



# Analysis of insect nuclear small heat shock proteins and interacting proteins

Mohamed Taha Moutaoufik<sup>1,2</sup> · Robert M. Tanguay<sup>1</sup>

Received: 7 June 2020 / Revised: 13 August 2020 / Accepted: 19 August 2020 / Published online: 4 September 2020  
© Cell Stress Society International 2020

## Abstract

The small heat shock proteins (sHsps) are a ubiquitous family of ATP-independent stress proteins found in all domains of life. *Drosophila melanogaster* Hsp27 (DmHsp27) is the only known nuclear sHsp in insect. Here analyzing sequences from HMMER, we identified 56 additional insect sHsps with conserved arginine-rich nuclear localization signal (NLS) in the N-terminal region. At this time, the exact role of nuclear sHsps remains unknown. DmHsp27 protein-protein interaction analysis from iRefIndex database suggests that this protein, in addition to a putative role of molecular chaperone, is likely involved in other nuclear processes (i.e., chromatin remodeling and transcription). Identification of DmHsp27 interactors should provide key insights on the cellular and molecular functions of this nuclear chaperone.

**Keywords** Small heat shock protein (sHsp) · DmHsp27 · Chaperone · Alpha-crystallin domain (ACD) · *Drosophila melanogaster* · Insect

## Introduction

The nucleus is the cellular organelle that distinguishes eukaryotes from prokaryotes. A large number of biological activities occur in the nucleus, including DNA replication and damage repair, the biogenesis of ribosomal RNA precursors, transcription, and splicing of pre-mRNAs. Many proteins are involved in nuclear transport in processes dependent on Karyopherin- $\beta$  proteins (both importins and exportins).

The small heat shock proteins (sHsps) are a ubiquitous family of ATP-independent stress proteins found in all domains of life (Caspers et al. 1995; de Jong et al. 1998; Fu et al. 2006; Maaroufi and Tanguay 2013). These proteins are upregulated in response to a variety of stresses that negatively

impact protein homeostasis. sHsp sequence analysis indicates a tripartite architecture, with a conserved  $\alpha$ -crystallin domain (ACD) flanked by a variable non-conserved N- and C-terminal regions (NTR and CTR) (Kappé et al. 2010; Basha et al. 2012). sHsp number, level and stage of expression, tissue distribution, and intracellular localization can show differences between species, suggesting that sHsps have adapted to their environment and to specific cell needs (Michaud et al. 2002; Morrow and Tanguay 2015).

Relocation of cytosolic sHsps to the nucleus is observed under certain conditions of stress in virtually all mammals (Van De Klundert and De Jong 1999; Borrelli et al. 2002), but this is distinct from nuclear sHsps that are present all the time in the nucleus in *Drosophila melanogaster* and plant, due to the presence of a nuclear localization signal (NLS) (Beaulieu et al. 1989; Marin and Tanguay 1996; Scharf et al. 2001; Siddique et al. 2003, 2008; Michaud et al. 2008).

Determining the cellular localization of proteins is important for understanding protein functions. Here we focused on the nuclear-localized small Hsp27 of *Drosophila melanogaster* (DmHsp27). This sHsp was the first reported to be localized in the nucleus (Beaulieu et al. 1989), and since this discovery, other sHsps, especially from plants, have been shown to have a nuclear localization such as *Solanum peruvianum* LpHsp16.1, *Arabidopsis thaliana* AtHsp17.4, *Gossypium arboreum* GaHsp17.3, *Medicago truncatula*

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12192-020-01156-3>) contains supplementary material, which is available to authorized users.

✉ Robert M. Tanguay  
robert.tanguay@fmed.ulaval.ca

<sup>1</sup> Lab of Cell & Developmental Genetics, Department of Cellular and Molecular Biology, Medical Biochemistry & Pathology, Medical School, Université Laval, Quebec G1K 7P4, Canada

<sup>2</sup> Present address: Department of Biochemistry, University of Regina, Regina, SK S4S 0A2, Canada

MtHsp17.1, *Lactuca sativa* LsHsp17.8, *Triticum aestivum* TaHsp19.0, *Hordeum vulgare* HvHsp19.1, and *Oryza sativa* OsHsp18.6 (Scharf et al. 2001; Siddique et al. 2003, 2008). The import of proteins into the nucleus is dependent on a nuclear localization signals (NLS), and members of the nuclear transport receptor (importin  $\beta$ -like) superfamily. In plants, the nuclear localization depends on a short motif (N(G)KRKR) located between  $\beta$ -strands 5 and 6 in the conserved ACD (Siddique et al. 2003). This mode requires importin  $\alpha$  to serve as a bridge between the NLS and the import receptor importin  $\beta$  (Marfori et al. 2012). In *Drosophila melanogaster* DmHsp27, the NLS motif is composed of three arginine (Arg54-Arg55-Arg56) in the NTR (Michaud et al. 2008). Authors suggested that arginine 54 and 55 form a single functional group that acts in concert with arginine 56 to dictate nuclear localization of DmHsp27 (Michaud et al. 2008). Arginine-rich NLS uses importin  $\beta$  for import by a mechanism independent of importin  $\alpha$  (Palmeri and Malim 1999). The difference between the nuclear localization of sHsps in mammals, plants, and *D. melanogaster* suggests that they adapt to their environments and their specific cellular needs. While we know the existence of several sHsps with an NLS in plants, DmHsp27 is the only member reported in insect. NTR in sHsp was reported as a disordered non-conserved region due to the lack of sequence conservation and structure among this family (Kim et al. 1998; van Montfort et al. 2001). Therefore, localization of a nuclear signal in this disordered and non-conserved region is surprising. Here we characterize nuclear insect sHsps containing an arginine-rich NLS and suggest possible functions of this group of sHsps based on DmHsp27 interacting partners.

## Materials and methods

### Sequence collection of the insect sHsps

All insect protein sequences containing one or more instances of ACD (accession number PF00011) were extracted from HMMER web server (Finn et al. 2011). Proteins containing the term “fragment” in their description were removed. From this data, we extracted only sequences containing arginine-rich NLS (residues XRR and RXR) in NTR as defined in *Drosophila melanogaster* by Michaud et al. (2008). After alignment using ClustalW (Larkin et al. 2007), only proteins showing a conservation of NLS with DmHsp27 have been kept.

### Lengths, amino acid composition, alignment, and logos of insect sHsp

The beginning and end locations ACD of all sequences were extracted from PFAM database (Finn et al. 2014). The

sequences were split into the 3 building blocks (ACD, NTR, and CTR). The length frequencies were analyzed using R ggplot2 package (Wickham 2009), and amino acid distribution average was evaluated using R stringi package (<https://cran.r-project.org/web/packages/stringi/>).

The sequences were aligned with ClustalW default parameters. Secondary structure of the alignment generated by ESPript (Robert and Gouet 2014) according to DmHsp27 ACD predicted structure by Moutaoufik et al. (2017b). UniProt accession numbers of sequences used in the alignments are indicated. To identify conserved regions, we used WebLogo 3 program (Crooks et al. 2004) to create blocks (LOGOS) of conserved amino acid residues from the multiple protein sequences.

### Phylogenetic analysis of the insect sHsps

Multiple alignments of insect sHsp sequences were calculated in MEGAX ((Kumar et al. 2018)) using ClustalW default parameters. Maximum likelihood phylogenetic trees were constructed using MEGAX default parameters and 50 bootstrap replicates. Phylogenetic trees were visualized using FigTree v1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

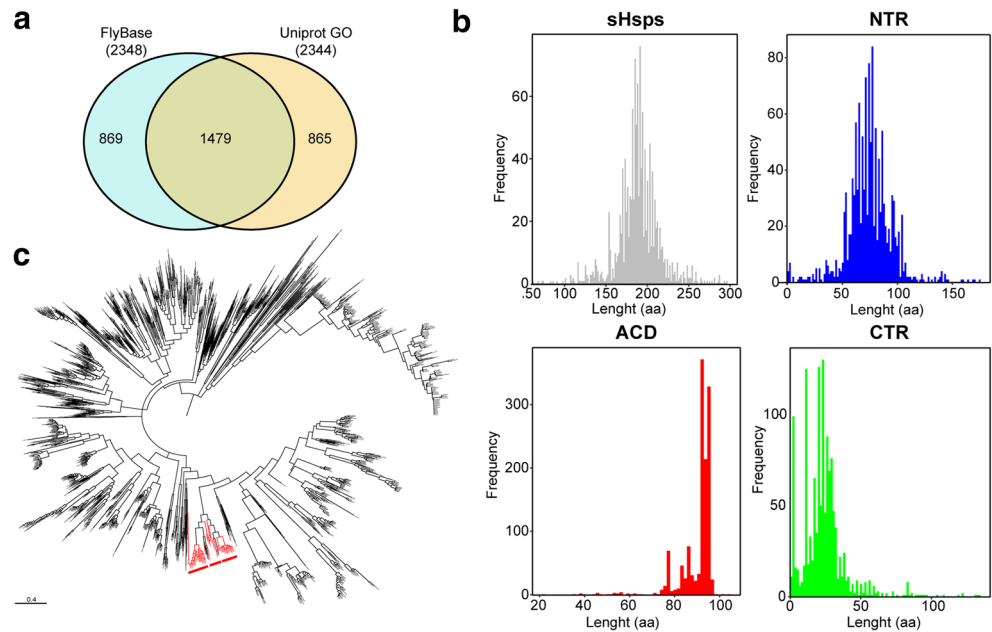
### Protein-protein interaction data

In order to understand nuclear functions of insect sHsps, we searched protein-protein interaction (PPI) data of insect nuclear sHsps using iRefIndex database (Razick et al. 2008). From all other insect species identified in this study, only *Drosophila melanogaster* PPI data was available in iRefIndex. DmHsp27 experimentally verified PPI data were downloaded, filtered by nuclear localization, and imported into Cytoscape (Shannon et al. 2003) to construct PPI network.

## Results and discussion

Nucleus plays a central role in eukaryotes. However, its protein content is far from complete. A mouse nuclear proteomic prediction reveals that mouse nucleus is the house of between 4084 and 9122 proteins (Fink et al. 2008). Using FlyBase (Thurmond et al. 2019) and UniProt Gene Ontology (GO) (2019), we identified reported nuclear proteins in *Drosophila melanogaster*. The total number of nuclear proteins reported in FlyBase and UniProt Go is constant with 2348 and 2344 entries, respectively (Fig. 1a, Supplementary Table 1), although more than half proteins (1479) are shared between two databases.

**Fig. 1** *Drosophila melanogaster* proteome and insect sHsps. **A** Venn diagram showing common and specific *Drosophila melanogaster* nuclear proteins between public FlyBase and UniProt GO databases. **B** Frequency of nuclear sHsp length. Full length sequence presented in (gray), NTR (blue), ACD (red), and CTR (green). **C** Nuclear sHsp evolutionary analysis by Maximum Likelihood method. Nuclear sHsps highlighted in red



## Collection of nuclear insect sHsp sequences

It is accepted that the presence of the conserved ACD is a sufficient criterion for assigning a sequence to the sHsp family (Caspers et al. 1995; Kappé et al. 2010). To collect the sequences of sHsps, we searched the presence of ACD by using the domain PF00011 in HMMER web server (Finn et al. 2011). We obtained 1321 insect sequences containing ACD (Supplementary Table 2), with an average length of 188.82 ( $\pm 26.89$ ) (Fig. 1b). NTR have a length average of 74.50 ( $\pm 20.67$ ) and CTR 23.98 ( $\pm 16.26$ ); both show a variation in the length and frequency (Fig. 1b). ACD less variable region with a length average of ACD 90.33 ( $\pm 7.95$ ).

Analysis of NTR shows that 57 sequences contain an arginine-rich NLS in NTR (Table 1). The potential nuclear localization was confirmed by Euk-mPLoc (Chou and Shen 2007), a powerful tool, developed by hybridizing the functional domain information and sequential evolutionary information through three different modes of pseudo amino acid composition (Eq.4, Eq.6, and Eq.12).

Next, we studied the evolutionary history of insect sHsp sequences (Fig. 1c). It is interesting that insect nuclear sHsps were clustered in one phylogenetic group, suggesting that all nuclear insect sHsps evolved from the same ancestral sequence.

## Analysis of insect nuclear sHsps

Insect sHsp sequences that presented arginine-rich NLS include *Drosophila* genus: *Drosophila simulans* (2 sequences) and 1 sequence for *Drosophila albomicans*, *Drosophila ananassae*, *Drosophila busckii*, *Drosophila buzzatii*,

*Drosophila erecta*, *Drosophila ficusphila*, *Drosophila grimshawi*, *Drosophila guanche*, *Drosophila melanogaster*, *Drosophila mojavensis*, *Drosophila navajoa*, *Drosophila persimilis*, *Drosophila pseudoobscura pseudoobscura*, *Drosophila repletoides*, *Drosophila sechellia*, *Drosophila sulfurigaster albostrigata*, *Drosophila virilis*, *Drosophila willistoni*, and *Drosophila yakuba*. In addition to other genera: *Haematobia irritans* (20 sequences), *Musca domestica* (6 sequences), *Culicoides sonorensis* (2 sequences) and 1 sequence for *Bactrocera cucurbitae*, *Bactrocera dorsalis*, *Bactrocera latifrons*, *Ceratitis capitata*, *Glossina austeni*, *Glossina morsitans morsitans*, *Lucilia cuprina*, and *Sarcophaga crassipalpis* (Table 1) suggesting that arginine-rich NLS is not specific to drosophila.

The average mass in Dalton is 23,489.00 ( $\pm 1086.55$ ). The average amino acid length is 207.77 ( $\pm 11.45$ ). The longest sequence is B4MN48 from *Drosophila willistoni* with 227 amino acids (aa) and the shortest is A0A336MGY1 from *Culicoides sonorensis* with sequence 167 aa (Table 1).

The average length of amino acids in NTR is 84.40 ( $\pm 2.61$ ) (Fig. 2a). The analysis of NTR revealed underrepresentation of cysteine and an overrepresentation of leucine, proline, aspartic acid, and arginine (Fig. 2a and b). ACD has an average length of 88.56 ( $\pm 2.74$ ) (Fig. 2a). In contrast, within the ACD, a significant enrichment of charged positively (lysine) and negatively (aspartic acid and glutamic acid) residues is observed, whereas tryptophan is absent (Fig. 2a and c). Most sequences present one cysteine in ACD, suggesting that under oxidative conditions, a disulfide crosslink may be formed between ACDs to form a dimer as shown for DmHsp27 (Moutaoufik et al. 2017b). In CTR, the shortest region has an average amino acid number of 34.80 ( $\pm 1.50$ )

**Table 1** Insect nuclear sHsps

Protein name	UniProt ID	Species	Length	Mass (Dalton)	Arginine-rich NLS localization
Hsp 27	P02518	<i>Drosophila melanogaster</i>	213	23,617	Arg 54,55,56
GM25143	B4HKP4	<i>Drosophila sechellia</i>	213	23,625	Arg 54,55,56
GD14177	A0A0J9UHX1	<i>Drosophila simulans</i>	213	23,625	Arg 54,55,56
GD14177	B4QN53	<i>Drosophila simulans</i>	213	23,566	Arg 54,55,56
GG15371	B3NCF2	<i>Drosophila erecta</i>	212	23,650	Arg 54,55,56
GE20833	B4PEX9	<i>Drosophila yakuba</i>	212	23,594	Arg 54,55,56
ND	A0A1I8M7V4	<i>Musca domestica</i>	216	24,395	Arg 52–53
Hsp20	T1PAE8	<i>Musca domestica</i>	216	24,364	Arg 52–53
Hsp27	A0A1L8EIU6	<i>Haematobia irritans</i>	203	23,200	Arg 48–50
Hsp27	A0A1L8EY3	<i>Haematobia irritans</i>	203	23,250	Arg 48–50
Hsp27	A0A1L8EJ25	<i>Haematobia irritans</i>	203	23,327	Arg 48–50
Hsp27	A0A1L8EJ27	<i>Haematobia irritans</i>	203	23,223	Arg 48–50
Hsp27	A0A1L8EJ69	<i>Haematobia irritans</i>	203	23,350	Arg 48–50
Hsp27	A0A1L8EJ80	<i>Haematobia irritans</i>	203	23,212	Arg 48–50
Hsp27	A0A1X9JRV5	<i>Musca domestica</i>	203	23,031	Arg 47–49
ND	A0A1I8M631	<i>Musca domestica</i>	203	23,040	Arg 47–49
Hsp27	A0A1L8EJ49	<i>Haematobia irritans</i>	203	23,261	Arg 48–50
Hsp27	A0A1L8EIP8	<i>Haematobia irritans</i>	202	23,112	Arg 49–50–51
Hsp27	A0A1L8EIV9	<i>Haematobia irritans</i>	202	23,126	Arg 49–50–51
Hsp27	A0A1L8EJ16	<i>Haematobia irritans</i>	202	23,098	Arg 49–50–51
Hsp27	A0A1L8EIS9	<i>Haematobia irritans</i>	203	22,956	Arg 48–50
Hsp27	A0A1L8EIV6	<i>Haematobia irritans</i>	203	23,117	Arg 48–50
Hsp27	A0A1L8EIW9	<i>Haematobia irritans</i>	203	23,078	Arg 48–50
Hsp27	A0A1L8EIX8	<i>Haematobia irritans</i>	203	23,097	Arg 48–50
Hsp27	A0A1L8EIY9	<i>Haematobia irritans</i>	203	23,022	Arg 48–50
Hsp27	A0A1L8EJ10	<i>Haematobia irritans</i>	203	23,068	Arg 48–50
Hsp27	A0A1L8EJ18	<i>Haematobia irritans</i>	203	23,038	Arg 48–50
Hsp27	A0A1L8EJ35	<i>Haematobia irritans</i>	203	23,123	Arg 48–50
Hsp27	A0A1L8EJ40	<i>Haematobia irritans</i>	203	23,105	Arg 48–50
Hsp27	A0A1L8EJ47	<i>Haematobia irritans</i>	203	23,127	Arg 48–50
ND	T1PK02	<i>Musca domestica</i>	198	22,646	Arg 47–49
Hsp27	A0A1X9JUH8	<i>Musca domestica</i>	198	22,628	Arg 47–49
Hsp27	A0A0L0CSL4	<i>Lucilia cuprina</i>	203	23,331	Arg 47–49–50
GK17637	B4MN48	<i>Drosophila willistoni</i>	227	25,285	Arg 58,59,60
Hsp27	A9UEZ2	<i>Drosophila buzzatii</i>	211	23,695	Arg 51,52,53
ND	A0A1B0G3A3	<i>Glossina morsitans morsitans</i>	212	24,164	Arg 55–57
ND	A0A1A9UGE3	<i>Glossina austeni</i>	184	21,061	Arg 55–57
Hsp27	I1T1H1	<i>Bactrocera dorsalis</i>	210	23,831	Arg 51–52–53
Hsp27	A0A0K8UP97	<i>Bactrocera latifrons</i>	212	24,037	Arg 53–54–55
Hsp27	A0A0A1XCZ4	<i>Bactrocera cucurbitae</i>	212	23,920	Arg 53–54–55
Hsp27	B3GK93	<i>Ceratitidis capitata</i>	214	23,944	Arg 54–55–56
Hsp27	A0A0M4EGE5	<i>Drosophila busckii</i>	217	24,445	Arg 51–52–53
Hsp27	A0A3B0JUV4	<i>Drosophila guanche</i>	225	25,016	Arg 58–59–60
GL22445	B4H1L8	<i>Drosophila persimilis</i>	225	25,041	Arg 58,59,60
GA18205	Q29F98	<i>Drosophila pseudoobscura pseudoobscura</i>	225	25,041	Arg 58,59,60
GF10692	B3M6F4	<i>Drosophila ananassae</i>	219	24,490	Arg 57,58,59
Hsp27	F5B960	<i>Drosophila albomicans</i>	222	24,770	Arg 52,53,54
Hsp27	F5B961	<i>Drosophila sulfurigaster albostrigata</i>	221	24,675	Arg 52,53,54
GJ13835	B4LGS9	<i>Drosophila virilis</i>	211	23,792	Arg 51,52,53

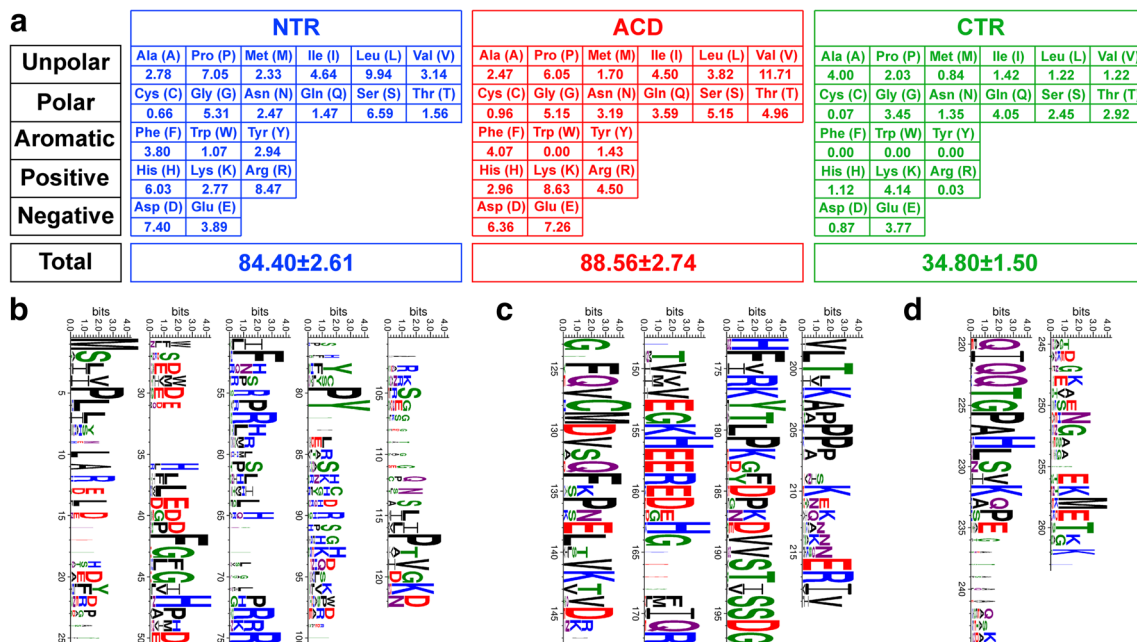
**Table 1** (continued)

Protein name	UniProt ID	Species	Length	Mass (Dalton)	Arginine-rich NLS localization
GH17009	B4IYY1	<i>Drosophila grimshawi</i>	218	24,442	Arg 54,55,56
Hsp25	A1E385	<i>Sarcophaga crassipalpis</i>	221	24,982	Arg 60–61
ND	A0A484BBT9	<i>Drosophila navojoa</i>	211	23,628	Arg 51–52–53
GI13087	B4L112	<i>Drosophila mojavensis</i>	209	23,459	Arg 51,52,53
Hsp27	F5B962	<i>Drosophila repletoidea</i>	220	24,504	Arg 48,49,50
CSON012237	A0A336MGY1	<i>Culicoides sonorensis</i>	167	19,445	Arg 40–41
CSON11881	A0A336MMP2	<i>Culicoides sonorensis</i>	167	19,401	Arg 40–41
Hsp27	A0A1W4VN65	<i>Drosophila ficusphila</i>	213	23,784	Arg 53–54–55

with absence of tryptophan, phenylalanine, and tyrosine. Additionally, there is underrepresentation of cysteines and arginine (Fig. 2a and d). This is consistent with a previous study from 1935 metazoan proteins which contain ACD by Kriehuber et al. (2010).

Blocks of the most conserved residues were made as logos to visualize the conserved regions in three regions—NTR, ACD, CTR—of nuclear insect sHsps (Fig. 2c, d, and e) in addition to a complete sequence alignment which (Supplementary Fig. 1) shows that motif characteristics of sHsps were conserved. This is the case of the arginine in the Beta 6 + 7 strand: DmHsp27-R122 equivalent of HspB1-R127, DmHsp27-R131 equivalent of HspB1-R136, and DmHsp27-R135 equivalent of HspB1-R140, HspB5-R120, and HspB4-R116 associated with human pathologies (Vicart

et al. 1998; Bera and Abraham 2002; Evgrafov et al. 2004; Inagaki et al. 2006; Gentil and Cooper 2012; Ylikallio et al. 2015). In addition to the conserved I/V/L-X-I/V/L motif in ACD, implicated in sHsps higher assembly by interaction with  $\beta 4/\beta 8$  pocket in ACD of adjacent dimer (de Jong et al. 1998; van Montfort et al. 2001; Stamler et al. 2005; Poulain et al. 2010; Clark et al. 2018) and the FGFG motif in NTR, important for the oligomeric structure and chaperone-like activity (Moutaoufik et al. 2017a). In addition, B4PEX9 (*Drosophila yakuba*), B4HKP4 (*Drosophila sechellia*), B3NCF2 (*Drosophila erecta*), A1E385 (*Sarcophaga crassipalpis*), and A0A0J9UHX1 and B4QN53 (*Drosophila simulans*) show a conservation in both phosphorylated serines in DmHsp27 (S58 and S75) reported by Moutaoufik et al. (2017a). T1PAE8 (*Musca domestica*), B4L112 (*Drosophila*



**Fig. 2** Insect nuclear sHsps. **A** Distribution of residues within the nuclear sHsps in insect. Average number of each amino acid within the NTR (blue), CTR (green), and the ACD (red) is presented. **B**, **C**, and **D** Logo presentation using WebLogo 3 program (Crooks et al. 2004). The height of each letter is made proportional to its frequency. Amino acids are

colored according to their chemical properties: acidic (D,E) in red, basic (K,R,H) in blue, polar (G,S,T,Y,C) in green and (N,Q) in purple, and hydrophobic (A,V,L,I,P,W,F,M) in black. Logo representation for Logos presentation of NTR (**B**), ACD (**C**), and CTR (**D**) of insect nuclear sHsp sequences

**Table 2** Proteins associated with DmHsp27

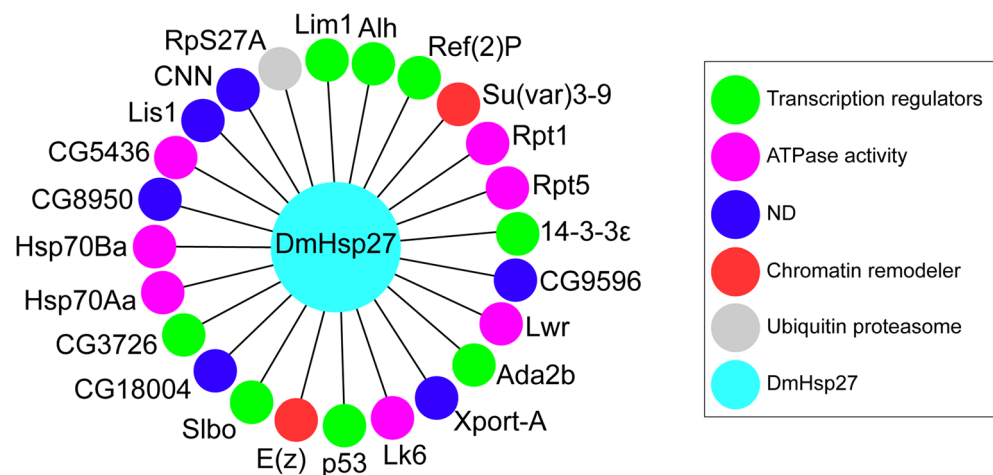
Protein 1	UniprotID1	Protein 2	UniprotID2	Method	Reference
14-3-3ε	P92177	DmHsp27	P02518	Colocalization	Müller et al. (2010)
Ada2b	Q8I8V0	Dmhsp27	P02518	Two-hybrid	Giot et al. (2003)
Alh	Q9VI61	DmHsp27	P02518	Two-hybrid	Giot et al. (2003)
CG18004	A1Z8D4	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
CG3726	X2JDV1	DmHsp27	P02518	AP-MS	Rhee et al. (2014)
CG8950	Q7JVV5	DmHsp27	P02518	AP-MS	Rhee et al. (2014)
CG9596	Q8MS69	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
CG5436	O97125	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
CNN	P54623	DmHsp27	P02518	MS	Habermann et al. (2012)
E(z)	P42124	DmHsp27	P02518	AP-MS	Kang et al. (2015)
DmHsp27	P02518	DmHsp27	P02518	BN gel	Moutaoufik et al. (2017a)
Hsp70 Aa	P82910	Dmhsp27	P02518	AP-MS	Guruharsha et al. (2011)
Hsp70Ba	Q8INI8	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
Lim1	M9PJE4	DmHsp27	P02518	Two-hybrid	Giot et al. (2003)
Lis1	Q7KNS3	DmHsp27	P02518	Two-hybrid	Giot et al. (2003)
Lk6	Q9VGI4	DmHsp27	P02518	Two-hybrid	Giot et al. (2003)
Lwr	Q7KNM2	DmHsp27	P02518	AP-WB/Two-hybrid	Joannis et al. (1998)
p53	A0A0B4K7P1	DmHsp27	P02518	AP-WB	Lei et al. (2017)
Rpt1	Q7KMQ0	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
Ref(2)P	P14199	DmHsp27	P02518	AP-MS	Guruharsha et al. (2011)
RpS27A	P15357	DmHsp27	P02518	AP-WB	Lee et al. (2014)
Rpt5	Q9V3V6	DmHsp27	P02518	AP-WB	Cho-Park and Steller (2013)
Slbo	Q02637	DmHsp27	P02518	Genetic Interactions	Rørth et al. (1998)
Su(var)3-9	I0DHL3	DmHsp27	P02518	Two-hybrid	Giot et al. (2003)
Xport-A	Q9VDS3	DmHsp27	P02518	AP-WB	Rosenbaum et al. (2011)

*mojavensis*), A9UEZ2 (*Drosophila buzzatii*), A0A1I8M7V4 (*Musca domestica*), and A0A484BBT9 (*Drosophila navojoa*) show only a conservation in phosphorylated serine (S58). However, A0A1L8EIU6, A0A1L8EY3, A0A1L8EJ25, A0A1L8EJ27, A0A1L8EJ69, A0A1L8EJ80, A0A1L8EJ49,

A0A1L8EIP8, A0A1L8EIV9, A0A1L8EJ16, A0A1L8EIS9, A0A1L8EIV6, A0A1L8EIV9, A0A1L8EIX8, A0A1L8EY9, A0A1L8EJ10, A0A1L8EJ18, A0A1L8EJ35, A0A1L8EJ40 and A0A1L8EJ47 (*Haematobia irritans*), A0A1X9JRV5, A0A1I8M631, T1PK02 and A0A1X9JUH8

**Fig. 3** DmHsp27 and its nuclear-associated proteins.

Representation of DmHsp27 and associated protein assemblies. Green color indicates transcription regulators, purple color designate proteins with ATPase activity, blue color indicate proteins with non-defined nuclear functions, red color indicate chromatin remodeler, and gray indicates protein involved in ubiquitin proteasome system



(*Musca domestica*), A0A3B0JUV4 (*Drosophila guanche*), B3M6F4 (*Drosophila ananassae*), A0A1W4VN65 (*Drosophila ficusphila*), A0A0L0CSL4 (*Lucilia cuprina*), A0A1B0G3A3 (*Glossina morsitans morsitans*), A0A1A9UGE3 (*Glossina austeni*), Q29F98 (*Drosophila pseudoobscura pseudoobscura*), B4MN48 (*Drosophila willistoni*), and B4H1L8 (*Drosophila persimilis*) show only a conservation in phosphorylated serine (S75).

### DmHsp27 interactome and nuclear functions

The nuclear functions of insect sHsps remain unknown. A study had reported an association of DmHsp22, DmHsp23, DmHsp26, and DmHsp27 to heterogeneous nuclear ribonucleoprotein (hnRNP) (Kloetzel and Bautz 1983). In addition, under heat shock conditions, DmHsp27 and cytoplasmic DmHsps were co-purified with the proteasome (Arrigo et al. 1985).

We searched the protein partners of insect sHsps identified in this study, using iRefIndex (Razick et al. 2008). Only *Drosophila melanogaster* protein-protein interactions are presented in the used database. DmHsp27 shows 46 associations with both nuclear and cytoplasmic proteins. After filtering by subcellular localization, we identified 24 nuclear proteins that tend to interact with DmHsp27 in addition to self-interacting to form oligomeric structures (Fig. 3, Table 2). Third (8, 33%) of DmHsp27 nuclear partners were involved in regulation of transcription, more than one-quarter (7, 29%) in ATPase activity, (2, 0.08%) were implicated in chromatin remodeling, (1, 0.04%) in ubiquitin proteasome complex and quarter (6, 25%) with no known nuclear function (Fig. 3b). These include proteins involved in many biological processes such as fly development (14-3-3 $\epsilon$ , Lk6, enhancer of zeste E(Z), centrosomin (cnn), lissencephaly-1 (Lis-1), lesswright (lwr), p53, slow border cells (slbo), and suppressor of variegation 3-9 (Su(var)3-9)) (Ashton-Beaucage et al. 2014; Arquier et al. 2005; Jones et al. 1998; Eisman et al. 2009; Liu et al. 1999; Miles et al. 2008; Bauer et al. 2007; Borghese et al. 2006; Kunert et al. 2003). In addition, LIM homeobox 1 (Lim1) implicated in eye development (Roignant et al. 2010). However, 14-3-3 $\epsilon$ , Alhambra (Alh), and transcriptional adaptor 2 Ada2b were associated with locomotor rhythm and muscle and wing development (Bahri et al. 2001; Gause et al. 2006; Bejarano et al. 2008). Finally, 14-3-3 $\epsilon$ , p53, and CG5436 (Hsp68) were related to fly viability (Bauer et al. 2005; Nielsen et al. 2008; Biteau et al. 2010), lwr to immune system response (Bhaskar et al. 2002), and CG5436 (Hsp68), Hsp70Aa, and Hsp70Ba to chaperone activity (Gong and Golic 2006; Gaudet et al. 2011). This is consistent with the related expression of DmHsp27 during oogenesis and spermatogenesis (Marin and Tanguay 1996; Michaud et al. 1997), eye development (Chen et al. 2012), lifespan (Hao et al. 2007;

Liao et al. 2008), and defense against pathogens (Chen et al. 2010).

### Conclusion

Here we have shown that the NLS signal rich in arginine is not a peculiarity of *Drosophila melanogaster*. The progress of genome sequencing projects should confirm the presence of other insect sHsp sequences with NLS signal, of the same type or not. The exact role of nuclear sHsps remains unknown. The interaction network of DmHsp27 suggests that this protein does not only play the role of molecular chaperone, but it is likely involved in different nuclear processes. Such studies open the perspectives to establish the functional activities of DmHsp27 and associated proteins in different biological processes.

### References

- Arquier N, Bourouis M, Colombani J, Léopold P (2005) *Drosophila* Lk6 kinase controls phosphorylation of eukaryotic translation initiation factor 4E and promotes normal growth and development. *Curr Biol* 15:19–23. <https://doi.org/10.1016/j.cub.2004.12.037>
- Arrigo AP, Darlix JL, Khandjian EW, Simon M, Spahr PF (1985) Characterization of the prosome from *Drosophila* and its similarity to the cytoplasmic structures formed by the low molecular weight heat-shock proteins. *EMBO J* 4:399–406
- Ashton-Beaucage D, Udell CM, Gendron P, Sahmi M, Lefrançois M, Baril C, Guenier AS, Duchaine J, Lamarre D, Lemieux S, Therrien M (2014) A functional screen reveals an extensive layer of transcriptional and splicing control underlying RAS/MAPK signaling in *Drosophila*. *PLoS Biol* 12:e1001809. <https://doi.org/10.1371/journal.pbio.1001809>
- Bahri SM, Chia W, Yang X (2001) The *Drosophila* homolog of human AF10/AF17 leukemia fusion genes (Dalf) encodes a zinc finger/leucine zipper nuclear protein required in the nervous system for maintaining EVE expression and normal growth. *Mech Dev* 100:291–301. [https://doi.org/10.1016/S0925-4773\(00\)00539-6](https://doi.org/10.1016/S0925-4773(00)00539-6)
- Basha E, O'Neill H, Vierling E (2012) Small heat shock proteins and  $\alpha$ -crystallins: dynamic proteins with flexible functions. *Trends Biochem Sci* 37:106–117. <https://doi.org/10.1016/j.tibs.2011.11.005>
- Bauer JH, Poon PC, Glatt-Deeley H, Abrams JM, Helfand SL (2005) Neuronal expression of p53 dominant-negative proteins in adult *Drosophila melanogaster* extends life span. *Curr Biol* 15:2063–2068. <https://doi.org/10.1016/j.cub.2005.10.051>
- Bauer JH, Chang C, Morris SNS, Hozier S, Andersen S, Waitzman JS, Helfand SL (2007) Expression of dominant-negative Dmp53 in the adult fly brain inhibits insulin signaling. *Proc Natl Acad Sci U S A* 104:13355–13360. <https://doi.org/10.1073/pnas.0706121104>
- Beaulieu JF, Arrigo AP, Tanguay RM (1989) Interaction of *Drosophila* 27,000 Mr heat-shock protein with the nucleus of heat-shocked and ecdysone-stimulated culture cells. *J Cell Sci* 92:29–36
- Bejarano F, Luque CM, Herranz H, Sorrosal G, Rafel N, Pham TT, Milán M (2008) A gain-of-function suppressor screen for genes involved in dorsal-ventral boundary formation in the *Drosophila* wing. *Genetics* 178:307–323. <https://doi.org/10.1534/genetics.107.081869>

- Bera S, Abraham EC (2002) The alphaA-crystallin R116C mutant has a higher affinity for forming heteroaggregates with alphaB-crystallin. *Biochemistry* 41:297–305. <https://doi.org/10.1021/bi011010v>
- Bhaskar V, Smith M, Courey AJ (2002) Conjugation of Smt3 to dorsal may potentiate the *Drosophila* immune response. *Mol Cell Biol* 22:492–504
- Biteau B, Karpac J, Supoyo S, DeGennaro M, Lehmann R, Jasper H (2010) Lifespan extension by preserving proliferative homeostasis in *Drosophila*. *PLoS Genet* 6:e1001159. <https://doi.org/10.1371/journal.pgen.1001159>
- Borghese L, Fletcher G, Mathieu J, Atzberger A, Eades WC, Cagan RL, Rørth P (2006) Systematic analysis of the transcriptional switch inducing migration of border cells. *Dev Cell* 10:497–508. <https://doi.org/10.1016/j.devcel.2006.02.004>
- Borrelli MJ, Bemock LJ, Landry J, Spitz DR, Weber LA, Hickey E, Freeman ML, Corry PM (2002) Stress protection by a fluorescent Hsp27 chimera that is independent of nuclear translocation or multimeric dissociation. *Cell Stress Chaperones* 7:281–296
- Caspers GJ, Leunissen JA, de Jong WW (1995) The expanding small heat-shock protein family, and structure predictions of the conserved “alpha-crystallin domain”. *J Mol Evol* 40:238–248
- Chen J, Xie C, Tian L, Hong L, Wu X, Han J (2010) Participation of the p38 pathway in *Drosophila* host defense against pathogenic bacteria and fungi. *Proc Natl Acad Sci U S A* 107:20774–20779. <https://doi.org/10.1073/pnas.1009223107>
- Chen S-F, Kang M-L, Chen Y-C, Tang HW, Huang CW, Li WH, Lin CP, Wang CY, Wang PY, Chen GC, Wang HD (2012) Autophagy-related gene 7 is downstream of heat shock protein 27 in the regulation of eye morphology, polyglutamine toxicity, and lifespan in *Drosophila*. *J Biomed Sci* 19:52. <https://doi.org/10.1186/1423-0127-19-52>
- Cho-Park PF, Steller H (2013) Proteasome regulation by ADP-ribosylation. *Cell* 153:614–627. <https://doi.org/10.1016/j.cell.2013.03.040>
- Chou K-C, Shen H-B (2007) Euk-mPLOC: a fusion classifier for large-scale eukaryotic protein subcellular location prediction by incorporating multiple sites. *J Proteome Res* 6:1728–1734. <https://doi.org/10.1021/pr060635i>
- Clark AR, Vree Egberts W, Kondrat FDL, Hilton GR, Ray NJ, Cole AR, Carver JA, Benesch JLP, Keep NH, Boelens WC, Slingsby C (2018) Terminal regions confer plasticity to the tetrameric assembly of human HspB2 and HspB3. *J Mol Biol* 430:3297–3310. <https://doi.org/10.1016/j.jmb.2018.06.047>
- Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190. <https://doi.org/10.1101/gr.849004>
- de Jong WW, Caspers G-J, Leunissen JAM (1998) Genealogy of the  $\alpha$ -crystallin—small heat-shock protein superfamily. *Int J Biol Macromol* 22:151–162. [https://doi.org/10.1016/S0141-8130\(98\)00013-0](https://doi.org/10.1016/S0141-8130(98)00013-0)
- Eisman RC, Phelps MAS, Kaufman TC (2009) Centrosomin: a complex mix of long and short isoforms is required for centrosome function during early development in *Drosophila melanogaster*. *Genetics* 182:979–997. <https://doi.org/10.1534/genetics.109.103887>
- Evgrafov OV, Mersyanova I, Irobi J, van den Bosch L, Dierick I, Leung CL, Schagina O, Verpoorten N, van Impe K, Fedotov V, Dadali E, Auer-Grumbach M, Windpassinger C, Wagner K, Mitrovic Z, Hilton-Jones D, Talbot K, Martin JJ, Vasserman N, Tverskaya S, Polyakov A, Liem RKH, Gettemans J, Robberecht W, de Jonghe P, Timmerman V (2004) Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy. *Nat Genet* 36:602–606. <https://doi.org/10.1038/ng1354>
- Fink JL, Karunaratne S, Mittal A, Gardiner DM, Hamilton N, Mahony D, Kai C, Suzuki H, Hayashizaki Y, Teasdale RD (2008) Towards defining the nuclear proteome. *Genome Biol* 9:R15. <https://doi.org/10.1186/gb-2008-9-1-r15>
- Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39:W29–W37. <https://doi.org/10.1093/nar/gkr367>
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. *Nucleic Acids Res* 42:D222–D230. <https://doi.org/10.1093/nar/gkt1223>
- Fu X, Jiao W, Chang Z (2006) Phylogenetic and biochemical studies reveal a potential evolutionary origin of small heat shock proteins of animals from bacterial class A. *J Mol Evol* 62:257–266. <https://doi.org/10.1007/s00239-005-0076-5>
- Gaudet P, Livstone MS, Lewis SE, Thomas PD (2011) Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. *Brief Bioinform* 12:449–462. <https://doi.org/10.1093/bib/bbr042>
- Gause M, Eissenberg JC, MacRae AF et al (2006) Nipped-A, the Tra1/TRRAP subunit of the *Drosophila* SAGA and Tip60 complexes, has multiple roles in notch signaling during wing development. *Mol Cell Biol* 26:2347–2359. <https://doi.org/10.1128/MCB.26.6.2347-2359.2006>
- Gentil BJ, Cooper L (2012) Molecular basis of axonal dysfunction and traffic impairments in CMT. *Brain Res Bull* 88:444–453. <https://doi.org/10.1016/j.brainresbull.2012.05.003>
- Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N, Carroll S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL Jr, White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna M, Chant J, Rothberg JM (2003) A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736. <https://doi.org/10.1126/science.1090289>
- Gong WJ, Golic KG (2006) Loss of Hsp70 in *Drosophila* is pleiotropic, with effects on thermotolerance, recovery from heat shock and neurodegeneration. *Genetics* 172:275–286. <https://doi.org/10.1534/genetics.105.048793>
- Guruharsha KG, Rual J-F, Zhai B, Mintseris J, Vaidya P, Vaidya N, Beekman C, Wong C, Rhee DY, Cenaj O, McKillip E, Shah S, Stapleton M, Wan KH, Yu C, Parsa B, Carlson JW, Chen X, Kapadia B, VijayRaghavan K, Gygi SP, Celniker SE, Obar RA, Artavanis-Tsakonas S (2011) A protein complex network of *Drosophila melanogaster*. *Cell* 147:690–703. <https://doi.org/10.1016/j.cell.2011.08.047>
- Habermann K, Mirgorodskaya E, Gobom J, Lehmann V, Muller H, Blumlein K, Deery MJ, Czogiel I, Erdmann C, Ralsler M, von Kries JP, Lange BMH (2012) Functional analysis of centrosomal kinase substrates in *Drosophila melanogaster* reveals a new function of the nuclear envelope component otefin in cell cycle progression. *Mol Cell Biol* 32:3554–3569. <https://doi.org/10.1128/MCB.00814-12>
- Hao X, Zhang S, Timakov B, Zhang P (2007) The Hsp27 gene is not required for *Drosophila* development but its activity is associated with starvation resistance. *Cell Stress Chaperones* 12:364–372
- Inagaki N, Hayashi T, Arimura T, Koga Y, Takahashi M, Shibata H, Teraoka K, Chikamori T, Yamashina A, Kimura A (2006) Alpha B-crystallin mutation in dilated cardiomyopathy. *Biochem Biophys Res Commun* 342:379–386. <https://doi.org/10.1016/j.bbrc.2006.01.154>
- Joanisse DR, Inaguma Y, Tanguay RM (1998) Cloning and developmental expression of a nuclear ubiquitin-conjugating enzyme (DmUbc9) that interacts with small heat shock proteins in *Drosophila melanogaster*. *Biochem Biophys Res Commun* 244:102–109. <https://doi.org/10.1006/bbrc.1998.8214>



- Jones CA, Ng J, Peterson AJ, Morgan K, Simon J, Jones RS (1998) The *Drosophila* esc and E(z) proteins are direct partners in polycomb group-mediated repression. *Mol Cell Biol* 18:2825–2834
- Kang H, McElroy KA, Jung YL et al (2015) Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in *Drosophila*. *Genes Dev* 29:1136–1150. <https://doi.org/10.1101/gad.260562.115>
- Kappé G, Boelens WC, de Jong WW (2010) Why proteins without an  $\alpha$ -crystallin domain should not be included in the human small heat shock protein family HSPB. *Cell Stress Chaperones* 15:457–461. <https://doi.org/10.1007/s12192-009-0155-4>
- Kim KK, Kim R, Kim SH (1998) Crystal structure of a small heat-shock protein. *Nature* 394:595–599. <https://doi.org/10.1038/29106>
- Kloetzel P-M, Bautz EK (1983) Heat-shock proteins are associated with hnRNA in *Drosophila melanogaster* tissue culture cells. *EMBO J* 2:705–710
- Kriehuber T, Rattei T, Weinmaier T, Bepperling A, Haslbeck M, Buchner J (2010) Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *FASEB J* 24:3633–3642. <https://doi.org/10.1096/fj.10-156992>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Kunert N, Marhold J, Stanke J, Stach D, Lyko F (2003) A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development* 130:5083–5090. <https://doi.org/10.1242/dev.00716>
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948. <https://doi.org/10.1093/bioinformatics/btm404>
- Lee SY, Ramirez J, Franco M, Lectez B, Gonzalez M, Barrio R, Mayor U (2014) Ube3a, the E3 ubiquitin ligase causing Angelman syndrome and linked to autism, regulates protein homeostasis through the proteasomal shuttle Rpn10. *Cell Mol Life Sci* 71:2747–2758. <https://doi.org/10.1007/s00018-013-1526-7>
- Lei Y, Liu K, Hou L, Ding L, Li Y, Liu L (2017) Small chaperons and autophagy protected neurons from necrotic cell death. *Sci Rep* 7:5650. <https://doi.org/10.1038/s41598-017-05995-6>
- Liao P-C, Lin H-Y, Yuh C-H, Yu LK, Wang HD (2008) The effect of neuronal expression of heat shock proteins 26 and 27 on lifespan, neurodegeneration, and apoptosis in *Drosophila*. *Biochem Biophys Res Commun* 376:637–641. <https://doi.org/10.1016/j.bbrc.2008.08.161>
- Liu Z, Xie T, Steward R (1999) Lis1, the *Drosophila* homolog of a human lissencephaly disease gene, is required for germline cell division and oocyte differentiation. *Development* 126:4477–4488
- Maaroufi H, Tanguay RM (2013) Analysis and phylogeny of small heat shock proteins from marine viruses and their cyanobacteria host. *PLoS One* 8:e81207. <https://doi.org/10.1371/journal.pone.0081207>
- Marfori M, Lonhienne TG, Forwood JK, Kobe B (2012) Structural basis of high-affinity nuclear localization signal interactions with importin- $\alpha$ . *Traffic* 13:532–548. <https://doi.org/10.1111/j.1600-0854.2012.01329.x>
- Marin R, Tanguay RM (1996) Stage-specific localization of the small heat shock protein Hsp27 during oogenesis in *Drosophila melanogaster*. *Chromosoma* 105:142–149. <https://doi.org/10.1007/BF02509495>
- Michaud S, Marin R, Westwood JT, Tanguay RM (1997) Cell-specific expression and heat-shock induction of Hsps during spermatogenesis in *Drosophila melanogaster*. *J Cell Sci* 110(Pt 17):1989–1997
- Michaud S, Morrow G, Marchand J, Tanguay RM (2002) *Drosophila* small heat shock proteins: cell and organelle-specific chaperones? In: Arrigo A-P, Müller WEG (eds) *Small Stress Proteins*. Springer, Berlin Heidelberg, pp 79–101
- Michaud S, Lavoie S, Guimond M-O, Tanguay RM (2008) The nuclear localization of *Drosophila* Hsp27 is dependent on a monopartite arginine-rich NLS and is uncoupled from its association to nuclear speckles. *Biochim Biophys Acta Mol Cell Res* 1783:1200–1210. <https://doi.org/10.1016/j.bbamcr.2008.01.031>
- Miles WO, Jaffray E, Campbell SG, Takeda S, Bayston LJ, Basu SP, Li M, Raftery LA, Ashe MP, Hay RT, Ashe HL (2008) Medea SUMOylation restricts the signaling range of the Dpp morphogen in the *Drosophila* embryo. *Genes Dev* 22:2578–2590. <https://doi.org/10.1101/gad.494808>
- Morrow G, Tanguay RM (2015) *Drosophila* small heat shock proteins: an update on their features and functions. In: Tanguay RM, Hightower LE (eds) *The big book on small heat shock proteins*. Springer International Publishing, Cham, pp 579–606
- Moutaoufik MT, Morrow G, Finet S, Tanguay RM (2017a) Effect of N-terminal region of nuclear *Drosophila melanogaster* small heat shock protein DmHsp27 on function and quaternary structure. *PLoS One* 12:e0177821. <https://doi.org/10.1371/journal.pone.0177821>
- Moutaoufik MT, Morrow G, Maaroufi H, Féraud C, Finet S, Tanguay RM (2017b) Oligomerization and chaperone-like activity of *Drosophila melanogaster* small heat shock protein DmHsp27 and three arginine mutants in the alpha-crystallin domain. *Cell Stress Chaperones* 22:455–466. <https://doi.org/10.1007/s12192-016-0748-7>
- Müller H, Schmidt D, Steinbrink S, Mirgorodskaya E, Lehmann V, Habermann K, Dreher F, Gustavsson N, Kessler T, Lehrach H, Herwig R, Gobom J, Ploubidou A, Boutros M, Lange BMH (2010) Proteomic and functional analysis of the mitotic *Drosophila* centrosome. *EMBO J* 29:3344–3357. <https://doi.org/10.1038/emboj.2010.210>
- Nielsen MD, Luo X, Biteau B, Syverson K, Jasper H (2008) 14-3-3 Epsilon antagonizes FoxO to control growth, apoptosis and longevity in *Drosophila*. *Aging Cell* 7:688–699. <https://doi.org/10.1111/j.1474-9726.2008.00420.x>
- Palmeri D, Malim MH (1999) Importin  $\beta$  can mediate the nuclear import of an arginine-rich nuclear localization signal in the absence of importin  $\alpha$ . *Mol Cell Biol* 19:1218–1225
- Poulain P, Gelly J-C, Flatters D (2010) Detection and architecture of small heat shock protein monomers. *PLoS One* 5:e9990. <https://doi.org/10.1371/journal.pone.0009990>
- Razick S, Magklaras G, Donaldson IM (2008) iRefIndex: a consolidated protein interaction database with provenance. *BMC Bioinformatics* 9:405. <https://doi.org/10.1186/1471-2105-9-405>
- Rhee DY, Cho D-Y, Zhai B, Slattery M, Ma L, Mintseris J, Wong CY, White KP, Celniker SE, Przytycka TM, Gygi SP, Obar RA, Artavanis-Tsakonas S (2014) Transcription factor networks in *Drosophila melanogaster*. *Cell Rep* 8:2031–2043. <https://doi.org/10.1016/j.celrep.2014.08.038>
- Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320–W324. <https://doi.org/10.1093/nar/gku316>
- Roignant J-Y, Legent K, Janody F, Treisman JE (2010) The transcriptional co-factor Chip acts with LIM-homeodomain proteins to set the boundary of the eye field in *Drosophila*. *Development* 137:273–281. <https://doi.org/10.1242/dev.041244>
- Rørth P, Szabo K, Bailey A et al (1998) Systematic gain-of-function genetics in *Drosophila*. *Development* 125:1049–1057
- Rosenbaum EE, Brehm KS, Vasiljevic E, Liu CH, Hardie RC, Colley NJ (2011) XPORT-dependent transport of TRP and rhodopsin. *Neuron* 72:602–615. <https://doi.org/10.1016/j.neuron.2011.09.016>
- Scharf K-D, Siddique M, Vierling E (2001) The expanding family of Arabidopsis thaliana small heat stress proteins and a new family of proteins containing  $\alpha$ -crystallin domains (Acid proteins). *Cell Stress Chaperones* 6:225–237
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software

- environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504. <https://doi.org/10.1101/gr.1239303>
- Siddique M, Port M, Tripp J, Weber C, Zielinski D, Calligaris R, Winkelhaus S, Scharf KD (2003) Tomato heat stress protein Hsp16.1-CIII represents a member of a new class of nucleocytoplasmic small heat stress proteins in plants. *Cell Stress Chaperones* 8:381–394
- Siddique M, Gernhard S, von Koskull-Döring P, Vierling E, Scharf KD (2008) The plant sHSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties. *Cell Stress Chaperones* 13:183–197. <https://doi.org/10.1007/s12192-008-0032-6>
- Stamler R, Kappé G, Boelens W, Slingsby C (2005) Wrapping the alpha-crystallin domain fold in a chaperone assembly. *J Mol Biol* 353:68–79. <https://doi.org/10.1016/j.jmb.2005.08.025>
- Thurmond J, Goodman JL, Strelets VB, Attrill H, Gramates LS, Marygold SJ, Matthews BB, Millburn G, Antonazzo G, Trovisco V, Kaufman TC, Calvi BR, the FlyBase Consortium, Perrimon N, Gelbart SR, Agapite J, Broll K, Crosby L, Santos G, Emmert D, Gramates LS, Falls K, Jenkins V, Matthews B, Sutherland C, Tabone C, Zhou P, Zytkevich M, Brown N, Antonazzo G, Attrill H, Garapati P, Holmes A, Larkin A, Marygold S, Millburn G, Pilgrim C, Trovisco V, Urbano P, Kaufman T, Calvi B, Czoch B, Goodman J, Strelets V, Thurmond J, Cripps R, Baker P (2019) FlyBase 2.0: the next generation. *Nucleic Acids Res* 47:D759–D765. <https://doi.org/10.1093/nar/gky1003>
- UniProt Gene Ontology (GO) (2019) UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515. <https://doi.org/10.1093/nar/gky1049>
- Van De Klundert FAJM, De Jong WW (1999) The small heat shock proteins Hsp20 and  $\alpha$ B-crystallin in cultured cardiac myocytes: differences in cellular localization and solubilization after heat stress. *Eur J Cell Biol* 78:567–572. [https://doi.org/10.1016/S0171-9335\(99\)80022-3](https://doi.org/10.1016/S0171-9335(99)80022-3)
- van Montfort RL, Basha E, Friedrich KL et al (2001) Crystal structure and assembly of a eukaryotic small heat shock protein. *Nat Struct Biol* 8:1025–1030. <https://doi.org/10.1038/nsb722>
- Vicart P, Caron A, Guicheney P, Li Z, Prévost MC, Faure A, Chateau D, Chapon F, Tomé F, Dupret JM, Paulin D, Fardeau M (1998) A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy. *Nat Genet* 20:92–95. <https://doi.org/10.1038/1765>
- Wickham H (2009) *ggplot2: elegant graphics for data analysis*. Springer-Verlag, New York
- Ylikallio E, Konovalova S, Dhungana Y, Hilander T, Junna N, Partanen JV, Toppila JP, Auranen M, Tyynismaa H (2015) Truncated HSPB1 causes axonal neuropathy and impairs tolerance to unfolded protein stress. *BBA Clin* 3:233–242. <https://doi.org/10.1016/j.bbacli.2015.03.002>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.