761-768; Deprá, M., J.L. Poppe, H.J. Schmitz, D.C. De Toni, and V.L.S. Valente 2014, J. Pest Sci. 87: 379-383; De Toni, D.C., M.S. Gottschalk, J. Cordeiro, P.R.P. Hofmann, and V.L.S. Valente 2007, Neotrop. Entomol. 36: 356-375; Goñi, B., M.E. Martinez, V.L.S. Valente, and C.R. Vilela 1998, Rev. bras. Entomol. 42: 131-140; Goñi, B., M.E. Martinez, G. Techera, and P. Fresia 2002, Dros. Inf. Serv. 85: 75-80; Goñi, B., M. Remedios, P. Gonzalez-Vainer, M. Martinez, and C.R. Vilela 2012, Zoologia 29: 308-317; Hale, L.R., and R.S. Singh 1991, Genetics 129: 103-117; Martins, M.B. 1987, Bol. Mus. Para. Emílio Goeldi 3: 195-218; Mata, R.A., M. McGeoch, and R. Tidon 2008, Biodivers. Conserv. 17: 2899-2916; Poppe, J.L., H.J. Schmitz, S.M. Callegari-Jacques, and V.L.S. Valente 2015, Neotrop. Entomol. 44: 140-152; Poppe, J.L., H.J. Schmitz, D. Grimaldi, and V.L.S. Valente 2014, Zootaxa 3779: 215-245; Poppe, J.L., H.J. Schmitz, and V.L.S. Valente 2013, Neotrop. Entomol. 42: 269-277; Robe, L.J., S. Machado, A.R. Bolzan, J.P.J. Santos, F.B. Valer, A.P. Santos, M.L. Blauth, and M.S. Gottschalk 2014, Stud. Neotrop. Fauna Environ. 49: 79-94; Saavedra, C.C.R., S.M. Callegari-Jacques, M. Napp, and V.L.S. Valente 1995, J. Zool. Syst. Evol. Res. 33: 62-74; Schmitz, H.J., V.L.S. Valente, and P.R.P. Hofmann 2007, Neotrop. Entomol. 36: 53-64; Tidon-Sklorz, R., and F.M. Sene 1995, Iheringia Ser. Zool. 78: 85-94; Val, F.C., and M.D. Marques 1996, Pap. Avul. Zool. 39: 223-230; Wheeler, M.R., and L.E. Magalhães 1962, Univ. Texas Publ. 6205: 155-171.

A description of the adult *Drosophila* miRNome.

Fernández-Costa, J.M., M. Sabater-Arcís, and R. Artero. Incliva Health Research Institute, Valencia, Spain. Avda. Menéndez Pelayo 4 accesorio 46010 Valencia. Spain. Departament de Genética i Estructura de Recerca Interdisciplinar en Biotecnologia i Biomedicina (ERI BIOTECMED), Universitat de València. Dr Moliner 50, 46100 Burjassot, Spain. Correspondence to: ruben.artero@uv.es

Introduction

MicroRNAs (miRNAs, miRs) are short (18–24 nucleotides, nt) non-coding RNAs present in all eukaryotes, which play crucial roles during the post-transcriptional gene expression regulation (Ambros, 2004).

According to genomic localization, miRNAs can be classified as intergenic or intragenic. Most of the human miRNAs described up to date (deposited in miRBase) are intergenic (68%). Among the intragenic miRNAs, 12% are located in intronic regions and the others are located in coding exons, repetitive regions, long non-coding RNAs or non-coding regions. miRNA genes are often located near to other miRNA being part of clusters. Clusters of close miRNAs tend to be intergenic and are expressed as polycistronic, co-regulated units that contain their own promoters (Saini, Griffiths-Jones, and Enright, 2007). In contrast, intragenic miRNAs are typically co-expressed with their host gene, although instances of independent transcriptional regulation have also been reported in intronic miRNA. Additionally differences in the expression of polycistronic mRNA located in introns have been associated to regulated alternative splicing (Bell, Buvoli, and Leinwand, 2010).

miRNA biogenesis is a multi-step process requiring several enzymes. In the canonical biogenesis miRNAs are usually transcribed by RNA polymerase II as long primary transcripts (pri-miRNA) (Y. Lee *et al.*, 2004), although, it has also been described pri-miRNA transcription by RNA polymerase III (Babiarz, Ruby, Wang, Bartel, and Blelloch, 2008). Pri-miRNA presents 7-methylguanosine cap (m^7G) in 5' end and they are polyadenylated in 3' end (Cai, Hagedorn, and Cullen, 2004). First, the maturation process begins in the nucleus, where pri-miRNA is folded into a characteristic hairpin structure that is asymmetrically and specifically cut near the stem-loop by RNase III Drosha, generating one or more hairpin structures known as pre-miRNA (approximately 65 nt) (Y. Lee *et al*., 2003). Drosha needs the binding of RNA protein DGCR8 (Pasha in *Drosophila*) as co-factor, which together form the microprocessor complex (Denli, Tops, Plasterk, Ketting, and Hannon, 2004). Second, pre-miRNAs are then transported to the cytoplasm by exportin 5

(EXP5), and further processed by RNase III Dicer (Dicer-1 in *Drosophila*) to generate a small heteroduplex miRNA (15-21 nt) (Lee *et al*., 2003). Finally, one strand of the RNA duplex is released by Dicer and the resulting single strand is integrated together with Argonaute proteins into RISC (RNA-induced silencing complex) complex (Gregory, Chendrimada, Cooch, and Shiekhattar, 2005). By guiding the RISC complex to bind to target ''seed match'' sites within the 3' untranslated region (UTR) of mRNAs, the mature miRNA can silencing gene expression by mRNA cleavage, when the sequence is perfectly complementary with the target (Bartel, 2004), or by translation repression or RNA deadenylation, if complementarity is not perfect (Wu, Fan, and Belasco, 2006). Evidence has shown that some miRNAs can also suppress the expression of their target mRNA by binding to the 5'UTR (Lytle, Yario, and Steitz, 2007) or open reading frame (Forman, Legesse-Miller, and Coller, 2008).

Since the first miRNA was discovered in *Caenorhabditis elegans* in 1993 (R.C. Lee, Feinbaum, and Ambros, 1993), miRNAs have been identified in insects, vertebrates, plants, and virus. Computational methods have predicted that miRNAs account for ~1% of all eukaryotic genes, and more than 60% of human genes might be subject to regulation by miRNA (Zhonglin *et al*. 2015). Even though the biological functions of most animal miRNAs are little known, increasing evidence suggests that miRNAs play important roles in diverse physiological processes such as homeostasis, development, proliferation, differentiation, apoptosis, or immune defence (Lu and Liston, 2009). Therefore, dysregulation of gene expression that encode miRNAs can contribute to the development of human diseases including cancer, cardiophaties, metabolic diseases, and neurodegenerative diseases (Abe and Bonini, 2013; Fernández-Hernando *et al*., 2013; Iorio and Croce, 2012).

New massive sequencing techniques allow us to obtain millions of readings from different sequences of several samples in parallel. Due to its high yield, the results of these platforms can be used for different types of studies. For example, sequencing from small RNA libraries can be used to study miRNA transcriptome (miRNome) profile present in a sample, allowing their identification, characterization, and quantification.

Because of miRNA's crucial role in most biological processes, we decided to use massive sequencing technology to carry out a complete description of which miRNAs are expressed in *Drosophila melanogaster* adult flies, also determining miRNA expression levels. We have found 104 of the 153 miRNA previously identified in *Drosophila melanogaster* obtaining expression levels between 10 and 10^4 reads/million. Our data contribute to the description of miRNAs normal expression levels in adult flies, and provide information to analyse the expression of miRNA altered in many pathologies using *Drosophila melanogaster* as a model.

Experimental Procedure

Small RNA library generation and next-generation sequencing

In order to describe all miRNAs expressed in *Drosophila melanogaster*, we analysed the expression of wild-type strain *OrR*. Two biological replicates of *OrR* were used, each of which containing 50 *Drosophila* males of the same age (2-day-old). Total RNA was extracted and the small RNA fraction was enriched using the miRVana kit (Ambion). Small RNA was run in 15% acrylamide:bisacrylamide 19:1 gels and the 15–30 nt fraction was sliced out and eluted with 1 M NaCl overnight at 48ºC. Purification was carried out using the MEGAclear Kit (Ambion). The quality of purified small RNAs (50 ng) was analysed by capillary electrophoresis (Agilent 2100 Bioanalyzer). Libraries for SOLiD™ 3 sequencing were prepared following the manufacturer's protocol (Small RNA Expression Kit, Applied Biosystems). Briefly, small RNA samples (15 ng) were hybridized and ligated overnight with the adapter mix, reverse transcribed and PCR-amplified (15 cycles). The primers used in this PCR included a unique six-nucleotide barcode for each sample. A single emulsion PCR reaction was used to couple the barcoded libraries to P1-coated beads as per the standard Applied Biosystems protocol. After emulsion PCR, template beads were enriched in a glycerol gradient and deposited onto the surface of glass slides for SOLiD sequencing. Sequencing was performed using 35 bp chemistry on a version 3.0 SOLiD machine (SOLiD™ 3).

Bioinformatics analysis

An average of 29 million readings were obtained from the SOLiD™ 3 sequencing data (Table 1). Low-quality reads were first removed from the data set (at least QV≥ 10 in the first 10 bases) . Filtered reads

were then mapped against the *Drosophila melanogaster* genome (version r5.23), using the software Small_RNA_Tool_v0.5.0 (http://solidsoftwaretools.com/gf/) and allowing up to two mismatches in the first 8 nt and up to three mismatches in the entire read. A custom pipeline was then used to select reads that mapped uniquely to a point of the *Drosophila* genome, which represented the usable sequence data. mirBase (version 13.0) and custom scripts were applied in order to identify known miRNAs. Contaminations by protein-coding genes or other ncRNAs (rRNAs, tRNAS, snoRNAs, and so forth) were discarded from the final data set. miRNA counts were normalized per million of reads that mapped uniquely, in order to yield the relative transcript abundance in the original sample (Figure 1).

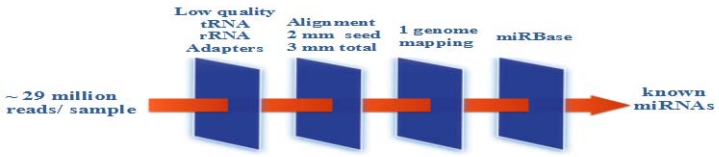

Figure 1. Representative diagram with the steps taken to miRNome characterization of adult flies by massive sequencing of small RNA libraries. mm = mismatch.

Results and Discussion

Using SOLiD™ 3 sequencing of small-RNA libraries and following bioinformatics analysis described above, we obtained the miRNome profile for two samples of wild-type adult *Drosophila melanogaster* flies (strain *OrR*, samples O_1 and O_2). Readings for 104 of the 153 known miRNA up to date of the analysis (mirBase v.13) were detected, so approximately 70% of miRNAs described are expressed in adult flies. In addition, there are large differences in the expression of these miRNA, setting a minimum of 10 reads per million. Therefore, there is a small group of seven miRNAs (*miR-1, miR-14, miR-277, miR-276a, miR-279, miR-8, and Bantam*) with a high expression (>10^4 reads) compared to the others, a group of 29 miRNAs with an intermediate expression (>10^3 reads), and two groups of 33 and 35 miRNAs with a very low expression (>10^2 and >10 reads) (Table 2).

Table 1. Number of reads obtained from massive sequencing and after bioinformatics processing.

			Mapping into 1 or more sites		Mapping into a single site	
Library	Total reads	QV[1]	Filtered reads[2]	Mapping into the genome	Filtered reads[2]	Mapping into the genome
O_1	28691891	21418676	3120769	2066628	2668549	**910374**
O_2	29572378	21965363	3254335	1719574	3175191	**867498**

[1]Reads after quality control [2]tRNA, rRNAs, adapters.

The expression differences between all miRNAs analysed suggests that the muscular and nervous tissues contribute the most to adult miRNA expression, considering that *miR-1* (miRNA with the highest expression in adult flies) is involved in myogenesis and cardiogenesis (Kwon, Han, Olson, and Srivastava, 2005), and five members belonging to this high expression group, *bantam, miR-8, miR-276a, miR-277,* and

Table 2. Number of reads per million for miRNAs in *Drosophila melanogaster* adult males.

>10^4 reads			>10^3 reads			>10^2 reads			>10 reads		
miRNA	average SOLiD reads	standard error	miRNA	average SOLiD reads	standard error	miRNA	average SOLiD reads	standard error	miRNA	average SOLiD reads	standard error
dme-mir-1	85238.5	387.5	*dme-mir-317*	9768	3229.0	*dme-mir-956*	967	661.0	*dme-mir-2b-1*	82	62.0
dme-mir-14	56861.5	1936.5	*dme-mir-33*	9642.5	2543.5	*dme-mir-79*	870.5	603.5	*dme-mir-927*	78	25.0
dme-mir-277	45842.5	6613.5	*dme-mir-252*	6718.5	2141.5	*dme-mir-87*	795.5	72.5	*dme-mir-1010*	77	8.0
dme-mir-276a	37744.5	7709.5	*dme-mir-11*	6700	85.0	*dme-mir-263b*	714.5	432.5	*dme-mir-13a*	72	2.0
dme-mir-279	27764.5	7966.5	*dme-mir-307*	6633	1066.0	*dme-mir-92a*	668	508.0	*dme-mir-965*	62.5	14.5
dme-mir-8	13732	7065.0	*dme-mir-263a*	6593	1132.0	*dme-mir-986*	638.5	287.5	*dme-mir-4*	52.5	52.5
dme-bantam	12388	3900.0	*dme-mir-34*	6442.5	1186.5	*dme-mir-957*	635.5	15.5	*dme-mir-2b-2*	48.5	34.5
			dme-mir-12	6080.5	2754.5	*dme-mir-274*	569.5	244.5	*dme-mir-2c*	48	4.0
			dme-mir-210	4767.5	802.5	*dme-mir-190*	567	218.0	*dme-mir-137*	46	14.0
			dme-mir-125	4443	1841.0	*dme-mir-iab-4*	551.5	310.5	*dme-mir-978*	44	15.0
			dme-mir-184	3786	391.0	*dme-mir-304*	469.5	224.5	*dme-mir-318*	43	43.0
			dme-let-7	3749.5	1139.5	*dme-mir-960*	440.5	24.5	*dme-mir-963*	37.5	2.5
			dme-mir-958	2374.5	1256.5	*dme-mir-375*	393	212.0	*dme-mir-1015*	35.5	0.5
			dme-mir-989	2248	2248.0	*dme-mir-285*	365.5	50.5	*dme-mir-984*	33	1.0
			dme-mir-7	2132.5	497.5	*dme-mir-2a-2*	335.5	39.5	*dme-mir-999*	32	11.0
			dme-mir-305	1966.5	869.5	*dme-mir-987*	295.5	28.5	*dme-mir-976*	30.5	2.5
			dme-mir-124	1897	1436.0	*dme-mir-995*	293	180.0	*dme-mir-1004*	28	28.0
			dme-mir-100	1847.5	314.5	*dme-mir-932*	290	100.0	*dme-mir-313*	25.5	25.5
			dme-mir-306	1797.5	417.5	*dme-mir-312*	288.5	162.5	*dme-mir-977*	23	2.0
			dme-mir-284	1781.5	174.5	*dme-mir-316*	265	5.0	*dme-mir-1007*	23	5.0
			dme-mir-1003	1677	239.0	*dme-mir-9b*	256.5	145.5	*dme-mir-985*	22	12.0
			dme-mir-9a	1627.5	447.5	*dme-mir-998*	248	54.0	*dme-mir-929*	21.5	1.5
			dme-mir-10	1476	22.0	*dme-mir-193*	240	117.0	*dme-mir-286*	20.5	20.5
			dme-mir-276b	1377	295.0	*dme-mir-133*	236	147.0	*dme-mir-282*	19.5	8.5
			dme-mir-31a	1298	252.0	*dme-mir-1000*	231	11.0	*dme-mir-982*	19	4.0
			dme-mir-994	1286	1286.0	*dme-mir-31b*	209	11.0	*dme-mir-975*	19	0.0
			dme-mir-988	1173.5	312.5	*dme-mir-9c*	190	112.0	*dme-mir-283*	18	18.0
			dme-mir-993	1090.5	358.5	*dme-mir-311*	182.5	17.5	*dme-mir-962*	17.5	5.5
			dme-mir-970	1035.5	505.5	*dme-mir-92b*	170.5	72.5	*dme-mir-964*	16	2.0
						dme-mir-2a-1	169	38.0	*dme-mir-959*	16	4.0
						dme-mir-1012	141.5	27.5	*dme-mir-968*	13	13.0
						dme-mir-1017	138	40.0	*dme-mir-219*	10.5	10.5
						dme-mir-310	125.5	20.5	*dme-mir-1006*	10	10.0
						dme-mir-278	118	29.0			
						dme-mir-996	117	48.0			

miR-279, are involved in neural processes (Karres, Hilgers, Carrera, Treisman, and Cohen, 2007; Li *et al*., 2013; Tan, Poidevin, Li, Chen, and Jin, 2012; Sun, Jee, de Navas, Duan, and Lai, 2015). Furthermore, *miR-14*, the miRNA that completes this group, and *bantam*, were shown to regulate programmed cell death in *Drosophila* (Jovanovic and Hengartner, 2006).

These evidences may be relevant to the study of miRNA involvement in some diseases in which nervous and muscular tissues are most affected using *Drosophila* as a model organism. Neuromuscular diseases caused by repeat expansions like Myotonic Dystrophy type 1 (DM1) are a case in point. Fernandez-Costa *et al*. studying the changes in the muscle miRNome of a *Drosophila* model expressing CTG repeats by SOLiD™ 3 sequencing, demonstrate that the expression of 20 miRNA was affected by expression of CTG repeats in DM1 flies (Fernandez-Costa *et al*., 2013). The dysregulation of a specific miRNA conserved between *Drosophila* and human, miR-1, had been reported previously in DM1 patients (Rau *et al.*, 2011). Therefore, the information obtained in the present work could be helpful to perform a comparative study of the miRNA expression levels between this DM1 fly model and *OrR* wild-type flies. In addition to comparative studies, the characterization of adult *OrR* miRNome and the expression levels for known miRNAs, could serve to identify new miRNA using prediction algorithms like miRDeep (Friedländer *et al*., 2008).

In summary, in this study we characterized which miRNAs are expressed in wild-type adult flies, providing information to compare the miRNA expression profile altered in different diseases, such as cancer, cardiophaties, metabolic diseases, or neurodegenerative diseases, or to select miRNA altered in human disease to study the implication of these miRNAs in the pathology using *Drosophila* as a model.

References: Abe, M., and N.M. Bonini 2013, Trends in Cell Biology 23: 30-36; Ambros, V., 2004, Nature 431: 350–5; Babiarz, J.E., J.G. Ruby, Y. Wang, D.P. Bartel, and R. Blelloch 2008, Genes and Development 22: 2773–2785; Bartel, D.P., 2004, Cell 116: 281–97; Bell, M.L., M. Buvoli, and L.A. Leinwand 2010, Molecular and cellular biology 30: 1937–1945; Cai, X., C.H. Hagedorn, and B.R. Cullen 2004, RNA (New York, N.Y.) 10: 1957–1966; Denli, A.M., B.B.J. Tops, R.H. Plasterk, R.F. Ketting, and G.J. Hannon 2004, Nature 432: 231–235; Fernandez-Costa, J.M., A. Garcia-Lopez, S. Zuñiga, V. Fernandez-Pedrosa, A. Felipo-Benavent, M. Mata, O. Jaka, A. Aiastui, B. Aguado, M. Perez-Alonso, J.J. Vilchez, A. Lopez de Munain, and R.D. Artero 2013, Human molecular genetics 22: 704–16; Forman, J.J., A. Legesse-Miller, and H.A. Coller 2008, Proceedings of the National Academy of Sciences of the United States of America 105: 14879–84; Fernández-Hernando, C., C.M. Ramírez, L. Goedeke, and Y. Suárez 2013, Arteriosclerosis, thrombosis, and vascular biology 33: 178-185; Friedländer, M., W. Chen, C. Adamidi, J. Maaskola, R. Einspanier, S. Knespel, and N. Rajewsky 2008, Nature biotechnology 26: 407-415; Gregory, R.I., T.P. Chendrimada, N. Cooch, and R. Shiekhattar 2005, Cell 123: 631–640; Iorio, M.V., and C.M. Croce 2012, Carcinogenesis 33: 1126–33; Jovanovic, M., and M.O. Hengartner 2006, Oncogene 25: 6176–6187; Karres, J.S., V. Hilgers, I. Carrera, J. Treisman, and S.M. Cohen 2007, Cell 131: 136–45; Kwon, C., Z. Han, E.N. Olson, and D. Srivastava 2005, Proceedings of the National Academy of Sciences of the United States of America 102: 18986–18991; Lee, R.C., R.L. Feinbaum, and V. Ambros 1993, Cell 843-54; Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Rådmark, S. Kim, and V.N. Kim 2003, Nature 425: 415–419; Lee, Y., M. Kim, J. Han, K.-H. Yeom, S. Lee, S.H. Baek, and V.N. Kim 2004, The EMBO journal 23: 4051–4060; Li, W., M. Cressy, H. Qin, T. Fulga, D. Van Vactor, and J. Dubnau 2013, The Journal of neuroscience : the official journal of the Society for Neuroscience 33: 5821–33; Lu, L.-F., and A. Liston 2009, Immunology 127: 291–8; Lytle, J.R., T.A. Yario, and J.A. Steitz 2007, Proceedings of the National Academy of Sciences of the United States of America, 104: 9667–9672; Rau, F., F. Freyermuth, C. Fugier, J.-P. Villemin, M.-C. Fischer, B. Jost, D. Dembele, *et al.* 2011, Nature structural and molecular biology 18: 840–845; Saini, H.K., S. Griffiths-Jones, and A.J. Enright 2007, Proceedings of the National Academy of Sciences of the United States of America 104: 17719–17724; Sun, K., D. Jee, L.F. de Navas, H. Duan, and E.C. Lai 2015, PLoS genetics 11: e1005245; Tan, H., M. Poidevin, H. Li, D. Chen, and P. Jin 2012, PLoS genetics 8: e1002681; Wu, L., J. Fan, and J.G. Belasco 2006, Proceedings of the National Academy of Sciences of the United States of America, 103: 4034–4039; Zhonglin, T., Y. Yalan, W. Zishuai, Z. Shuanping, M. Yulian, and L. Kui 2015, Nature 5: 15544.