Supplementary Figures

Supplementary data include 6 Supplementary Figures (Figure S1-S6).

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.*

Figure S1

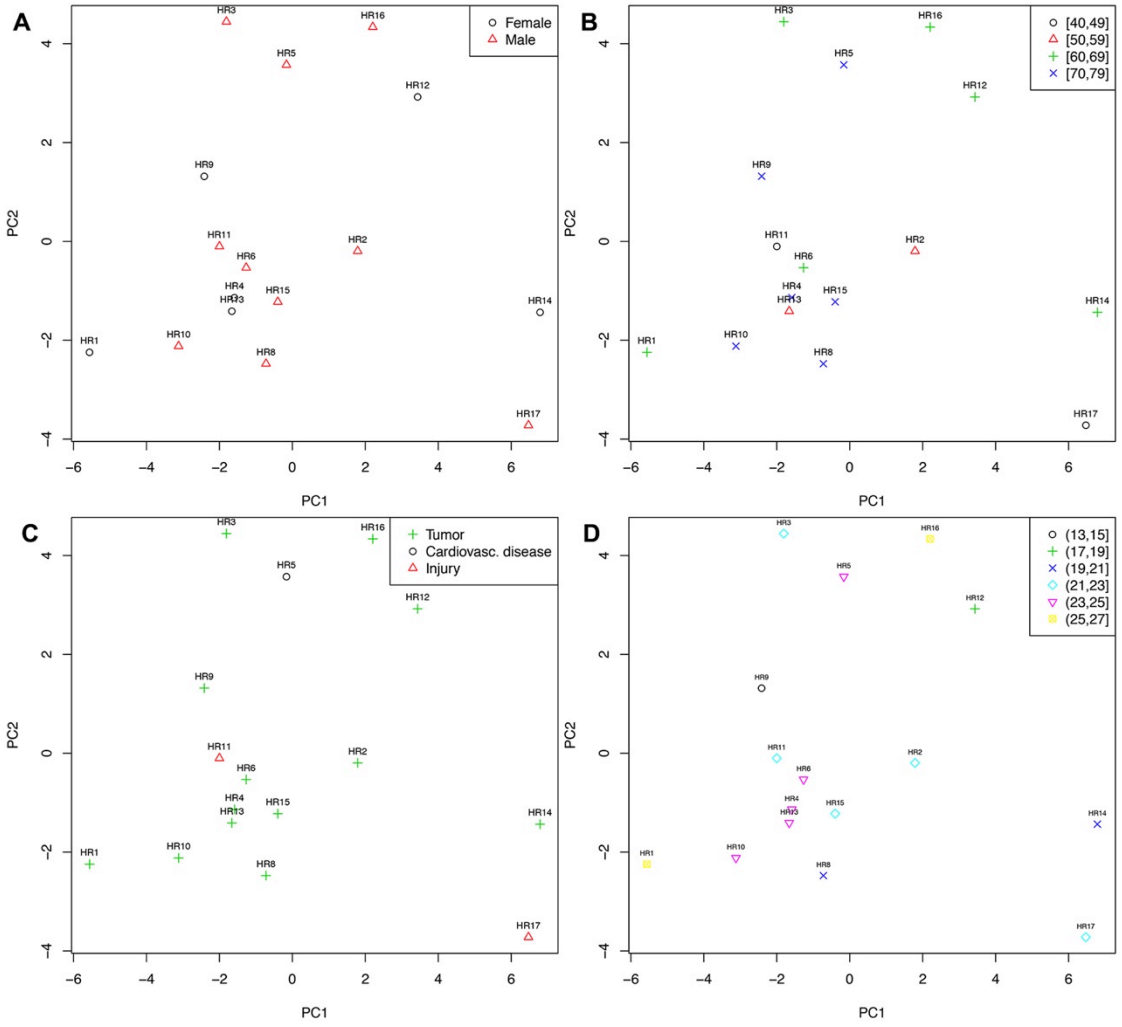

Figure S1: Two-dimensional principal component analysis (PCA) plot of the 16 retina samples based on miRNA expression profiles.

In each plot the samples are annotated based on potential sources of variation, namely gender (A), age (B), cause of death (C) and total post-mortem time (D). Symbols in B and D indicate ranges of years and hours, respectively. The contribution of PC1 and PC2 to expression data variance was 26% and 15 %, respectively.

Figure S2

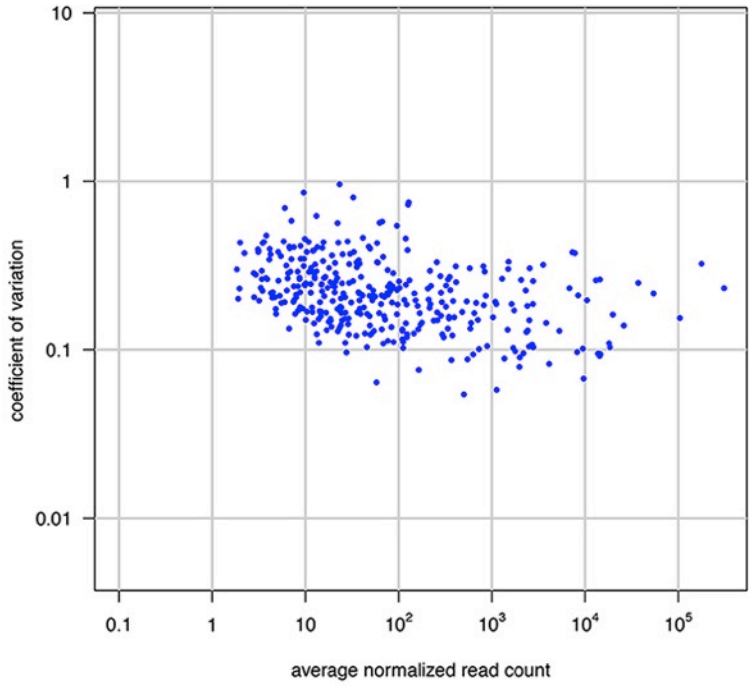

Figure S2: Scatter plot of the coefficient of variation and the average normalised read count for each miRNA (indicated by a blue dot).

Figure S3

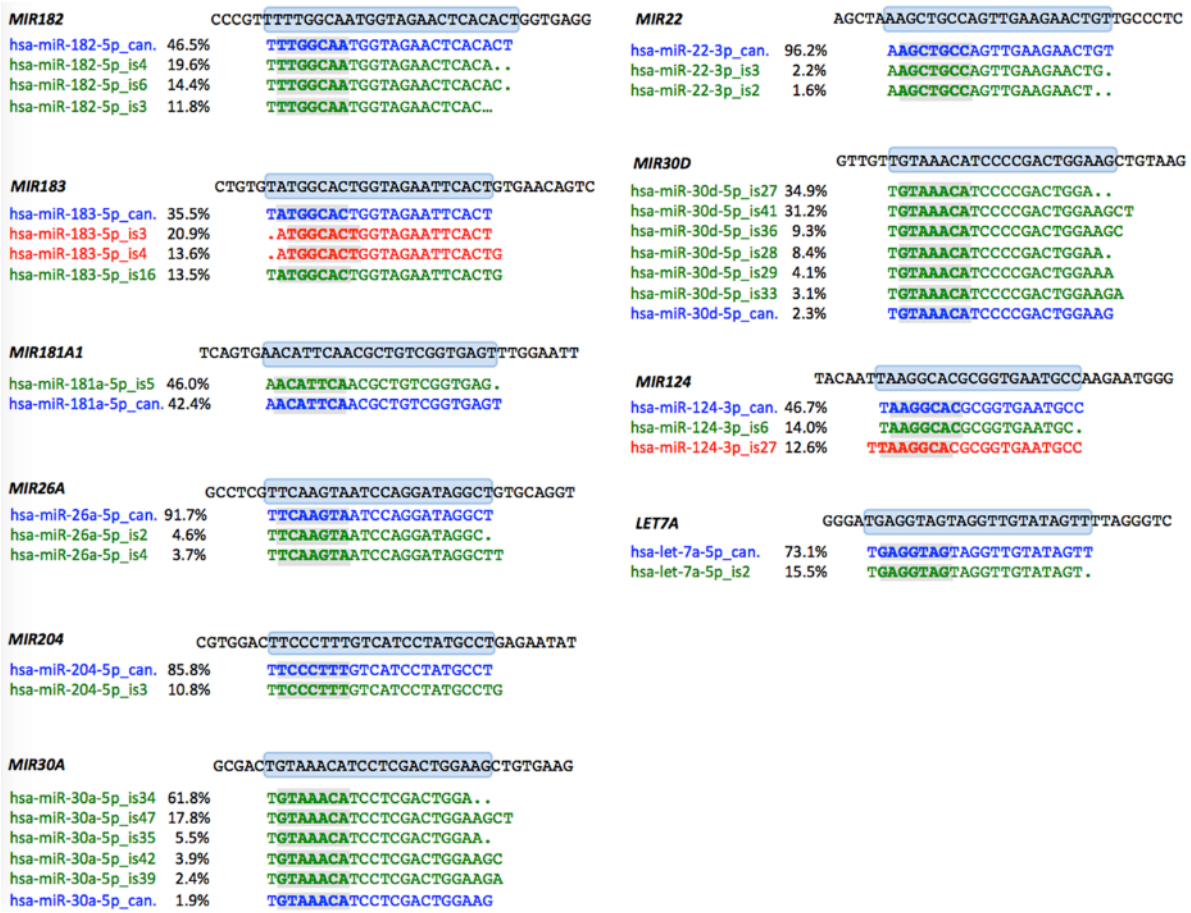

Figure S3: Alignment of the most abundant isomiR to the canonical sequence for the ten top-ranked miRNAs in the human retina.

The sequences of the isomiRs (_is) that contribute at least 10% to the total read counts of a given miRNA are aligned to the canonical sequence annotated in the miRBase (_can.; in blue). When the canonical form contributes with less than 10% to the total read counts (e.g. miR-30a-5p and miR-30d-5p), all isomiRs that are more abundant than the canonical miRNA are listed. In the cases where no isomiRs contributed to >10% (e.g. miR-26a-5p and miR-22-3p), the isomiRs that contribute with at least 1% are shown. The percentages represent the contribution of each isoform to the total read counts of a given miRNA. In green: isomiRs bearing the same seed as the canonical miRNA; in red: isomiRs with a modified seed sequence; in black: precursor sequence. The seed sequence (nucleotides 2 to 8) is shaded. The analysis was performed on isomiRs detected in all retina samples (n=16).

Figure S4

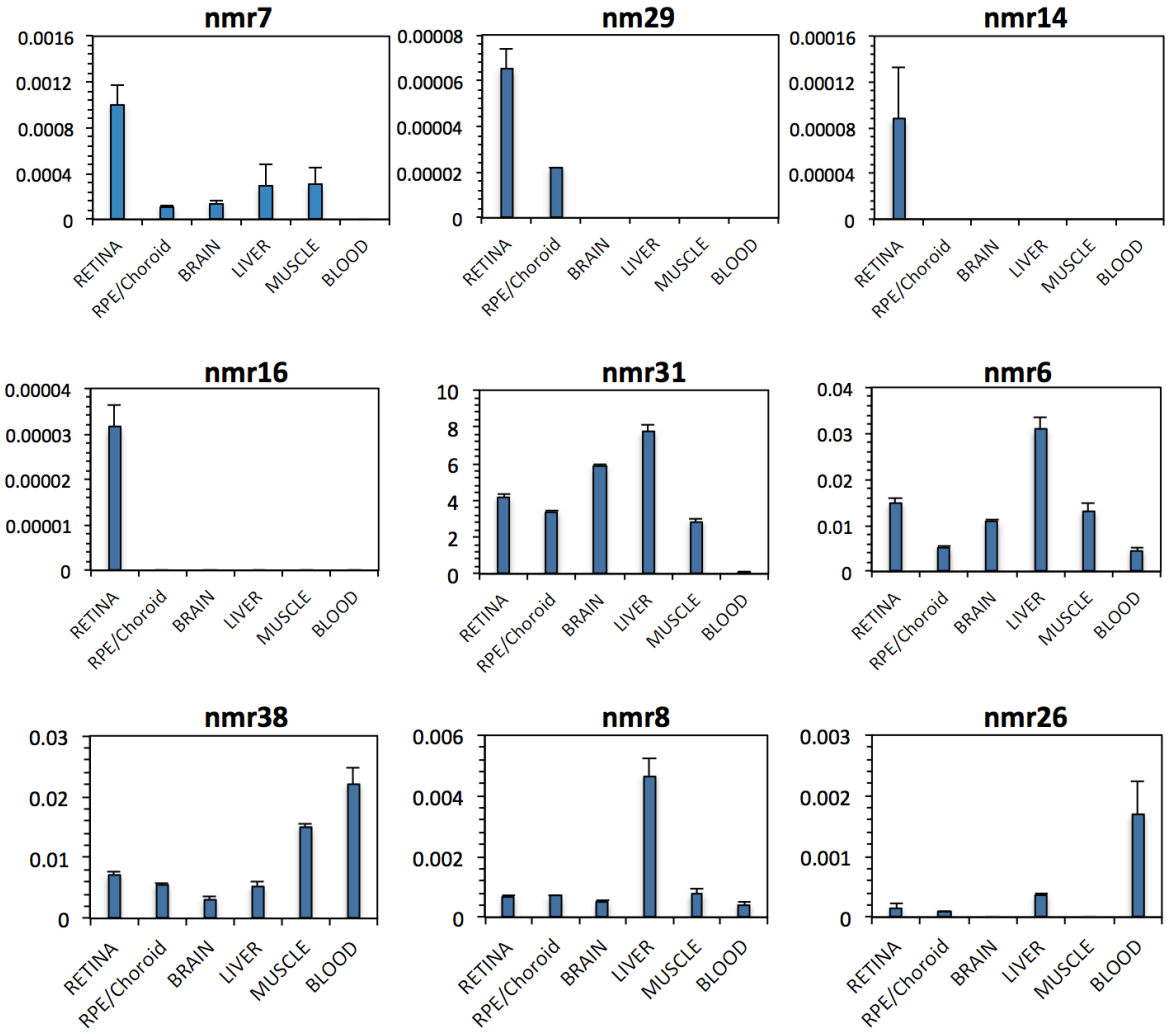

Figure S4: qRT-PCR analysis of nine putative novel miRNAs in the human retina and in other human tissues.

The expression of nine predicted novel miRNAs (see Table 2 for their genomic location) was validated by qRT-PCR using the miRCURY LNA™ microRNA PCR System on total RNA from human retina, RPE/choroid, brain, liver, muscle and blood. The ubiquitous miR-26a was used as a reference endogenous control for normalisation in the context of each tissue. Values are shown as $2^{\wedge\text{-dCT}}$. The novel miRNAs are shown in the same order as in Figure 5A.

Figure S5

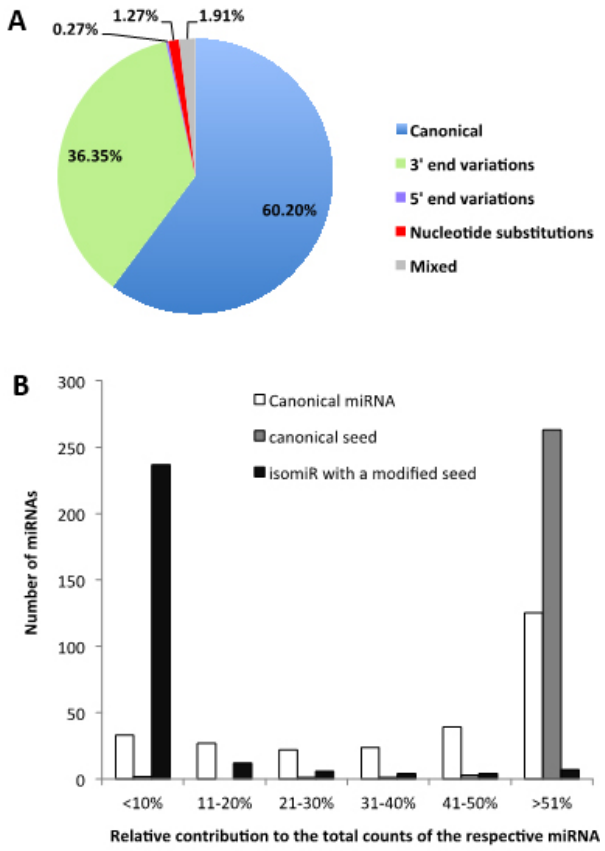

Figure S5: Characterisation of miRNA sequence variations and their contribution to miRNA expression levels in the RPE/choroid.

(A) Pie-chart illustrating the contribution of the different isomiR categories to total read counts in the RPE/choroid. (B) Contribution of canonical miRNAs (white bars) and of isomiRs with a canonical (grey bars) or modified seed (black bars) to overall miRNA expression in the human RPE/choroid. The bar reports the number of miRNAs for which the expression of each sequence class, relative to total counts, is within the five categories indicated on the x-axis. The data reported in this figure refer to the analysis of the isomiRs identified in the two human RPE/choroid samples.

Figure S6

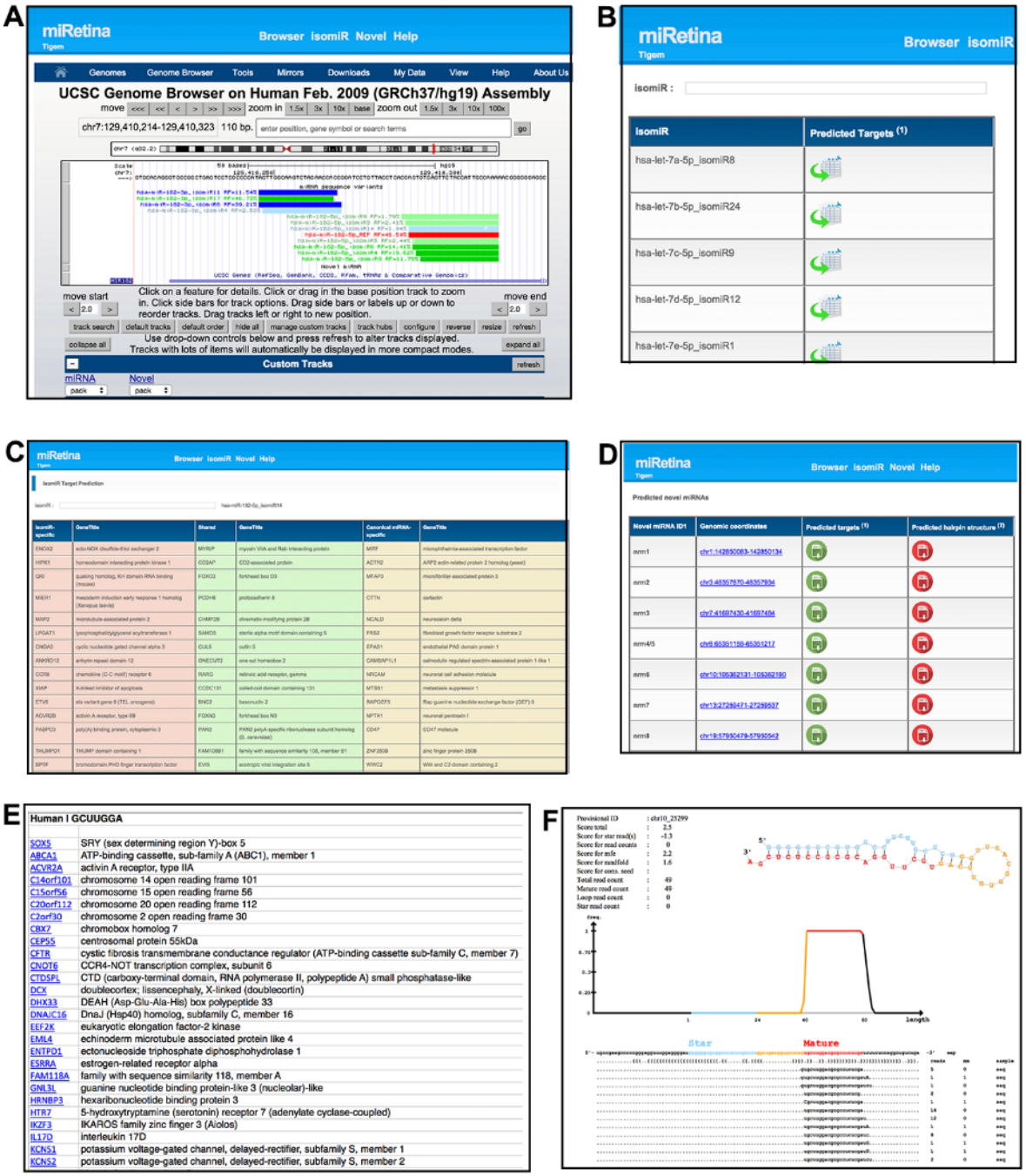

Figure S6: The miRetina database.

(A) miRetina enables queries by miRNA gene name (e.g. *MIR182* or miR-182) or genomic coordinates and provides graphical information on the genomic location by presenting an alignment of the reference miRNA in a Genome Browser window along with the respective isomiRs. Only sequence variants that were identified in all 16 individuals are depicted. The relative expression level of each miRNA variant in the retina is shown as relative frequency (RF,%) out of total miRNA reads. A colour code discriminates between the reference miRNA sequence (in red) and the isomiRs (green for the isomiRs with a canonical

seed, blue for the ones with a modified seed). Nucleotide substitutions are indicated within the isomiR field (see a C to A substitution within the blue coloured isomiR182-5p). miRNA variants that contribute with more than 10% to the total miRNA reads are labelled in darker shade (e.g. see RF of light and dark green isomiRs).

(B, C) For the isomiRs of the 100 top-ranked miRNAs that bear a modified seed and have an RF≥1%, the user can download the list of predicted targets, generated using the TargetScan custom tool. For each isomiR, the database provides information on the genes predicted to be targeted only by the isomiR, those targeted both by the isomiR and its respective reference miRNA and genes predicted to be targeted only by the reference miRNA.

(D) The 51 predicted novel miRNAs are listed in a tabular format and can be visualised in the genome browser interface of miRetina.

(E) The user can retrieve the list of predicted target genes for each novel miRNA, obtained using the TargetScan custom software, and download this information as a table with active hyperlinks to NCBI.

(F) Information on the predicted secondary structure of the hairpin and alignment of reads across the predicted precursor (as provided by the miRDeep2 software) is also downloadable.