bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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 **Held out wings RNA binding activity in the cytoplasm during early spermatogenesis**

2 Michaela Agapiou[1,2], Tayah Hopes[1,2], Fruzsina Hobor [1,3], Amanda Bretman[4], Thomas A.
3 Edwards[1,3] and Julie L. Aspden[1,2] #
4
5 [1] School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds,
6 LS2 9JT, UK.
7 [2] LeedsOmics, University of Leeds, UK
8 [3] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK.
9 [4] School of Biology, Faculty of Biological Sciences, University of Leeds, LS2 9JT, UK
10 #corresponding author: j.aspden@leeds.ac.uk
11
12 Abstract

13 Held out wings (HOW) is an RNA-binding protein essential for spermatogenesis in *Drosophila*
14 *melanogaster*. HOW is a signal transduction and activation of RNA (STAR) protein, regulating
15 post-transcriptional gene expression. The characteristics of RNA-binding by the conserved
16 short cytoplasmic isoform, HOW(S), are unknown. *In vivo* RIP-seq identified 121 novel
17 transcripts bound by HOW(S) in germ stem cells and spermatogonia, many with signal
18 transduction functions. (A/G/U)CUAAC motifs were enriched in 3'-UTRs and GCG(A/U)G in
19 5'-UTRs. HOW binds with high-affinity to sites containing CUAAC motifs from *lola* and *hipk*
20 mRNAs. This study provides new insight into STAR protein-RNA interactions and functions in
21 spermatogenesis.
22
23
24

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

25 Introduction

26 Post-transcriptional gene regulation plays an essential role in stem cells contributing to
27 pluripotency, differentiation and self-renewal. Molecular signals contribute to germ stem cell
28 regulation through a variety of RNA processes controlled by a network of RNA-binding
29 proteins including the signal transduction and activation of RNA (STAR) protein family. STAR
30 proteins characteristically contain a single RNA-binding STAR domain, composed of one maxi-
31 K-homology (KH) domain, and one or two flanking regions (Vernet & Artzt, 1997). As a result
32 of their functions in signal transduction pathways, the family contributes to the regulation of
33 gene expression during key stages of development. For example, the STAR protein Quaking
34 (QKI) regulates RNA processing during gametogenesis and myelination in mammals (Kondo
35 et al., 1999). RNA-binding proteins in general are important throughout spermatogenesis,
36 including for stem cell maintenance and meiosis, across a variety of organisms (Legrand &
37 Hobbs, 2018).
38
39 Held out wings (HOW) is the ortholog of QKI in *Drosophila melanogaster.* HOW is essential
40 but partial loss of function mutants exhibit a variety of phenotypes including defective wing
41 development, glial maturation, tendon maturation and sterility (Baehrecke, 1997; Edenfeld
42 et al., 2006; Monk et al., 2010; Nabel-Rosen et al., 1999). HOW regulates RNA processing,
43 including pre-mRNA splicing, localisation and mRNA degradation (Graindorge et al., 2013;
44 Nabel-Rosen et al., 2002; Rodrigues et al., 2012). There are multiple protein isoforms of HOW,
45 produced through alternative splicing, all of which contain the STAR domain but differing at
46 the C-terminus (Larkin et al., 2021). The different isoforms of QKI and HOW are differentially
47 localised within cells, and the shortest isoforms, QKI-6 and HOW(S), are predominantly
48 cytoplasmic, where their function is unknown. The HOW and QKI ortholog in *C. elegans*, GLD-

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

49 1, has a single cytoplasmic isoform that regulates mRNA translation in germ cells (Jan et al.,
50 1999). Mirroring QKI's expression in mammals, HOW is expressed in *D. melanogaster* testis,
51 and it is important during spermatogenesis (Monk et al., 2010). In germ cells, HOW's
52 expression is unusual and highly restricted to the earliest stages of gametogenesis, where it
53 contributes to the maintenance of germ stem cells (GSCs); without HOW, GSCs in the testis
54 do not survive (Monk et al., 2011; Monk et al., 2010). Expression of HA-tagged HOW(S) in the
55 testis indicates that this isoform is restricted to the cytoplasm in germ cells (Monk et al., 2010).
56 The restricted expression of HOW during spermatogenesis suggests a specific function in
57 regulating the balance between stem cell maintenance versus the proliferation and
58 differentiation of spermatogonia.
59
60 The majority of work on HOW thus far has focused on HOW(L) and its role in nuclear RNA
61 processing events. Current understanding of HOW RNA binding comes from a handful of
62 mRNAs (e.g. *dpp, miple1, bam*), which contain 3'-UTR ACUAA motifs typically bound by
63 HOW(L) (Israeli et al., 2007). The only mRNA previously known to be bound specifically by
64 HOW(S) rather than HOW(L) is *dgrasp* in oocytes (Giuliani et al., 2014). An optimal binding
65 preference has yet to be characterised for HOW but a consensus of NCUAACN has been
66 generated from *in vitro* binding experiments (Ray et al., 2013). Here, we sought to globally
67 identify novel mRNA targets of HOW in the testis, specifically by HOW(S) in the cytoplasm, to
68 understand the binding characteristic of this interaction *in vivo*.
69
70
71
72

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

73 ## Results & Discussion

74 **Expression and pull-down of cytoplasmic HA-tagged HOW(S) in testis germ cells**

75 To understand the role of HOW(S) and the RNAs it binds in the cytoplasm, HA-tagged HOW(S)
76 was expressed in the early stages of spermatogenesis in *Drosophila melanogaster* testes. This
77 was achieved by using the UAS/GAL4 system, whereby HOW(S)-HA expression was driven by
78 *nanos*-GAL4, and thus, HOW(S)-HA was expressed in GSCs and early spermatogonia (Fig. 1A).
79 This HA-tagged HOW(S) is localised to the cytoplasm of these cells (Fig. 1A), recapitulating the
80 presence of cytoplasmic HOW in these specific germ cells (Monk et al., 2011; Monk et al.,
81 2010). We were able to successfully pull down HOW(S)-HA from the lysates of these testes
82 using anti-HA beads (Fig. 1B). To isolate RNA bound to HOW(S), RNP-IPs were performed from
83 large scale testes lysates in triplicate (Fig. 1C and Supplemental Fig. 1A-C). RNA from the
84 HOW(S)-HA pull-down was purified along with RNA from input testis lysates and RNA from
85 pull-downs performed from the *nanos*-GAL4 parent testes, which had no HA-tagged HOW(S)
86 (Supplemental Fig. 1D). These RNA samples then underwent RNA-seq.

87

88

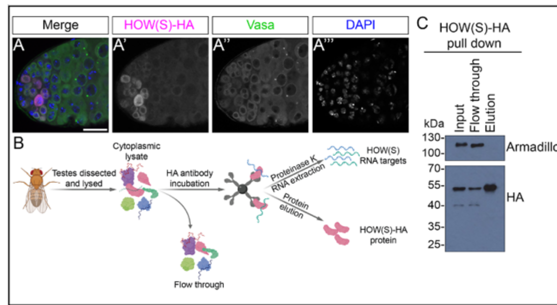
89

90

91

92

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.



93

94 *Figure 1: Expression and pull-down of HA-tagged HOW(S) from cytoplasm of germ cells*

95 *A) Confocal images showing cytoplasmic expression of HOW(S)-HA (magenta)in GSCs and*

96 *spermatogonia, driven by the nanos-GAL4 line. Germ cells are positive for vasa (green), and*

97 *blue is DAPI. B) Schematic of HOW(S)-HA RIP-seq. C) Western blot showing specific depletion*

98 *of HOW(S)-HA from testis the input lysate and enrichment in the elution using anti-HA*

99 *beads.*

100

101 **HOW(S) bound RNAs are enriched for signalling functions**

102 To identify which RNAs were specifically bound to HOW(S)-HA, differential transcript

103 enrichment analysis was performed on RNAs in HOW(S)-HA pulldown compared to those

104 present in input lysate (Fig. 2A, Supplemental Table 1, Supplemental Fig. 2A-B). The majority

105 of transcripts showed no enrichment, having a log_2(Fold Change) between -0.5 and 0.5 (Fig.

106 2B). We defined enriched transcripts as those with a log_2(Fold Change) ≥ 1 and an FDR-

107 corrected *p*-value ≤ 0.05. Through this analysis we identified 403 transcripts enriched in the

108 HOW(S)-HA pull-down RNA compared with input RNA. Non-specific background RNAs

109 enriched in the parental *nanos*-GAL4 RIP-seq sample (with no HOW(S)-HA present) were then

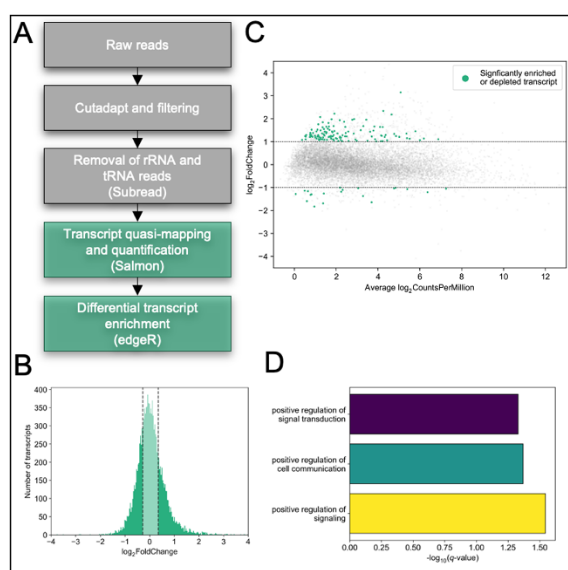
110 subtracted, leaving those RNA transcripts specifically bound to HOW(S)-HA. 121 mRNAs were

111 identified as specifically enriched by HOW(S) RIP-seq (Fig. 2C – green dots). To understand the

112 role of the RNAs bound by HOW(S), GO analysis was performed on the genes corresponding

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

to these HOW(S) bound mRNAs. This revealed an enrichment for genes with functions in cell signalling, specifically in cell communication and signal transduction e.g. *hipk* (homeodomain interacting protein kinase) and *CycG* (cyclin G) (Fig. 2D, Supplemental Fig. 2C). This is entirely consistent with HOW's membership of the STAR family of proteins.



<u>Figure 2: Enrichment of HOW(S) bound RNA, with role in signal transduction</u>

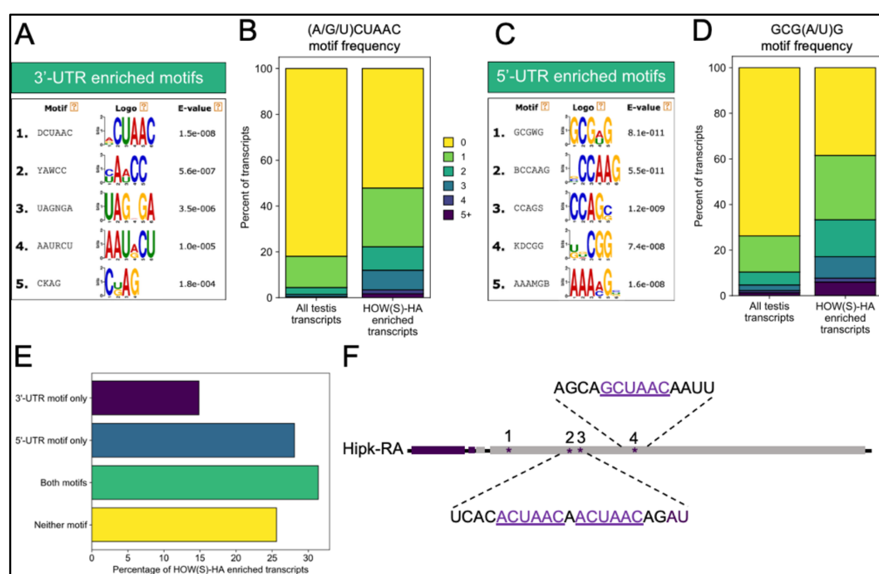
A) Pipeline to identify RNAs specifically bound to HOW(S). B) Histogram from differential enrichment analysis of pull-down versus input, middle quartiles are shaded in lighter green. C) Scatter plot from differential transcript analysis. Significantly enriched or depleted transcripts are those with log_2(Fold Change) of ≥ 1 or ≤ -1 (dotted lines), and an FDR-corrected p-value ≤ 0.05, additionally non-specific background RNAs enriched in the parental nanos-GAL4 RIP-seq sample were subtracted. Specifically enriched or depleted RNAs from the HOW(S) RIP-seq are highlighted in green. D) GO term analysis reveals over-representation of signaling related terms in transcripts bound by HOW(S).

HOW's ortholog QKI is a known regulator of circRNA processing (Conn et al., 2015), and while our method was not optimised for circRNA detection, further analysis revealed a

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

130 small number of circRNAs in our dataset, some of which HOW(S) may bind (Supplemental Fig.
131 2D and Supplemental Table 2).
132
133 **CUAAC motif enriched in 3'-UTRs of HOW(S) bound RNAs**
134 To determine the nature of HOW(S) binding to the 121 transcripts identified by RIP-seq, motif
135 analysis was performed on these mRNA targets. Within the 3'-UTRs the most enriched motif
136 was found to be (A/G/U)CUAAC (Fig. 3A), with 46% of the HOW(S)-HA bound mRNAs
137 containing this motif (Fig. 3B). (A/G/U)CUAAC is very similar to the other previously identified
138 HOW binding sequence in *stripe* and *dpp* mRNAs: ACUAA (Israeli et al., 2007). Rather than
139 characterising a small number of HOW targets and identifying individual binding sites, the
140 (A/G/U)CUAAC motif generated here is from a much larger number of RNAs bound by HOW(S)
141 in the cytoplasm.
142

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

143 

144 *Figure 3: Enrichment of CUAAC motifs in HOW(S) bound RNAs*

145 *A) Motifs found to be enriched in 3'-UTRs of mRNAs bound by HOW(S). B) Frequency at*
146 *which (A/G/U)CUAAC motif identified in 3'-UTR of HOW(S) bound transcripts compared to*
147 *the 3'-UTRs of all testis expressed transcripts. C) Motifs found to be enriched in 5'-UTRs of*
148 *mRNAs bound by HOW(S). D) Frequency at which GCG(A/U)G motif identified in 5'-UTR of*
149 *HOW(S) bound transcripts compared to the 5'-UTRs of all testis expressed transcripts. E)*
150 *Percentage of transcripts bound by HOW(S) that had either of the top UTR motifs, both, or*
151 *neither in their respective UTRs. F) Schematic of the 3' end of hipk transcript, purple is CDS,*
152 *grey is UTR, and asterisks mark (A/G/U)CUAAC sites in the 3'-UTR.*

153

154 Other distinct motifs were found enriched in the 5'-UTRs of HOW(S) bound mRNAs,
155 including GCG(A/U)G (Fig. 3C), but these do not show similarity with previously identified
156 HOW binding sites (Graindorge et al., 2013; Israeli et al., 2007; Ray et al., 2013). Although
157 these are unlikely to be HOW(S) binding sites, because their sequence differs dramatically
158 from the previously characterised ACUAA motif, these motifs may be important for HOW(S)
159 RNA-protein complexes as they are highly enriched in the HOW(S) bound mRNAs – the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

160 GCG(A/U)G motif is present in 60% of HOW(S) bound mRNAs (Fig. 3D). Instead, these motifs
161 may represent sequences that are bound to another RNA-binding protein (RBP) associated
162 with HOW(S), since RIP-seq will identify both indirect and direct interactions. This type of co-
163 binding has been demonstrated for HOW(L), which can interact with the RBP SXL, with the
164 two proteins binding distinct sequences in the 5'-UTR of *msl-2* (Graindorge et al., 2013). We
165 also observed that 31% of HOW(S) bound mRNAs contain enriched motifs both in their 5'-
166 and 3'-UTRs (Fig. 3E), which indicates that HOW(S) molecules could bind in both UTRs of an
167 mRNA transcript simultaneously, potentially regulating the mRNA in numerous ways. This has
168 been shown for other RBPs during development e.g. SXL during sex determination (Duncan
169 et al., 2006).

170 Further analysis of the HOW(S) bound transcripts that have the (A/G/U)CUAAC motif
171 reveals that 21% of them contain multiple repeats of this motif, which is substantially higher
172 than that found in all testis expressed transcripts (Fig. 3B). RBPs often bind RNAs in a modular
173 manner, and HOW, along with many of the STAR proteins, dimerise (Nir et al., 2012). Thus,
174 multiple binding sites in HOW(S) bound transcripts could indicate a modular binding mode.
175 This pattern is also seen for the top 5'-UTR enriched motif GCG(A/U)G (Fig. 3D), indicating
176 that this too may represent a modular RBP binding region.

177 One such mRNA which has multiple predicted binding sites is *hipk*. Hipk protein
178 functions in multiple signalling pathways, including Notch (Lee et al., 2009), Hippo (Chen &
179 Verheyen, 2012; Poon et al., 2012), and the JAK-STAT pathway (Tettweiler et al., 2019) and is
180 essential for germline development in *C. elegans* and mouse spermatogenesis (Berber et al.,
181 2013; Crapster et al., 2020). *Hipk-RA* transcript (FBtr0072552) was identified as enriched in
182 the HOW(S)-HA RIP-seq and contains four (A/G)CUAAC sites within its 3'-UTR (Fig. 3F),

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

183 therefore likely represents a key HOW(S) target with downstream roles in signalling pathways
184 essential in maintaining the stem cell niche and cell fate.
185

186 **Recombinant HOW-STAR domain binds enriched motifs within HOW(S) bound RNAs with**
187 **high affinity**

188 To test the binding of HOW to the motifs identified as enriched from RIP-seq we
189 recombinantly expressed HOW's STAR domain (Fig. 4A, Supplemental Fig. 3A-C) and
190 performed fluorescence anisotropy (FA) binding assays. Three different mRNAs that we
191 identified as being bound by HOW(S) in GSCs and early spermatogonia were selected to test
192 HOW binding to sequences from within these transcripts that were identified as highly similar
193 to the enriched motifs.
194

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

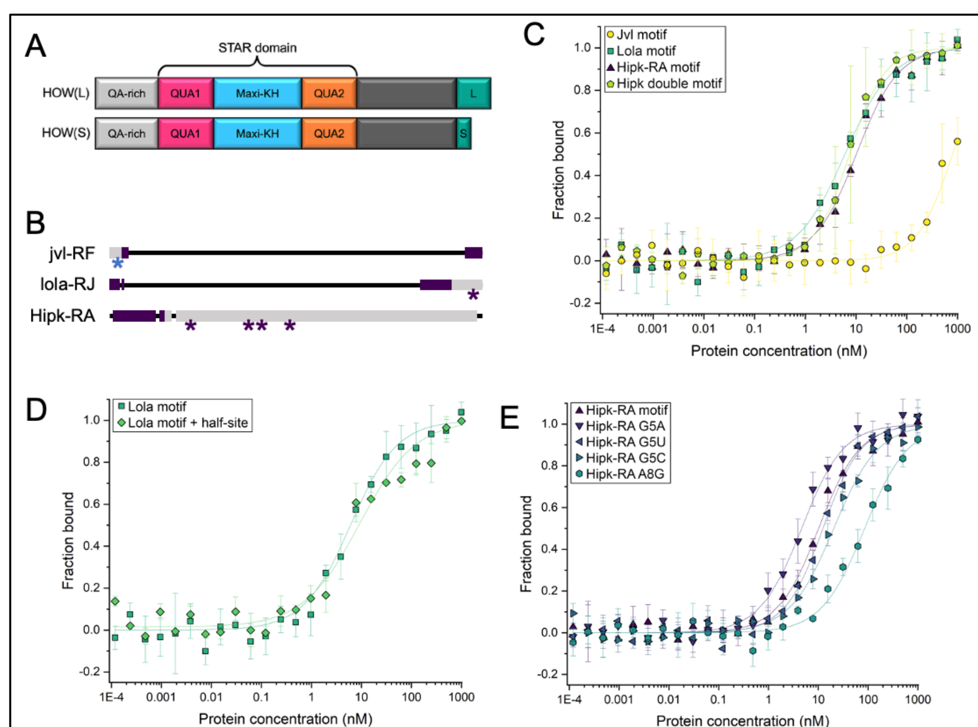


Figure 4: HOW KH domain binds to HOW(S) targets with high affinity and specificity

A) Schematic of two HOW protein isoforms, indicating recombinant protein STAR domain expressed. B) Schematic of three transcripts bound by HOW(S): the 5' end of jvl-RF, and 3' ends of lola-RF and Hipk-RA. Asterisks mark the 5'-UTR (blue) and 3'-UTR (purple) motifs. Purple bars are CDS, grey is UTRs. C) FA binding plots from jvl motif, lola motif and Hipk-RA motif, as well as the Hipk double site. D) lola motif and lola motif + half-site. E) Mutations in the Hipk-RA motif.

A site was selected from the 5'-UTR of *jvl* transcript (FBtr0305694) that contained the most highly enriched 5'-UTR motif within HOW(S) bound mRNAs: GCGUG (Fig. 4B). FA analysis of this *jvl* RNA oligo revealed that HOW-STAR did not bind to this site in the low nanomolar range (Fig. 4C & Table 1). This suggests that although the GCG(A/U)G motif is highly enriched in the 5'-UTRs of HOW(S) bound mRNAs, it is unlikely to be an *in vivo* HOW(S) binding site. An alternative explanation could be that HOW(S) interacts with a second RBP, which provides the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

RNA-binding specificity for these interactions. One potential candidate is SLIRP1 protein, which has a binding motif that is highly similar to the one found in the 5'-UTR of HOW(S) bound RNAs: GCG(U/C)(G>A/C/U) (Ray et al., 2013) and is highly expressed in the testis.

Oligo name	**Sequence**	**Apparent K_D (nM) to 3 sig. fig.**	**Standard error (±)**
Jvl motif	GUUGGCGUGUUUU	803	1064
Lola motif	ACACACUAACUCGU	6.05	0.62
Lola motif + half-site	ACACACUAACUCGUAACUAUG	9.25	1.57
Hipk-RA motif	AGCAGCUAACAAUU	10.1	0.73
Hipk-RA G5A	AGCA**A**CUAACAAUU	4.65	0.56
Hipk-RA G5U	AGCA**U**CUAACAAUU	12.4	1.25
Hipk-RA G5C	AGCA**C**CUAACAAUU	20.4	2.86
Hipk-RA A8G	AGCAGCU**G**ACAAUU	85.3	22.1
Hipk double motif	UACAACUAACAACUAACAGAU	7.46	1.76

Table 1: Summary of HOW binding affinities for HOW(S) predicted binding sites

Sequence of oligos based on the predicted binding sites within HOW(S) targets, along with apparent K_D and standard error calculated from FA. Sequences that match the top motifs from the motif enrichment analysis are underlined, and nucleotides mutated are in bold.

We also tested RNA oligos based on regions from *lola* and *hipk*, two transcripts bound by HOW(S), which contain sequences matching the 3'-UTR HOW(S) motif. The *lola* element contains an ACUAAC motif and HOW-STAR exhibits high affinity binding to this RNA oligo by FA (Fig. 4C, Supplemental Fig. 3D). Just downstream of this motif is a potential half-site and such half-sites have been suggested to enable stronger binding by STAR proteins (Galarneau & Richard, 2009). However, the RNA oligo including this additional half-site did not increase

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

225 the affinity of HOW-STAR for the HOW binding site (Fig. 4D & Table 1). Therefore, we focused
226 simply on HOW(S) binding sites with a complete motif.

227 Multiple sites were found in Hipk-RA mRNA that match the 3'-UTR enriched motif, so
228 we characterised the binding of HOW(S) to RNA oligos containing the sequences of two of
229 these regions (Fig. 3F). The first site contains GCUAAC and exhibits high affinity binding (10.1
230 +/- 0.73 nM) with HOW-STAR (Fig. 4C). Our motif enrichment analysis indicated that the first
231 position of the top 3'-UTR motif, (A/G/C)CUAAC, appeared to be the most flexible (Fig 3A).
232 Therefore, the G from this GCUAAC site was mutated to A, C or U, and binding tested to
233 understand the effect on HOW-STAR binding affinity in the context of the same 14mer oligo
234 used in these FA experiments. This revealed that there is a preference for adenosine at this
235 site over guanosine, uridine and cytidine (Fig. 4E & Table 1). This is consistent with the motif
236 enrichment analysis, where adenosine is more commonly found compared with other
237 nucleotides (Fig. 3A). To show specificity of HOW binding to these NCUAAC motifs, the
238 adenosine in the fourth position was mutated to a guanosine, which resulted in a large
239 decrease (approximately 20 fold vs ACUAAC) in HOW-STAR binding affinity (Fig. 4E; Table 1).
240 This is also consistent with the mutational analysis performed on HOW's binding site in
241 *stripe*'s 3'-UTR, which can be bound by either HOW(S) or HOW(L) (Israeli et al., 2007).

242 The second Hipk-RA site contains two HOW(S) motifs in close proximity in the 3'-UTR.
243 The binding of HOW-STAR with this region, the Hipk 'double site', was of very similar affinity
244 to the site only containing one CUAAC motif, suggesting that only a monomer of HOW is
245 binding to this region, rather than dimer (Fig. 4C, Table 1). Other STAR proteins can bind as
246 dimers and this may be the case for HOW but at more distant sites than those within this
247 'double site' (Table 1), although avidity effects may also be in play here.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

248 Overall, we have identified 121 novel targets of cytoplasmic HOW(S) in testis germ
249 cells, several of which are involved in signal transduction. Within these HOW(S) bound RNAs
250 the (A/G/U)CUAAC motif is enriched in 3'-UTRs and GCG(A/U)G in 5'-UTRs. This 3'-UTR motif
251 is similar to potential HOW binding sites identified in *stripe* and *dpp* mRNAs, as well the
252 consensus sequences of other STAR proteins. The STAR domain of HOW has strong
253 (nanomolar K_D) affinity for RNA elements containing the CUAAC motif from *lola* and *hipk*
254 mRNAs but not with GCG(A/U)G 5'-UTR motif. Together these results provide new insight into
255 STAR protein-RNA interactions and potential importance to spermatogenesis. Given the
256 importance of HOW to germ stem cells, understanding these novel cytoplasmic protein-RNA
257 networks and their functional consequences will shed new light on how the balance of self-
258 renewal and differentiation is regulated in stem cells.
259
260 Methods
261 **Fly husbandry and stocks**
262 Flies were kept in a 25°C humidified room with a 12:12 hour light:dark cycle and raised on
263 10ml standard sugar-yeast-agar medium (Bass et al., 2007). Crosses were generated by
264 collecting unmated flies, then 5-10 flies of each sex were placed into a vial together with
265 grains of active baker's yeast. UAS-HOW-S-HA line generously gifted by Prof T. Volk, *nanos-*
266 *GAL4* line used is #64277 from BDSC.
267
268 **Immunofluorescence**
269 Testes from 0-3 day old unmated flies were fixed with 4% paraformaldehyde and stained
270 sequentially with anti-vasa then anti-HA. Slides were imaged using a Zeiss LSM880 Upright

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

271 Confocal Microscope with the 40X oil-immersion objective and Zen imaging software. See
272 the Supplemental Methods for details regarding tissue dissection, staining and antibodies.
273

274 **Ribonucleoprotein immunoprecipitation**

275 1000 pairs of testes from 0–3 day old flies (mated or unmated) were dissected per sample.
276 RIP pull-downs were adapted from (Keene et al., 2006), for full details see the Supplemental
277 Methods.
278

279 **Sequencing and computational analysis**

280 The RNA from the lysates and elutions of three HOW(S)-HA pull-downs and one *nanos*-GAL4
281 parental control pull-down was prepared using the Ribo-Zero rRNA Removal Kit (Illumina)
282 followed by the TruSeq Stranded Total RNA Library Prep (Illumina), sequencing was 75 bp
283 single-end. Differential transcript enrichment was carried out using edgeR, and motif
284 enrichment with DREME. See Supplemental Methods for details regarding sequencing,
285 filtering, mapping, and analyses.
286

287 **Data availability**

288 RIP-seq data have been deposited to the Gene Expression Omnibus with accession ID:
289 GSE201319.
290

291 **Protein expression and purification**

292 Residues 72–266 of *D. melanogaster* HOW (the STAR domain) was expressed in BL21(DE3)
293 cells with a His_6-GST tag and purified via affinity chromatography and size exclusion

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

294 chromatography. See Supplemental Methods for additional protein expression and
295 purification details.
296
297 **Fluorescence anisotropy**
298 RNA oligonucleotides were 3' labelled with 6-carboxyfluorescein and mixed with purified
299 STAR protein, the top protein concentration was 1 nM. See Supplemental Methods for assay
300 and calculation details.
301
302 Funding
303 MA was funded from BBSRC DTP, BB/M011151/1. JA was funded by the University of Leeds
304 (University Academic Fellow scheme).
305
306 Competing Interest Statement
307 Authors have no competing interests.
308
309 Acknowledgments
310 We thank the Next Generation Sequencing facility, at St James University Hospital, Leeds, UK
311 for performing Next Generation Sequencing. Parts of this work were undertaken on ARC3,
312 part of the High Performance Computing facilities at the University of Leeds, UK. We thank
313 the Protein Production Facility, especially Dr Brian Jackson. We also thank the Imaging Facility,
314 Faculty of Biological Sciences, University of Leeds, UK for their assistance in confocal
315 microscopy. We also thank the Biomolecular Mass Spectrometry Facility, University of Leeds,
316 UK for their assistance.
317

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

Author contributions

MA conceived, designed and performed experiments for the study. FH performed experiments for the study. TH performed experiments for the study. AB conceived the work, interpreted data, drafted and revised the manuscript. TE conceived the work, interpreted data, drafted and revised the manuscript. JLA conceived the work, interpreted data, drafted and revised the manuscript. All authors contributed to manuscript writing, revision and have approved the submitted version.

References

Baehrecke, E. H. (1997). who encodes a KH RNA binding protein that functions in muscle development. *Development*, *124*(7), 1323-1332.

Bass, T. M., Grandison, R. C., Wong, R., Martinez, P., Partridge, L., & Piper, M. D. W. (2007). Optimization of dietary restriction protocols in Drosophila. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *62*(10), 1071-1081. https://doi.org/10.1093/gerona/62.10.1071

Berber, S., Llamosas, E., Thaivalappil, P., Boag, P. R., Crossley, M., & Nicholas, H. R. (2013). Homeodomain interacting protein kinase (HPK-1) is required in the soma for robust germline proliferation in <i>C. elegans</i>. *Developmental Dynamics*, *242*(11), 1250-1261. https://doi.org/10.1002/dvdy.24023

Chen, J., & Verheyen, E. M. (2012). Homeodomain-interacting protein kinase regulates yorkie activity to promote tissue growth. *Current Biology*, *22*(17), 1582-1586. https://doi.org/10.1016/j.cub.2012.06.074

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

342 Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., . . . Goodall,
343 G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell*,
344 *160*(6), 1125-1134. https://doi.org/10.1016/j.cell.2015.02.014
345 Crapster, J. A., Rack, P. G., Hellmann, Z. J., Le, A. D., Adams, C. M., Leib, R. D., . . . Chen, J. K.
346 (2020). HIPK4 is essential for murine spermiogenesis. *eLife*, *9*, e50209-e50209.
347 https://doi.org/10.7554/eLife.50209
348 Duncan, K., Grskovic, M., Strein, C., Beckmann, K., Niggeweg, R., Abaza, I., . . . Hentze, M. W.
349 (2006). Sex-lethal imparts a sex-specific function to UNR by recruiting it to the msl-2
350 mRNA 3' UTR: translational repression for dosage compensation. *Genes Dev*, *20*(3),
351 368-379. https://doi.org/10.1101/gad.371406
352 Edenfeld, G., Volohonsky, G., Krukkert, K., Naffin, E., Lammel, U., Grimm, A., . . . Klämbt, C.
353 (2006). The splicing factor crooked neck associates with the RNA-binding protein
354 HOW to control glial cell maturation in Drosophila. *Neuron*, *52*(6), 969-980.
355 https://doi.org/10.1016/j.neuron.2006.10.029
356 Galarneau, A., & Richard, S. (2009). The STAR RNA binding proteins GLD-1, QKI, SAM68 and
357 SLM-2 bind bipartite RNA motifs. *BMC Molecular Biology*, *10*(47), doi:10.1186/1471-
358 2199-1110-1147. https://doi.org/10.1186/1471-2199-10-47
359 Giuliani, G., Giuliani, F., Volk, T., & Rabouille, C. (2014). The <i>Drosophila</i> RNA-binding
360 protein HOW controls the stability of <i>dgrasp</i> mRNA in the follicular
361 epithelium. *Nucleic Acids Research*, *42*(3), 1970-1986.
362 https://doi.org/10.1093/nar/gkt1118
363 Graindorge, A., Carré, C., & Gebauer, F. (2013). Sex-lethal promotes nuclear retention of
364 <i>msl2</i> mRNA via interactions with the STAR protein HOW. *Genes &*
365 *Development*, *27*(12), 1421-1433. https://doi.org/10.1101/gad.214999.113

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

366 Israeli, D., Nir, R., & Volk, T. (2007). Dissection of the target specificity of the RNA-binding
367 protein HOW reveals dpp mRNA as a novel HOW target. *Development*, *134*(11),
368 2107-2114. https://doi.org/10.1242/dev.001594
369 Jan, E., Motzny, C. K., Graves, L. E., & Goodwin, E. B. (1999). The STAR protein, GLD-1, is a
370 translational regulator of sexual identity in Caenorhabditis elegans. *The EMBO*
371 *Journal*, *18*(1), 258-269.
372 Keene, J. D., Komisarow, J. M., & Friedersdorf, M. B. (2006). RIP-Chip: The isolation and
373 identification of mRNAs, microRNAs and protein components of ribonucleoprotein
374 complexes from cell extracts. *Nature Protocols*, *1*(1), 302-307.
375 https://doi.org/10.1038/nprot.2006.47
376 Kondo, T., Furuta, T., Mitsunaga, K., Ebersole, T. A., Shichiri, M., Wu, J., . . . Abe, K. (1999).
377 Genomic organization and expression analysis of the mouse qkI locus. *Mammalian*
378 *Genome*, *10*(7), 662-669. https://doi.org/10.1007/s003359901068
379 Larkin, A., Marygold, S. J., Antonazzo, G., Attrill, H., Dos Santos, G., Garapati, P. V., . . .
380 Consortium, F. (2021). FlyBase: updates to the Drosophila melanogaster knowledge
381 base. *Nucleic Acids Res*, *49*(D1), D899-D907. https://doi.org/10.1093/nar/gkaa1026
382 Lee, W., Andrews, B. C., Faust, M., Walldorf, U., & Verheyen, E. M. (2009). Hipk is an
383 essential protein that promotes Notch signal transduction in the Drosophila eye by
384 inhibition of the global co-repressor Groucho. *Developmental Biology*, *325*(1), 263-
385 272. https://doi.org/10.1016/j.ydbio.2008.10.029
386 Legrand, J. M. D., & Hobbs, R. M. (2018). RNA processing in the male germline: Mechanisms
387 and implications for fertility. *Seminars in Cell and Developmental Biology*, *79*, 80-91.
388 https://doi.org/10.1016/j.semcdb.2017.10.006

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

389 Monk, A. C., Siddall, N. A., Fraser, B., McLaughlin, E. A., & Hime, G. R. (2011). Differential
390 roles of HOW in male and female Drosophila germline differentiation. *PLoS ONE*,
391 *6*(12), e28508-e28508. https://doi.org/10.1371/journal.pone.0028508
392 Monk, A. C., Siddall, N. A., Volk, T., Fraser, B., Quinn, L. M., McLaughlin, E. A., & Hime, G. R.
393 (2010). HOW Is Required for Stem Cell Maintenance in the Drosophila Testis and for
394 the Onset of Transit-Amplifying Divisions. *Cell Stem Cell*, *6*(4), 348-360.
395 https://doi.org/10.1016/j.stem.2010.02.016
396 Nabel-Rosen, H., Dorevitch, N., Reuveny, A., & Volk, T. (1999). The balance between two
397 isoforms of the Drosophila RNA-binding protein How controls tendon cell
398 differentiation. *Molecular Cell*, *4*(4), 573-584. https://doi.org/10.1016/S1097-
399 2765(00)80208-7
400 Nabel-Rosen, H., Volohonsky, G., Reuveny, A., Zaidel-Bar, R., & Volk, T. (2002). Two isoforms
401 of the Drosophila RNA binding protein, How, act in opposing directions to regulate
402 tendon cell differentiation. *Developmental Cell*, *2*(2), 183-193.
403 https://doi.org/10.1016/S1534-5807(01)00118-6
404 Nir, R., Grossman, R., Paroush, Z. e., & Volk, T. (2012). Phosphorylation of the <i>Drosophila
405 melanogaster</i> RNA-binding protein HOW by MAPK/ERK enhances its dimerization
406 and activity. *PLoS Genetics*, *8*(3), e1002632-e1002632.
407 https://doi.org/10.1371/journal.pgen.1002632
408 Poon, C. L. C., Zhang, X., Lin, J. I., Manning, S. A., & Harvey, K. F. (2012). Homeodomain-
409 interacting protein kinase regulates Hippo pathway-dependent tissue growth.
410 *Current Biology*, *22*(17), 1587-1594. https://doi.org/10.1016/j.cub.2012.06.075

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.19.492257; this version posted May 19, 2022. 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.

411 Ray, D., Kazan, H., Cook, K. B., Weirauch, M. T., Najafabadi, H. S., Li, X., . . . Hughes, T. R.
412 (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature*,
413 *499*(7457), 172-177. https://doi.org/10.1038/nature12311
414 Rodrigues, F., Thuma, L., & Klämbt, C. (2012). The regulation of glial-specific splicing of
415 Neurexin IV requires HOW and Cdk12 activity. *Development*, *139*(10), 1765-1776.
416 https://doi.org/8223268
417 Teplova, M., Hafner, M., Teplov, D., Essig, K., Tuschl, T., & Patel, D. J. (2013). Structure-
418 function studies of STAR family Quaking proteins bound to their in vivo RNA target
419 sites. *Genes & Development*, *27*(8), 928-940.
420 https://doi.org/10.1101/gad.216531.113
421 Tettweiler, G., Blaquiere, J. A., Wray, N. B., & Verheyen, E. M. (2019). Hipk is required for
422 JAK/STAT activity during development and tumorigenesis. *PLoS One*, *14*(12),
423 e0226856. https://doi.org/10.1371/journal.pone.0226856
424 Vernet, C., & Artzt, K. (1997). STAR, a gene family involved in signal transduction and
425 activation of RNA. *Trends in Genetics*, *13*(12), 479-484.
426 https://doi.org/10.1016/S0168-9525(97)01269-9
427
428
429
430
431
432
433