Supplementary Tables

Supplementary data includes 11 Supplementary Tables (Table S1-S11).

High-resolution analysis of the human retina miRNome reveals isomiR variations and novel microRNAs

Marianthi Karali[1,2], Maria Persico[1], Margherita Mutarelli[1], Annamaria Carissimo[1], Mariateresa Pizzo[1], Veer Singh Marwah[1], Concetta Ambrosio[1], Michele Pinelli[1], Diego Carrella[1], Stefano Ferrari[3], Diego Ponzin[3], Vincenzo Nigro[1,2], Diego di Bernardo[1,4*] and Sandro Banfi[1,2*]

[1]Telethon Institute of Genetics and Medicine, via Campi Flegrei 34, 80078 Pozzuoli (NA), Italy.

[2]Department of Biochemistry, Biophysics and General Pathology, Second University of Naples, via Luigi De Crecchio 7, 80138 Naples (NA), Italy.

[3] Eye Bank of Venice,, Padiglione Rama, via Paccagnella 11, 30174 Zelarino (VE), Italy.

[4]Department of Chemical, Materials and Production Engineering, University of Naples "Federico II", via Claudio 21, 80125 Naples (NA), Italy.

Table S1

Table S1: Eye donor description and sample features

Sample	Gender	Age	Cause of death	Post-mortem time[1]	Available Material
1	Female	68	Neoplastic Disease	25h 15min	Retina
2	Male	52	Neoplastic Disease	21h 30min	Retina
3	Male	67	Neoplastic Disease	22h 50min	Retina
4	Female	70	Neoplastic Disease	25h 06min	Retina
5	Male	72	Cardiovascular Disease	24h 55min	Retina
6	Male	60	Neoplastic Disease	25h 25min	Retina
8	Male	71	Neoplastic Disease	21h 35min	Retina
9	Female	71	Neoplastic Disease	14h 15min	Retina
10	Male	70	Neoplastic Disease	23h 25min	Retina
11	Male	49	Injury/Poisoning	23h 30min	Retina
12	Female	63	Neoplastic Disease	19h 15min	Retina
13	Female	59	Neoplastic Disease	24h 40min	Retina
14	Female	69	Neoplastic Disease	21h	Retina
15	Male	71	Neoplastic Disease	22h 15min	Retina
16	Male	64	Neoplastic Disease	26h	Retina
17	Male	42	Injury/Poisoning	21h 50min	Retina
27	Male	49	Neoplastic Disease	7h 15min	Retina; RPE/Choroid
34	Male	59	Neoplastic Disease	22h 45min	Retina; RPE/Choroid
51	Female	62	Neoplastic Disease	4h 15min	Sections for RNA ISH

[1] Total time from the moment of death to the storage of the dissected tissue in an RNA stabilisation (samples 1-34; Qiagen RNAlater) or fixative (sample 51; 4% Paraformaldehyde) solution

Table S2

Table S2: Read mapping report of the retina and RPE/choroid libraries

	Raw Read Trimming		Trimmed Read Mapping			Arm Prevalence	
small RNA Library	**Total Reads**	**After Trimming**	**to mature miRNAs[1]**	**to miRNA precursors[1]**	**to genomic locations only**	**mapping to 5p**	**mapping to 3p**
Human Retina 1	35,470,443	31,231,604	10,567,173	15,698,451	1,547,352	9,307,817	1,219,268
Human Retina 2	47,457,813	39,523,926	12,875,238	18,530,728	3,092,401	11,191,662	1,635,484
Human Retina 3	37,909,111	33,652,419	11,299,080	16,721,061	1,943,334	9,737,532	1,506,302
Human Retina 4	46,574,069	40,152,704	12,769,350	19,336,565	3,076,529	11,182,232	1,536,986
Human Retina 5	37,066,117	30,958,377	9,396,509	15,326,313	1,945,722	8,343,981	1,014,286
Human Retina 6	39,976,532	34,812,214	12,393,637	17,215,615	1,518,832	10,874,102	1,467,740
Human Retina 8	25,455,497	22,185,206	7,317,591	9,544,375	2,024,157	6,444,875	847,585
Human Retina 9	28,517,167	25,005,644	9,241,450	12,695,083	1,052,165	8,013,415	1,189,058
Human Retina 10	36,739,708	29,832,632	9,053,951	12,916,238	2,524,634	7,975,294	1,044,006
Human Retina 11	30,207,308	26,169,267	8,463,469	12,827,856	1,512,293	7,431,817	995,924
Human Retina 12	25,209,504	21,517,631	7,745,961	11,751,155	1,936,099	6,627,586	1,079,885
Human Retina 13	24,852,883	22,800,625	8,589,950	11,224,437	832,508	7,421,037	1,127,562
Human Retina 14	28,818,023	25,100,992	8,204,659	12,668,614	1,251,767	7,040,462	1,114,862
Human Retina 15	27,761,855	25,035,277	8,530,789	11,959,915	1,283,845	7,459,308	1,039,818
Human Retina 16	27,438,557	23,686,032	8,035,645	10,995,729	1,356,042	6,926,467	1,071,127
Human Retina 17	29,619,602	24,912,781	8,431,289	11,105,179	1,710,869	7,319,226	1,075,786
Human RPE/Choroid 27	28,954,136	24,861,947	10,184,418	7,837,396	2,332,605	8,725,343	1,407,301
Human RPE/Choroid 34	23,667,762	21,025,935	8,751,517	6,059,284	1,582,292	7,678,015	1,039,878

[1] Mapped to the miRBase v20

Table S3

Table S3: Complete ranking of retina expressed miRNAs and analysis of variation

Complete ranking of retina expressed miRNAs (averaged across 16 samples) prior to and following isomiR inclusion (see spreadsheets labelled **“prior to isomiR inclusion”** and **“following isomiR inclusion”**, respectively). The miRNA expression profiles of each individual retina sample (following isomiR inclusion) and the coefficient of variation across all samples is shown in the spreadsheet labelled **“Coefficient of variation”**.

(see separate Excel file containing three spreadsheets)

Table S4

Table S4: Comparison of miRNA expression profiles between the human and mouse retina

We compared the miRNA expression profiles of the human retina established in this study with the expression profile of miRNAs in the murine retina reported by Soundara Pandi et al (2013) *Investigative ophthalmology & visual science*, 54, 8140-8151 (25). We obtained a list of miRNAs that were identified in both the human and mouse retina (see spreadsheet labelled **"common_mouse_human_retina"**). We also obtained a list of miRNAs that were detected either in the human retina alone (see spreadsheet labelled **"human_retina_only")** or in the mouse retina alone (see spreadsheet labelled **"mouse_retina_only"**).

The last rows of each spreadsheet show the sum of expression values (as reads) across all miRNAs listed therein and the total number of reads obtained when considering both common and species-specific miRNAs. Mean expression values in the human retina were normalised by TMM, as previously described.

(see separate Excel file containing three spreadsheets)

Table S5

Table S5: List of identified isomiRs for each retina-expressed miRNA in all analysed samples and contribution of isomiRs with a modified or reference seed sequence to the total read counts of a given miRNA

List of isomiRs for each retina-expressed miRNA identified either in at least one (see spreadsheet labelled **“identified in >=1 sample”**) or in all analysed retina samples (see spreadsheet labelled **“identified in all 16 samples”**). The spreadsheet header annotation is as follows: Sequence_ID: Identifier of each read sequence, which indicates whether the read corresponds to the reference sequence (_REF) or to an isomiR (_isomiR); miRNA Name: miRNA to which a given isomiR is associated; Read Sequence: Sequence of the identified read; Average Frequency: average frequency as counts per million of a given variant in the samples it was detected; Detection (out of 16): number of samples (out of the 16 analysed) in which a given sequence variant was detected; nucl. substitution: nucleotide substitution position-nucleotide at sequence-nucleotide at precursor; additions: non-templated nucleotides added at the 3’-end; t5: nucleotides at 5’-end different from the annotated mature sequence in miRBase (“u”: nucleotides added, “d”: nucleotides trimmed); t3: nucleotides at 3’-end different from the annotated mature sequence in the miRBase (“d”: nucleotides added, “u”: nucleotides trimmed); modified seed (Yes/No): isomiR variations resulting in a modified seed. The “Relative frequency” column (in the spreadsheet labelled “identified in all 16 samples”) reports the relative frequency of each sequence variant as a percentage of the total reads for the respective miRNA.

The spread-sheet **“contribution isomiRs mod. seed”** presents the contribution of the reference miRNA (column C) and the isomiRs (column D) as a percentage of total read counts for a given miRNA. The percentage of total read counts that derives from isomiRs bearing a modified seed sequence (column E) or from variants with the reference seed sequence (column F) is also provided. Only variants identified in all retina samples we considered for this analysis. The mean expression values of each miRNA after TMM normalisation are given in column B.

(see separate Excel file containing three spreadsheets)

Table S6

Table S6: Independent evidence of expression for the isomiRs of the 100 top-ranked miRNAs that bear a modified seed and have a relative frequency (RF) ≥1%

For the isomiRs of the 100 top-ranked miRNAs that bear a modified seed and have an RF≥1% (n=117) we looked for evidence of expression from two independent resources:

(a) the YM500v2 database (http://ngs.ym.edu.tw/ym500v2/isomir.php)
Cheng, W.C. et al. (2015). *Nucleic Acids Res*, 43, D862-867, and

(b) Loher, P. et al. (2014). *Oncotarget*, 5, 8790-8802.

The spreadsheet header annotation is as follows: Sequence_ID_Retina: Identifier of each read sequence, which indicates whether the read corresponds to the reference sequence (_REF) or to an isomiR (_isomiR); miRNA Name: miRNA to which a given isomiR is associated; Read Sequence: Sequence of the identified read; Average Frequency: average frequency as counts per million of a given variant in the samples it was detected; Relative frequency: relative frequency of each sequence variant as a percentage of total reads for the respective miRNA; reported in YM500v2 (Cheng et al.): evidence of expression of a given isomiR in the YM500v2 database; reported in Loher et al.: evidence of presence of a given isomiR in the Argonaute PAR-CLIP data from human lymphoblastoid cell lines reported in Loher et al..

(see separate Excel file)

Table S7

Table S7: Complete list of hsa-miR-124-3p sequence variants identified in the human retina

		Sequence variation					
hsa-miR-124-3p: read sequence (5' to 3') [1]	**Average frequency**[2]	**5'-end**[3]	**3'-end**[4]	**3'-end untemplated additions**[5]	**mismatch within seed**[6]	**Seed change**[7]	**Relative frequency**[8]
TAAG**GCACGCGGTGAAT**GCC	148002	0	0	0	0	No	46.7%
TAAG**GCACGCGGTGAAT**GC	44337	0	u-C	0	0	No	14.0%
TAAG**GCACGCGGTGAAT**GCCA	24311	0	d-A	0	0	No	7.7%
TAAG**GCACGCGGTGAAT**G	6927	0	u-CC	0	0	No	2.2%
TAAG**GCACGCGGTGAAT**GCCG	5090	0	0	u-G	0	No	1.6%
TAAG**GCACGCGGTGAAT**GCCT	4938	0	0	u-T	0	No	1.6%
TAAG**GCACGCGGTGAAT**GCCAA	2939	0	d-AA	0	0	No	0.9%
TAAG**GCACGCGGTGAAT**GCCAAG	2212	0	d-AAG	0	0	No	0.7%
TTAAG**GCACGCGGTGAAT**GCC	40022	u-T	0	0	0	Yes	12.6%
TTAAG**GCACGCGGTGAAT**GC	17501	u-T	u-C	0	0	Yes	5.5%
TTAAG**GCACGCGGTGAAT**	7288	u-T	u-GCC	0	0	Yes	2.3%
TTAAG**GCACGCGGTGAAT**GCCA	5225	u-T	d-A	0	0	Yes	1.7%
TTAAG**GCACGCGGTGAAT**G	3443	u-T	u-CC	0	0	Yes	1.1%
TAAT**GCACGCGGTGAAT**GCC	4403	0	0	0	4TG	Yes	1.4%
all reads with seed: AAGGCAC	**238755**	**0**	**various**	**various**	**0**	**No**	**75.4%**
all reads with seed: TAAGGCA	**73480**	**u-T**	**various**	**various**	**0**	**Yes**	**23.2 %**
all reads with seed: AATGCAC	**4403**	**0**	**0**	**0**	**4TG**	**Yes**	**1.4%**

[1] Entries are ranked based on the read abundance for each of the three different seed sequences (highlighted in bold)

[2] Average frequency: Average frequency in counts per million across 16 samples

[3] 5’-end: variation at 5’-end with respect to the canonical (u-: nucleotide addition)

[4] 3’-end: variation at 3’-end with respect to the canonical (d: templated nucleotide additions, u: nucleotides trimmed)

[5] 3’-end untemplated additions: additions at the 3’-end that do not match the precursor sequence in the miRBase

[6] mismatch within seed: mismatches within the seed sequence, number indicates position of mismatch within the isomiR sequence

[7] Seed change: altered seed sequence due to variations at the 5’-end and/or mismatches

[8] Relative frequency: percentage of reads of each sequence variant with respect to the total reads of this miRNA.

Data refer to the sequence variants commonly present across all samples analysed (n=16).

Table S8

Table S8: Complete ranking of miRNAs expressed in the RPE/choroid

Complete ranking of miRNAs expressed in the RPE/choroid, prior to and following isomiR inclusion (see spreadsheets labelled **“prior to isomiR inclusion”** and **“following isomiR inclusion”**, respectively).

(see separate Excel file containing two spreadsheets)

Table S9

Table S9: List of identified isomiRs for each RPE/choroid-expressed miRNA and contribution of isomiRs with a modified or reference seed sequence to the total read counts of a given miRNA

List of isomiRs for each RPE/choroid-expressed miRNA identified either in at least one (see spreadsheet labelled **“identified in >=1 sample”)** or in both analysed RPE/choroid samples (see spreadsheet labelled **“identified in both samples”**). The spreadsheet header annotation is as follows: Sequence_ID: Identifier of each read sequence, which indicates whether the read corresponds to the reference sequence (_REF) or to an isomiR (_isomiR); miRNA Name: miRNA to which a given isomiR is associated; Read Sequence: Sequence of the identified read; Average Frequency: average frequency as counts per million of a given variant in the samples it was detected; Detection (out of 2): number of samples (out of the 2 analysed) in which a given sequence variant was identified; nucl. substitution: nucleotide substitution position-nucleotide at sequence-nucleotide at precursor; additions: non-templated nucleotides added at the 3’-end; t5: nucleotides at 5’-end different from the annotated mature sequence in the miRBase (“u”: nucleotides added, “d”: nucleotides trimmed); t3: nucleotides at 3’-end different from the mature annotated sequence in the miRBase (“d”: nucleotides added, “u”: nucleotides trimmed); modified seed (Yes/No): isomiR variations resulting in a modified seed. The “Relative frequency” column (in the spreadsheet labelled “identified in both samples”) reports the relative frequency of each sequence variant as a percentage of total reads for the respective miRNA.

The spread-sheet **“contribution isomiRs mod. seed”** presents the contribution of the reference miRNA (column C) and the isomiRs (column D) as a percentage of total read counts for a given miRNA. The percentage of total read counts that derives from isomiRs bearing a modified seed sequence (column E) or from variants with the reference seed sequence (column F) is also provided. Only variants identified in both samples were considered for this analysis. The mean expression values of each miRNA after TMM normalisation are given in column B.

(see separate Excel file containing three spreadsheets)

Table S10

Table S10: Comparison of miRNA expression between the human and mouse RPE/choroid

We compared the miRNA expression profiles of the human RPE/choroid established in this study with the expression profile of miRNAs in the murine RPE/choroid reported by Soundara Pandi et al (25). We obtained a list of miRNAs that were identified in the RPE/choroid of both species (see spreadsheet labelled **“common_human_mouse_RPEchoroid”**). We also obtained a list of miRNAs that were detected either in the human RPE/choroid alone (see spreadsheet labelled **“human_RPEchoroid_only”**) or in the mouse RPE/choroid alone (see spreadsheet labelled **“mouse_RPEchoroid_only”**).

The last rows of each spreadsheet show the sum of expression values (as reads) across all miRNAs listed therein and the total number of reads obtained when considering both common and species-specific miRNAs. Mean expression values in the human RPE/choroid were normalised by TMM as previously described.

(see separate Excel file containing three spreadsheets)

Table S11

Table S11: Analysis of isomiR tissue-specificity for the miRNAs that are expressed both in the retina and RPE/choroid samples

List of isomiRs that were detected either only in the retina (see spreadsheet labelled **“Retina-specific isomiRs”**) or in the RPE/choroid samples (see spreadsheet labelled **“RPEchoroid-specific isomiRs”**). The list of isomiRs that were identified in both the retina and RPE/choroid samples are listed in the spreadsheet labelled **“Common isomiRs Retina-RPEchor”** .

N.B.: Only sequence variants (i.e. reference miRNA sequences and isomiRs) corresponding to miRNAs expressed both in the retina and RPE/choroid samples and detected across all analysed samples (i.e. 16 retina and 2 RPE/Choroid) were considered for these comparisons. The last rows of each spreadsheet show the average expression values (as reads) and average relative frequency (as %) across all variants listed therein, along with the sum of reads of the listed variants.

The spreadsheet header annotation is as follows: Sequence_ID_Retina or RPE/Choroid: Identifier of each read sequence, which indicates whether the read corresponds to the reference sequence (_REF) or to an isomiR (_isomiR); miRNA Name: miRNA to which a given isomiR is associated; Read Sequence: Sequence of the identified read; Nucl. substitution: nucleotide substitution position-nucleotide at sequence-nucleotide at precursor; additions: non-templated nucleotides added at the 3’-end; t5: nucleotides at 5’-end different from the annotated mature sequence in the miRBase (“u”: nucleotides added, “d”: nucleotides trimmed); t3: nucleotides at 3’-end different from the annotated mature sequence in the miRBase (“d”: nucleotides added, “u”: nucleotides trimmed); modified seed (Yes/No): isomiR variations resulting in a modified seed; Average frequency Retina or RPE/Choroid: average frequency as counts per million of a given variant in the samples it was detected; Relative frequency Retina or RPE/Choroid: relative frequency of each sequence variant shown as percentage of total reads for the respective miRNA; Rank Retina or RPE/Choroid: Ranking position of each sequence based on its relative abundance. The ranking was established in the list that comprised all isomiRs identified in the retina or RPE/choroid for the common miRNAs, prior to their separation into the tissue-specific or common lists.

(see separate Excel file containing three spreadsheets)