46. Hildebrandt A, Pflanz R, Behr M, Tarp T, Riedel D, Schuh R: **Bark beetle controls**
914 **epithelial morphogenesis by septate junction maturation in Drosophila**. *Dev Biol* 2015,
915 **400**(2):237-247.
916 47. Jennings BH, Shah S, Yamaguchi Y, Seki M, Phillips RG, Handa H, Ish-Horowicz D: **Locus-**
917 **specific requirements for Spt5 in transcriptional activation and repression in**
918 **Drosophila**. *Curr Biol* 2004, **14**(18):1680-1684.
919 48. Pers D, Buchta T, Ozuak O, Wolff S, Pietsch JM, Memon MB, Roth S, Lynch JA: **Global**
920 **analysis of dorsoventral patterning in the wasp Nasonia reveals extensive**
921 **incorporation of novelty in a regulatory network**. *BMC Biol* 2016, **14**:63.
922 49. Rago A, Gilbert DG, Choi JH, Sackton TB, Wang X, Kelkar YD, Werren JH, Colbourne JK:
923 **OGS2: genome re-annotation of the jewel wasp Nasonia vitripennis**. *BMC Genomics*
924 2016, **17**:678.
925 50. Field CM, Alberts BM: **Anillin, a Contractile Ring Protein That Cycles from the Nucleus**
926 **to the Cell Cortex**. *J Cell Biol* 1995, **131**(1):165-178.
927 51. Cinalli RM, Lehmann R: **A spindle-independent cleavage pathway controls germ cell**
928 **formation in Drosophila**. *Nature cell biology* 2013, **15**(7):839-845.
929 52. Anand A, Kai T: **The tudor domain protein kumo is required to assemble the nuage and**
930 **to generate germline piRNAs in Drosophila**. *EMBO J* 2012, **31**(4):870-882.
931 53. Zhang Z, Koppetsch BS, Wang J, Tipping C, Weng Z, Theurkauf WE, Zamore PD:
932 **Antisense piRNA amplification, but not piRNA production or nuage assembly, requires**
933 **the Tudor-domain protein Qin**. *EMBO J* 2014, **33**(6):536-539.
934 54. Anand A, Kai T: **Response to Zhang et al**. *EMBO J* 2014, **33**(6):540-541.
935 55. Lecuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR,
936 Tomancak P, Krause HM: **Global analysis of mRNA localization reveals a prominent role**
937 **in organizing cellular architecture and function**. *Cell* 2007, **131**(1):174-187.
938 56. Pan J, Goodheart M, Chuma S, Nakatsuji N, Page DC, Wang PJ: **RNF17, a component of**
939 **the mammalian germ cell nuage, is essential for spermiogenesis**. *Development* 2005,
940 **132**(18):4029-4039.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

941 57. Watanabe T, Kankel, D. R.: **Molecular Cloning and Analysis oIf(1)ogre, a Locus of**
942 **Drosophila melanogaster With Prominent Effects on the Postembryonic Development**
943 **of the Central Nervous System**. *Genetics* 1990, **126**:1033-1044.
944 58. van der Zee M, Benton MA, Vazquez-Faci T, Lamers GE, Jacobs CG, Rabouille C:
945 **Innexin7a forms junctions that stabilize the basal membrane during cellularization of**
946 **the blastoderm in Tribolium castaneum**. *Development* 2015, **142**(12):2173-2183.
947 59. Olesnicky EC, Brent AE, Tonnes L, Walker M, Pultz MA, Leaf D, Desplan C: **A caudal**
948 **mRNA gradient controls posterior development in the wasp Nasonia**. *Development*
949 2006, **133**(20):3973-3982.
950 60. Ozuak O, Buchta T, Roth S, Lynch JA: **Dorsoventral polarity of the Nasonia embryo**
951 **primarily relies on a BMP gradient formed without input from Toll**. *Curr Biol* 2014,
952 **24**(20):2393-2398.
953 61. Riesgo-Escovar JR, Hafen E: **Common and distinct roles of DFos and DJun during**
954 **Drosophila development**. *Science* 1997, **278**(5338):669-672.
955 62. Dorfman R, Glazer L, Weihe U, Wernet MF, Shilo BZ: **Elbow and Noc define a family of**
956 **zinc finger proteins controlling morphogenesis of specific tracheal branches**.
957 *Development* 2002, **129**(15):3585-3596.
958 63. Wernet MF, Meier KM, Baumann-Klausener F, Dorfman R, Weihe U, Labhart T, Desplan
959 C: **Genetic dissection of photoreceptor subtype specification by the Drosophila**
960 **melanogaster zinc finger proteins elbow and no ocelli**. *PLoS genetics* 2014,
961 **10**(3):e1004210.
962 64. Rybakin V, Clemen CS: **Coronin proteins as multifunctional regulators of the**
963 **cytoskeleton and membrane trafficking**. *Bioessays* 2005, **27**(6):625-632.
964 65. Gandhi M, Goode BL: **Coronin: the double-edged sword of actin dynamics**. *Subcell*
965 *Biochem* 2008, **48**:72-87.
966 66. Glater EE, Megeath LJ, Stowers RS, Schwarz TL: **Axonal transport of mitochondria**
967 **requires milton to recruit kinesin heavy chain and is light chain independent**. *J Cell Biol*
968 2006, **173**(4):545-557.
969 67. Tanaka T, Nakamura A: **The endocytic pathway acts downstream of Oskar in**
970 **Drosophila germ plasm assembly**. *Development* 2008, **135**(6):1107-1117.
971 68. Jaspers MH, Pflanz R, Riedel D, Kawelke S, Feussner I, Schuh R: **The fatty acyl-CoA**
972 **reductase Waterproof mediates airway clearance in Drosophila**. *Dev Biol* 2014,
973 **385**(1):23-31.
974 69. Buchta T, Ozuak O, Stappert D, Roth S, Lynch JA: **Patterning the dorsal-ventral axis of**
975 **the wasp Nasonia vitripennis**. *Dev Biol* 2013, **381**(1):189-202.
976 70. Ozuak O, Buchta T, Roth S, Lynch JA: **Ancient and diverged TGF-beta signaling**
977 **components in Nasonia vitripennis**. *Dev Genes Evol* 2014, **224**(4-6):223-233.
978 71. Tomancak P, Berman BP, Beaton A, Weiszmann R, Kwan E, Hartenstein V, Celniker SE,
979 Rubin GM: **Global analysis of patterns of gene expression during Drosophila**
980 **embryogenesis**. *Genome Biol* 2007, **8**(7):R145.
981 72. Lee DM, Wilk R, Hu J, Krause HM, Harris TJ: **Germ Cell Segregation from the Drosophila**
982 **Soma Is Controlled by an Inhibitory Threshold Set by the Arf-GEF Steppke**. *Genetics*
983 2015, **200**(3):863-872.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

984 73. Martinson EO, Mrinalini, Kelkar YD, Chang CH, Werren JH: **The Evolution of Venom by**
985 **Co-option of Single-Copy Genes**. *Curr Biol* 2017, **27**(13):2007-2013 e2008.
986 74. Hanyu-Nakamura K, Sonobe-Nojima H, Tanigawa A, Lasko P, Nakamura A: **Drosophila**
987 **Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells**. *Nature*
988 2008, **451**(7179):730-733.
989 75. Nakamura A, Amikura R, Mukai M, Kobayashi S, Lasko PF: **Requirement for a noncoding**
990 **RNA in Drosophila polar granules for germ cell establishment**. *Science* 1996,
991 **274**(5295):2075-2079.
992 76. Zhang F, Barboric M, Blackwell TK, Peterlin BM: **A model of repression: CTD analogs and**
993 **PIE-1 inhibit transcriptional elongation by P-TEFb**. *Genes Dev* 2003, **17**(6):748-758.
994 77. Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H: **P-TEFb-mediated**
995 **phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription**
996 **elongation**. *Mol Cell* 2006, **21**(2):227-237.
997 78. Lerit DA, Shebelut CW, Lawlor KJ, Rusan NM, Gavis ER, Schedl P, Deshpande G: **Germ**
998 **Cell-less Promotes Centrosome Segregation to Induce Germ Cell Formation**. *Cell Rep*
999 2017, **18**(4):831-839.
1000 79. Amarnath S, Stevens LM, Stein DS: **Reconstitution of Torso signaling in cultured cells**
1001 **suggests a role for both Trunk and Torso-like in receptor activation**. *Development* 2017,
1002 **144**(4):677-686.
1003 80. Guraya SS: **Recent advances in the morphology, cytochemistry, and function of**
1004 **Balbiani's vitelline body in animal oocytes**. *Int Rev Cytol* 1979, **59**:249-321.
1005 81. Kloc M, Bilinski S, Etkin LD: **The Balbiani body and germ cell determinants: 150 years**
1006 **later**. *Curr Top Dev Biol* 2004, **59**:1-36.
1007 82. Hurd TR, Herrmann B, Sauerwald J, Sanny J, Grosch M, Lehmann R: **Long Oskar Controls**
1008 **Mitochondrial Inheritance in Drosophila melanogaster**. *Dev Cell* 2016, **39**(5):560-571.
1009 83. Cox RT, Spradling AC: **Milton controls the early acquisition of mitochondria by**
1010 **Drosophila oocytes**. *Development* 2006, **133**(17):3371-3377.
1011 84. Patil VS, Anand A, Chakrabarti A, Kai T: **The Tudor domain protein Tapas, a homolog of**
1012 **the vertebrate Tdrd7, functions in the piRNA pathway to regulate retrotransposons in**
1013 **germline of Drosophila melanogaster**. *BMC Biol* 2014, **12**:61.
1014 85. Gao M, Thomson TC, Creed TM, Tu S, Loganathan SN, Jackson CA, McCluskey P, Lin Y,
1015 Collier SE, Weng Z *et al*: **Glycolytic enzymes localize to ribonucleoprotein granules in**
1016 **Drosophila germ cells, bind Tudor and protect from transposable**
1017 **elements**. *EMBO reports* 2015, **16**(3):379-386.
1018 86. Teng H, Sui X, Zhou C, Shen C, Yang Y, Zhang P, Guo X, Huo R: **Fatty acid degradation**
1019 **plays an essential role in proliferation of mouse female primordial germ cells via the**
1020 **p53-dependent cell cycle regulation**. *Cell Cycle* 2016, **15**(3):425-431.
1021 87. Pao SS, Paulsen IT, Saier MH: **Major facilitator superfamily**. *Microbiol Mol Biol R* 1998,
1022 **62**(1):1-+.
1023 88. Yamada Y, Davis KD, Coffman CR: **Programmed cell death of primordial germ cells in**
1024 **Drosophila is regulated by p53 and the Outsiders monocarboxylate**
1025 **transporter**. *Development* 2008, **135**(2):207-216.

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1026 89. Jaspers MHJ, Pflanz R, Riedel D, Kawelke S, Feussner I, Schuh R: **The fatty acyl-CoA**
1027 **reductase Waterproof mediates airway clearance in Drosophila**. *Developmental*
1028 *Biology* 2014, **385**(1):23-31.
1029 90. Hildebrandt A, Pflanz R, Behr M, Tarp T, Riedel D, Schuh R: **Bark beetle controls**
1030 **epithelial morphogenesis by septate junction maturation in Drosophila**. *Dev Biol* 2015,
1031 **400**(2):237-247.
1032 91. Tanaka T, Kato Y, Matsuda K, Hanyu-Nakamura K, Nakamura A: **Drosophila Mon2**
1033 **couples Oskar-induced endocytosis with actin remodeling for cortical anchorage of the**
1034 **germ plasm**. *Development* 2011, **138**(12):2523-2532.
1035 92. Lerit DA, Gavis ER: **Transport of germ plasm on astral microtubules directs germ cell**
1036 **development in Drosophila**. *Curr Biol* 2011, **21**(6):439-448.
1037 93. Ferree PM, McDonald K, Fasulo B, Sullivan W: **The origin of centrosomes in**
1038 **parthenogenetic hymenopteran insects**. *Curr Biol* 2006, **16**(8):801-807.
1039 94. van Oostende Triplet C, Jaramillo Garcia M, Haji Bik H, Beaudet D, Piekny A: **Anillin**
1040 **interacts with microtubules and is part of the astral pathway that defines cortical**
1041 **domains**. *J Cell Sci* 2014, **127**(Pt 17):3699-3710.
1042 95. Tse YC, Piekny A, Glotzer M: **Anillin promotes astral microtubule-directed cortical**
1043 **myosin polarization**. *Mol Biol Cell* 2011, **22**(17):3165-3175.
1044 96. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL: **TopHat2: accurate**
1045 **alignment of transcriptomes in the presence of insertions, deletions and gene fusions**.
1046 *Genome Biol* 2013, **14**(4):R36.
1047 97. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL,
1048 Pachter L: **Differential gene and transcript expression analysis of RNA-seq experiments**
1049 **with TopHat and Cufflinks**. *Nat Protoc* 2012, **7**(3):562-578.
1050 98. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL: **StringTie**
1051 **enables improved reconstruction of a transcriptome from RNA-seq reads**. *Nat*
1052 *Biotechnol* 2015, **33**(3):290-295.
1053 99. Brent AE, Schweitzer R, Tabin CJ: **A somitic compartment of tendon progenitors**. *Cell*
1054 2003, **113**(2):235-248.
1055

1056

1057 **Additional Files**

1058 **Additional File 1. (.xlsx)** Significantly enriched transcripts from both experiments. The
1059 transcripts listed in "Compilation of Transcripts significant in the main analyses" were generated
1060 using the sequencing data from the second experiment by NCBI annotation 102 with various
1061 normalization parameters. The "Transcripts from the second experiment only found in OGS 2.0"
1062 were the transcripts from the second experiment only found using the annotation OGS 2.0, but

bioRxiv preprint doi: https://doi.org/10.1101/384032; this version posted August 25, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1063 not found using NCBI annotation 102. The "Transcripts from the first experiment" were the
1064 transcripts only found in the first time sequencing data, but not in the second time sequencing
1065 data.

1066 **Additional File 2. (.docx)** Command jobs. The command jobs for analyses using the NCBI
1067 annotation 102 and the annotation OGS 2.0 with various normalization parameters.

1068 **Additional File 3. (.xlsx)** Original lists of the transcripts generated by Cuffdiff. This file
1069 includes the original lists of the transcripts generated by Cuffdiff using the NCBI annotation 102
1070 and the annotation OGS 2.0 with various normalization parameters.

1071 **Additional File 4. (.xlsx)** Primer list. All primers used in this manuscript are listed. This list
1072 includes the primers used for cloning the genes and making templates for the probes and
1073 dsRNAs.

1074
1075
1076
1077
1078