Characterization of novel 3'untranslated regions and related polymorphisms of the gene *NPPC*, encoding for the C-type natriuretic peptide.

Ombretta Melaiu, Maria Sole Facioni, Manuela Cabiati, Raffaele Caruso, Daniela Giannessi, Stefano Landi*, Federica Gemignani*, Silvia Del Ry*

*These authors contributed equally to this work

ABSTRACT

Elevated plasmatic levels of C-type natriuretic peptide (CNP) were found in patients with chronic heart failure (CHF), but its use as sensitive and specific clinical bio-marker is still controversial. In fact, high levels of CNP were also observed in patients classified in low severity New York Heart Association (NYHA) classes. CNP is encoded by a gene poorly studied (*NPPC*, natriuretic-precursor peptide C), where the regulatory regions are not well defined and the role of single nucleotide polymorphisms (SNPs) poorly ascertained. In the present work, we focused on the characterization of the 3'untranslated region (3'UTR) of the gene, using Rapid Amplification of cDNA 3'-End (3' RACE), and we identified two novel transcript isoforms (L-3'UTR; S-3'UTR; accession number JF420840, HQ419060 respectively). Since it could be hypothesized that genetic variations could explain the observed inter-patients differences, we searched for novel SNPs, by the use of High Resolution Melting (HRM). Moreover, we predicted relevant microRNAs binding to the novel 3'UTRs that could modulate the post-transcriptional regulation of *NPPC* and affect the plasmatic levels of CNP.

Introduction

The 22-aminoacids vasodilator C-type natriuretic peptide (CNP), released by the vascular endothelium and by the heart [Barr et al., 1996; Del Ry et al., 2011a], belongs to a family that includes also the urodilatin, and the ANP (atrial), BNP (brain), and DNP (dendroaspis) natriuretic peptides [de Bold et al., 1981; Schulz-Knappe et al., 1988; Sudoh et al., 1988; Sudoh et al., 1990; Schweitz et al., 1992; Del Ry et al., 2006]. In patients with chronic heart failure (CHF), plasma levels of ANP and BNP are elevated in relation to the clinical severity (measured according to the classes issued by the New York Heart Association, NYHA) and have diagnostic and prognostic relevance [Burnett et al., 1986; Ogawa et al., 1991; Clerico et al., 2009]. On the other hand, the significance of CNP is still controversial [Kalra et al., 2001]. In fact, also CNP shows elevated levels in CHF patients in relation to the functional and clinical severity [Kalra et al., 2003; Del Ry et al., 2005; Del Ry et al., 2007; Del Ry et al., 2008; Passino et al., 2008; Kalra et al., 2010; Del Ry et al., 2011b]. However, high levels of CNP were also observed in some patients with low NYHA classes, whereas others present low levels although with a severe disease, preventing its use as sensitive and specific clinical bio-marker [Del Ry et al., 2005].

It could be hypothesized that genetic variations could explain the observed inter-patients differences. In fact, CNP is encoded by a gene poorly studied (*NPPC*, natriuretic-precursor peptide C) where the regulatory regions were not well defined and the role of single nucleotide polymorphisms (SNPs) was never ascertained. In a single previous work Ono *et al.*, studied four SNPs within *NPPC* and one of them, the variant G2628A (rs5268), hypothesized to be located within the putative 3'UTR, was found associated with essential hypertension [Ono et al., 2002]. However, it should be noticed that the 3'UTR described by the authors showed a strange inconsistency when compared to the orthologous 3'UTRs of other mammals. The polymorphism rs5268 was considered to belong to the 3'UTR, simply because it was located 82 base pairs downstream of the stop codon. Experiments aimed to characterize the region were not performed and there was not any evidence that the SNP was actually located within it. Thus, in the present work, we focused on the characterization of the 3'UTR and we identified two novel transcript isoforms. Thus, we could search for novel polymorphisms and relevant microRNAs (miRNAs) that could modulate the post-transcriptional regulation of *NPPC* and affect the plasmatic levels of CNP.

Materials and Methods

Sample collection

Human cardiac tissues.

Patients undergone to left ventricular assist device (LVAD) implantation and at the end-stage heart failure (HF) (NYHA class III and IV; age: 57±11 yrs; LVEF%<20), provided biopsies necessary for the study. The cardiac biopsies (n=5) were collected from left ventricle (LV) at the time of the heart transplantation, thanks to the collaboration with the Institute of Clinical Physiology (IFC)-CNR, Niguarda Ca' Granda, Milan. Immediately after collection, samples were frozen in liquid nitrogen and stored -80°C until sample preparation. The study protocol was conform to the principles outlined in the Declaration of Helsinki and was approved by the local Ethics Committee of the Niguarda Ca' Granda Hospital. All subjects gave written informed consent to participate in the study.

Plasma samples.

Plasma samples were collected from 63 patients (mean age 61±1 years) with a diagnosis of chronic heart failure (CHF) provided by Institute of Clinical Physiology, CNR-Milan. Blood samples (10 ml), collected in ice-chilled disposable polypropylene tubes containing aprotinin, 500 KIU/ml, and EDTA, 1 mg/ml, were rapidly separated by centrifugation for 15 min at 4°C, and stored frozen at –20°C in ml aliquots in polypropylene tubes until assay, performed within 1 month from sampling.

Blood samples, used as controls, were obtained from 63 volunteers without heart disease and were provided by a routine biochemical analysis laboratory of the region. Plasma and blood samples were matched for the age of the donors. All subjects were informed and gave written consent to participate to the study and to allow their biological samples to be genetically analysed, according to the Helsinki declaration.

Neuroblastoma cells. The human neuroblastoma SKNBE cells (kindly donated by the National Research Council, Genova, Italy) were grown to confluence in RPMI 1640 medium (Lonza, Switzerland), containing 10% heat-inactivated fetal calf serum (Sigma Aldrich Corp. St Louis, MO, USA), 1% glutamine (Lonza, Switzerland), 1% PEN-STREP (Lonza, Switzerland), at 37°C in a 5% CO_2 humidified atmosphere.

Nucleic acid extraction and quantification

Total RNA was extracted by acid guanidinium thiocyanate-phenol-chloroform method from tissue samples obtained from human cardiac tissue with Rneasy Midi kit (Qiagen MI, Italy). Briefly, the tissue was homogenized with a Mixer Mill MM300 (Qiagen MI, Italy). In order to remove

contractile proteins, connective tissue and collagen, which can interfere with the procedure, the protocol was modified to include proteinase K digestion. Samples were lysed in a guanidine-isothiocynate-containing lysis buffer. After dilution of the lysate, the samples were treated with proteinase K. Debris was pelleted by centrifugation. Ethanol was added to the cleared lysate and RNA was bound to the Rneasy membrane. Contaminants were washed away, and total RNA was eluted in Rnase-free water. The presence of proteinase K and the tissue homogenization by Mixer Mill MM300 allowed us a high yield in RNA extraction. For the extraction of total RNA from the human neuroblastoma cell line, cells were washed twice with DPBS-A, treated with trypsin-EDTA (Invitrogen MI, Italy), and counted in a haemocytometer. Three millions cells were pelleted by centrifugation at 200 g, then the QIAgen Rneasy Mini Kit (Qiagen MI, Italy) was used according to the manufacturer's protocol for all large scale samples. RNA concentration was determined spectrophotometrically (Beckman DU 640) at 260 nm. The ratio of readings at 260 nm and 280 nm (A_{260}/A_{280}) provided an estimate of the purity of RNA with respect to contaminants that absorb in the UV spectrum, such as proteins and DNA. The integrity and purity of total RNA was also detected by electrophoresis of samples on ethidium bromide agarose gels. Only samples that showed clear and distinct 28S and 18S ribosomal RNA bands and had spectrophotometric OD 260/280 ratios of 1.9-2.1 were used. A known amount of total RNA (Ambion Inc. Austin, USA) was used as marker. DNA was extracted using the commercially available kit QIAmp DNA Mini and Blood Mini kit (Qiagen MI, Italy). Quantification of all experimental DNA samples was assessed using Qubit® Fluorometer (Invitrogen, MI, Italy). Briefly, total gDNA was extracted using the recommended protocol. Proteinase K (20 μl) was added to the sample and the mixture was further incubated at 56°C for 1 h. Lysis buffer (200 μl) was added to the sample, mixed thoroughly by vortexing, and incubated at 56°C for 10 min. Ethanol (95–100%, 200 μl) was then added to the sample, mixed thoroughly by vortexing, and the whole mixture was applied to a QiAmp spin column with a collection tube. The mixture was allowed to bind to the column by centrifugation at 6000 g for 1 min. The QiAmp spin column was washed with buffers followed by centrifugation at 6000 g for 1 min to remove excess ethanol. The spin column-bound gDNA was eluted with the elution buffer (50 μl) and was incubated at room temperature for 5 min, and then was centrifugated at 6000 g for 1 min. DNA was incubated at 65°C for 1 hour.

DNA sequences alignment and miRNA target prediction

The alignment between the available human consensus sequence (RefSeq) and its orthologous of mouse, rat, cow, pig, goat, and guinea pig was performed with UCSC Genome Browser (URL: http://genome.ucsc.edu/). The sequences obtained from our experiments were aligned against the

human genome as well as against mammals' genomes using Nucleotide BLAST (URL: http://blast.ncbi.nlm.nih.gov/Blast.cgi). Putative miRNA-binding sites within the 3'UTRs of *NPPC* were identified by means of miRanda (http://www.microrna.org). This algorithm evaluates the sequence complementarity between the mature miRNA and the target site, locates the binding position on the target, and calculates the binding free energy (expressed as ΔG, Gibbs free energy, in kcal/mol). The miRanda software was ran locally, with a threshold of $\Delta G \leq -15.00$ kcal/mol, using the two identified isoforms as target sequences and the most updated list of human mature miRNAs as query. MiRNAs were downloaded from miRBase a public database, available at the URL: http://www.mirbase.org/.

Rapid Amplification of cDNA 3'-End (3'-RACE)

Total RNA was reverse-transcribed with the commercially available kit Cloned AMV reverse transcriptase (Cloned AMV First-Strand cDNA Synthesis Kit, Invitrogen, MI, Italy), following the protocol suggested by the produced. For the identification of the 3'end, a rapid amplification of cDNA ends (3'-RACE) was performed. Briefly, the cDNA was synthesized from 4 µg of total RNA extracted from human cardiac tissue and 5 µg of RNA extracted from human neuroblastoma cells, using an universal primer (RACE-1) containing an anchor sequence, in a 20 µl reaction. Reaction protocol started with 42°C for 1 h 30 min and followed by 85°C for 5 min. cDNA was precipitated with glycogen, sodium acetate 3M, pH 5.2, and absolute ethanol. cDNA was then amplified using a gradient PCR as follows: an initial denaturation step (8 min at 95°C) followed by 35 amplification cycles (30 sec at 95°C, 30 sec at 53°C-60°C, 1 min and 30 sec at 72°C). The sequences of the primers used are reported in table 1, where the forward primer (RT-F) was gene-specific, whereas the reverse primer (RACE-2) contained a sequence complementary to the anchor sequence of the RACE-1 primer. Both were used for the first amplification, followed by a second nested PCR reaction with the gene-specific forward primer (RT-F') and the universal reverse primer (RACE-3), complementary to the anchor sequence of RACE-2, using the same amplification protocol as before. PCRs were performed in a volume of 25 µL per reaction. Reaction mixture included 2 µL of template cDNA, 2 µL of each primer (Bio-Fab Research RO, Italy), 3 µL of $MgCl_2$, 2.5 µL of Buffer, 0.25 µL of dNTP, 0.25 µL of Taq polimerase and sterile H_2O. The PCR products were separated on a 2% agarose gel. Bands were excised and purified with a QIAquick gel extraction kit (Qiagen MI, Italy) and sequenced with the standard Sanger reaction on an Applied Biosystem ABI PRISM 3730 XL automated sequencer.

PCR assay

To identify SNPs within the two identified 3'UTRs, specific primer couples SF/SR and LF/LR were designed within them, to allow PCR amplification of 386- or 846- base pairs long fragments,

respectively. Samples were amplified with a touch-down protocol. Touch-down amplification was performed with an initial step of 95°C for 8 min, followed by 20 cycles of 95°C for 30 sec, annealing temperatures starting at 68°C for 30 sec (decreasing by 1 C/cycle), and 72°C for 1 min and 15 sec for extension. This step was followed by 30 cycles of 95° C for 30 sec, 51°C for 30 sec, 72° C for 1 min and 15 sec, and a final extension at 72° C for 8 min. PCRs were performed in a volume of 20 μL per reaction. Reaction mixture included 1 μL of template DNA, 0.1 μL of each primer (Bio-Fab Research RO, Italy), 2.4 μL of MgCl$_2$ (Solis BioDyne Tartu, Estonia), 2 μL of Buffer (Solis BioDyne Tartu, Estonia), 4 μL of Solution S (Solis BioDyne Tartu, Estonia), 0.2 μL of dNTP (Solis BioDyne Tartu, Estonia), 0.3 μL of Taq polymerase (Solis BioDyne Tartu, Estonia), and sterile H$_2$O. The PCR products were separated on a 2% agarose gel. Bands of the expected sizes were purified with Microcon Centrifual Filter Device (Millipore) and sequenced.

High Resolution Melting (HRM)

To analyze exon 2 of *NPPC* gene, DNA extracted from plasma and blood samples was amplified using HRM-F and HRM-R primers reported in table 1, producing an amplicon of 162 bp. The PCR for HRM was carried out with the Type-it® HRM PCR kit (Qiagen MI, Italy) using the LightCycler® 480 (Roche) with the following conditions: 95°C for 5 min followed by 50 cycles of 95°C for 10 sec, 62°C for 30 sec and 72°C for 10 sec. Afterwards, amplicons were heated from 65°C to 95°C for 1 sec and cooled to 40 °C for 1 sec. Finally, the melting curves were generated by measuring the fluorescence emitted by the fluorochrome EvaGreen® during the increase of the temperature, from 65°C to 95°C, with an increment speed of 0.02°C/sec. Results were analyzed using LightCycler® 480 Gene Scanning software to detect sequence variations in PCR amplicons.

Results

When the human *NPPC* gene (RefSeq: NM_024409.2) was aligned together with the mammal orthologous (as listed in UCSC genome browser), its 3'UTR differed significantly. In fact, mouse, rat, and cow showed a splice junction, joining the portion of 20 nucleotides downstream of the stop codon with a genomic region distant 2845 nucleotides, whereas the human 3'UTR was reported as a simple short sequence of 40 nucleotides downstream of the stop codon. This corresponded, essentially, to the clones BC105065 and BC105067, as described by Strausberg *et al.* [Strausberg et al., 2002].

However, GenBank (URL: http://www.ncbi.nlm.nih.gov/gquery/?term=genbank) revealed also the presence of spliced ESTs for the human gene (Genbank ID DR003533). Thus, we defined more precisely the 3'UTR with the use of the 3'-RACE. The neuroblastoma cell line expressed a 3'UTR where 20 nucleotides downstream of the stop codon were joined (by splicing) with a region 1387

nucleotides far away. The figure 1 depicts the electropherogram showing the splice junction of this new isoform, never described before, where the 3'UTR is 356 nucleotides-length that we denominated "short" (S). When the human cardiac tissue was analysed, we found an alternative 3'UTR, where a splice junction joined a portion of 20 nucleotides downstream of the stop codon with a region 3089 nucleotides far away. The figure 2 shows the splice junction and depicts that this new isoform (with the 3'UTR of 480 nucleotides-long) resembles closely the conserved mammal orthologous. This isoform matches closely also with some previously described spliced ESTs and it was denominated as "long" (L). The S- and L- isoforms were submitted to GenBank. The former is annotated as "*Homo sapiens natriuretic peptide precursor type C mRNA, partial cds and 3'UTR, 318 bp*", with the accession number HQ419060, whereas the latter as "*Homo sapiens natriuretic peptide precursor type C mRNA, partial cds, 427 bp*" with the accession number JF420840. The figure 3 shows the UCSC Genome Browser window with the novel isoforms, the non-human genes, the spliced ESTs and the novel and corrected RefSeq. GenBank automatically assigned the protein sequences IDs ADQ54381.1 and ADZ36320.1 for the S- and L-, respectively. It should be noticed that the coding sequence does not differ to the previously established RefSeq (NM_024409.2). We confirmed the presence of the S- and L- isoforms with specific PCRs, revealing that, actually, both the isoforms are expressed in the cardiac tissue, with the L- apparently more expressed than the S-one, according to the band intensities obtained on the electrophoresis gel.

It should be also noticed that the polymorphism rs5268 did not fall within the transcribed mRNA, since it is located within the intronic sequence. However, we cannot exclude that it was associated with hypertension because of effects on the regulation of the splicing, or of the linkage disequilibrium (LD) with other SNPs nearby. Unfortunately, only two SNPs (rs3107179 and rs5262, the former located approximately 910 base pairs at the 5' region of the gene and the latter 1820 bp downstream of the stop codon), were genotyped in the contest of the HapMap project (URL: http://hapmap.ncbi.nlm.nih.gov/), thus preventing to define a precise LD map of the region. Moreover, using the available online databases dbSNP (URL: http://www.ncbi.nlm.nih.gov/projects/SNP/) and GenBank we could not predict whether the rs5268 was linked with other relevant SNPs in the gene region. In particular, we observed a poor characterization of the SNPs within the locus. Thus, we attempted to identify novel SNPs, in particular within the novel 3'UTRs. To identify any polymorphism, we used gene-specific primers for the S- and L- isoforms and the PCR products (386 bp and 846 bp, respectively) from eight independent chromosomes were sequenced. Although, the number of chromosomes was limited, we did not find any novel polymorphism and this finding agrees with the lack of genetic variations within dbSNP. Thus, we can conclude that among Caucasians, both the S- and L- isoforms are

devoid of common polymorphisms, whereas we cannot exclude the presence of rare variants. SNPs were not found either in controls or in patients. Following the limited information from HapMap project, it seemed that the two SNPs encompassing the locus were in LD each other (D'=0.979; r^2=0.34). Thus, we hypothesized that most of the SNPs in the gene region could be located within one LD-island. Then, we focused our attention to the central portion of the gene (i.e. the second exon), that putatively could bear SNPs in LD with the rs5268. In particular, the second exon encodes for the critical portion of the protein, and nine single nucleotide variations were described in dbSNP. Unfortunately, these variations were not validated experimentally and their allele frequencies were not available. It is predicted that rs5266, rs76093905, and rs138975043 do not encode for aminoacid substitutions, whereas the others (rs5267, rs79480591, rs13305994, rs80265630, rs13305993, rs80022541) do. No SNPs were described within the exon 1 in dbSNP. We focused on SNPs rs5267, rs79480591, rs13305994, and rs80265630, very close each others and we have undertaken a mutation screening in order to characterize whether these SNPs could have a meaningful frequency among Caucasians. We screened 126 independent chromosomes from CHF patients and 126 from healthy controls, with the use of HRM. A total of six samples showed HRM peaks slightly shifted as compared to the other samples and the related amplicons were sequenced. The results showed a complete lack of genetic variations among our series of samples.

After having excluded any novel polymorphisms within the 3'UTRs and within the second exon of the gene, we wondered whether the novel 3'UTRs could bear miRNA binding sites, important for the post-transcriptional regulation of *NPPC*. Thus, we predicted the putative miRNAs targeting the S- and L- 3'UTRs and we found 750 and 1024 miRNAs, respectively (see supplementary tables 1 and 2). When considering only those showing at least 10 contiguous bases perfectly matching the target, 41 miRNAs were predicted to bind the L- and 11 the S-form. When the human forms were aligned together with its mammal orthologous (rhesus, mouse, dog, and cow), the S- 3'UTR did not show any conserved sequence, whereas we found that L-form encompassing between chr2:232786879 to chr2:232786909 (fig. 3) was highly conserved showing >97% of homology. The miRNAs predicted to bind the conserved sequence are reported in table 2.

Discussion

The main findings of the present study are the identification of two novel isoforms of the 3'UTRs of the gene *NPPC*: the S-isoform (318 bp), firstly identified in a human neuroblastoma cell line, and the L-isoform (427 bp), identified in the human heart tissue from CHF patients. It should be noticed that the S-3'UTR was never observed among spliced human ESTs, nor among the orthologous

counterpart, and that this isoform shows low conservation among mammals. Since it was detected in a cancer cell line and, at low level, in heart tissues from patients undergone to LVAD implants, it could be speculated that the S-form is the result of a sort of "aberrant" splicing, consequence of cellular stress or disfunction. Alternatively, it cannot be excluded that the S-form has important regulatory functions peculiar of the human normal physiology of cardiomyocytes and neurons. In any case, further studies are warranted in order to better understand the meaning of this quite unconventional spliced mRNA. The L-3'UTR has a structure very similar to other mammalian species and we could observe that the rs5268, previously attributed to the 3'UTR of the gene, actually falls within the intervening intronic sequence. We did not identify any (novel) SNP within the L- or the S- 3'UTRs, as well as within the second exon of the gene, perhaps because the existence of an important selective pressure on these regions. Since the work from Ono *et al.* [Ono et al., 2002], was carried out in Japanese, it cannot be excluded that the association with the hypertension is due to polymorphisms private of the Asiatic population, in LD with rs5268. However, this explanation is quite unlikely. In fact, differences between Caucasians and Asians are quite limited and, according to dbSNP, there is a lack of genetic variants in the considered regions also among Asians. If the hypertension were really associated with the rs5268, the mechanism could be linked to a deregulated splicing of the last novel intron, or to a LD with other SNPs located at the 5' end of the gene. We suggest also that the inter-patients differences in plasmatic CNP levels are not due to SNPs within the 3'UTR or the coding sequence of the *NPPC* gene. The role of SNPs at the 5' end of the gene needs further investigations.

In order to better characterize the novel 3'UTRs, we pointed our attention on possible target sites for miRNAs. There is increasing evidence that miRNAs have distinct expression profiles and play crucial roles in various physiological and pathological processes. In this regard, it could be interesting to analyze the role of miRNAs in the post-transcriptional regulation of *NPPC*. Currently, information about the possible role of miRNA regulation in relation to CNP is not available. We found, following a literature search, only two studies on cardiomyocyte-specific miRNAs (miR-133a and miR-208) and the levels of ANP and BNP. Liu *et al.* showed that the levels of miR-133a are associated with the activity of ANP and BNP, in cardiovascular disease [Liu et al., 2011], whereas van Rooij *et al.* discovered that the BNP is up-regulated in the heart of miR-208 knockout mice [van Rooij et al., 2007]. Following our search with appropriate algorithms, we found that both the S- and L-forms are predicted to be target for a large number of miRNAs. Among them, we found the cardiomyocyte-specific miRNAs miR-133a and miR-208 (binding the L-form) [Babiarz et al., 2012]. Moreover, we detected also several miRNAs actively involved in growth, development, function, failure, and stress response of the heart [Ikeda and Pu, 2010; van Rooij,

2011], such as miR-21 [Thum et al., 2008], miR-24 [Sun et al., 2004], miR-33b [Najafi-Shoushtari et al., 2010], miR-125b [van Rooij et al., 2006], miR-133b [Sun et al., 2004], miR-143 [Esau et al., 2004; Boettger et al., 2009], miR-145 [Boettger et al., 2009], miR-150 [Zhou et al., 2007], miR-195 [Sun et al., 2004; van Rooij et al., 2006], miR-212 [Ucar et al., 2010], miR-296 [Sun et al., 2004].

Interestingly, the deletion of miR-133a1 and miR-133a2 caused lethal ventricular-septal defects, dilated cardiomyopathy, and heart failure in double-mutant mice [Liu et al., 2008]. Moreover, miR-208a is encoded by an intron of the *alphaMHC* gene, encoding for a major cardiac contractile protein involved in cardiomyocyte hypertrophy [van Rooij et al., 2007]. MiR-208a is also co-expressed with the *MYH6* gene whitin the adult heart, regulating the expression of two slow myosins [van Rooij et al., 2009]. On the other hand, miR-21 regulates the ERK-MAP kinase signaling pathway in cardiac fibroblasts, which affects global cardiac structure and function [Thum et al., 2008]. Finally, miR-125b, miR-24, miR-195 were found up-regulated in 2 independent mouse models of cardiac hypertrophy and specific analyses showed an increased expression of these miRNAs in idiopathic end-stage failing human hearts [van Rooij et al., 2006].

It should be stressed that the most conserved sequence within the L-form is predicted to bind to the miRNAs listed in table 2. Interestingly, these miRNAs show a high degree of complementarity to the target (at least 10 contiguous base pairs are perfectly matching the target), with a threshold of $\Delta G \leq -15.00$ kcal/mol. Thus, it could be hypothesized that a large number of them is crucial for the binding to the gene sequence. At the present time, information about most of the listed miRNAs is not available, in fact many of them have been recently identified. However, among them, we noticed the presence of let-7, one of the best characterized miRNAs, firstly identified in *C. elegans* as a regulator of development and cellular proliferation [Lagos-Quintana et al., 2001; Lee and Dutta, 2007]. In a future, the knowledge of the role played by the genetic background in the regulation of CNP, as well as of other markers, could be of extreme importance also for the clinical practice. In fact, in order to better distinguish between the physiological and the pathological levels of specific bio-markers, personalized thresholds will be defined with the help of the genetic information. More studies are warranted in order to better understand the genetic bases explaining the regulation of *NPPC*.

Figure 1: A) Electropherogram showing the splice junction of 3'UTR "short"; B) FASTA sequence 356 nucleotides long. In bold the stop codon, underlined the splice junction.

A

B

NNGNNGNGGGCCATCTCAGGGCGACCGGTCGCGACTGCTCCGGGACCTGCGCGTGGACACCAAGTCGCGGGCAGC
GTGGGCTCGCCTTCTGCAAGAGCACCCCAACGCGCGCAAATACAAAGGAGCCAACAAGAAGGGCTTGTCCAAGGG
CTGCTTCGGCCTCAAGCTGGACCGAATCGGCTCCATGAGCGGCCTGGGATGT**TAG**TGCGGCGCCCCCTGGCGGC<u>G</u>
<u>G</u>GAGAAGAATGATTCTGACACTTGGGGACCAGCCTTCAGTAGCTACCCTTGGAATGCCTTTGCTCTCTTCTCTCC
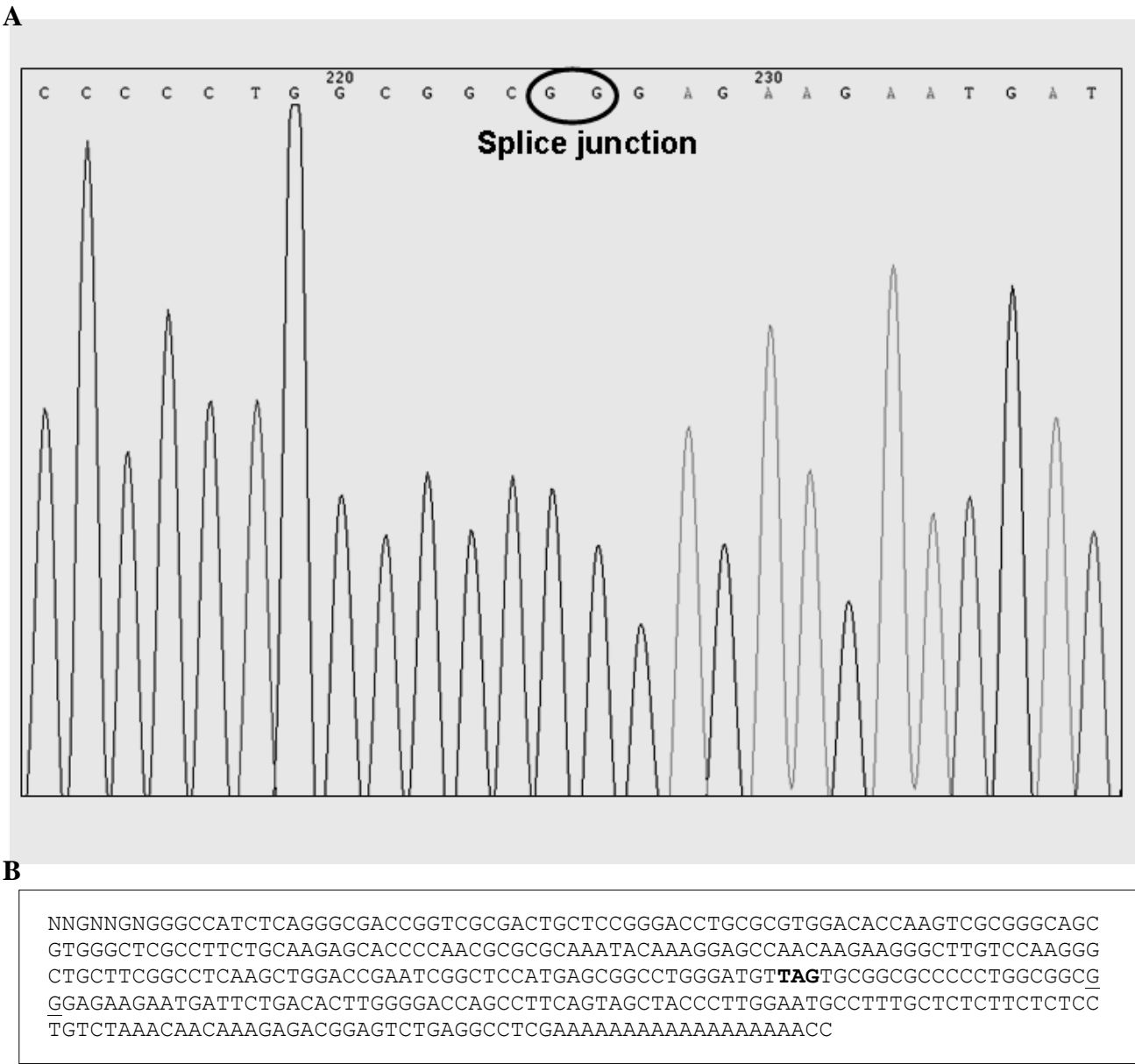TGTCTAAACAACAAAGAGACGGAGTCTGAGGCCTCGAAAAAAAAAAAAAAAAAAACC

Figure 2: A) Electropherogram showing the splice junction of 3'UTR "long"; B) FASTA sequence 480 nucleotides long. In bold the stop codon, underlined the splice junction.
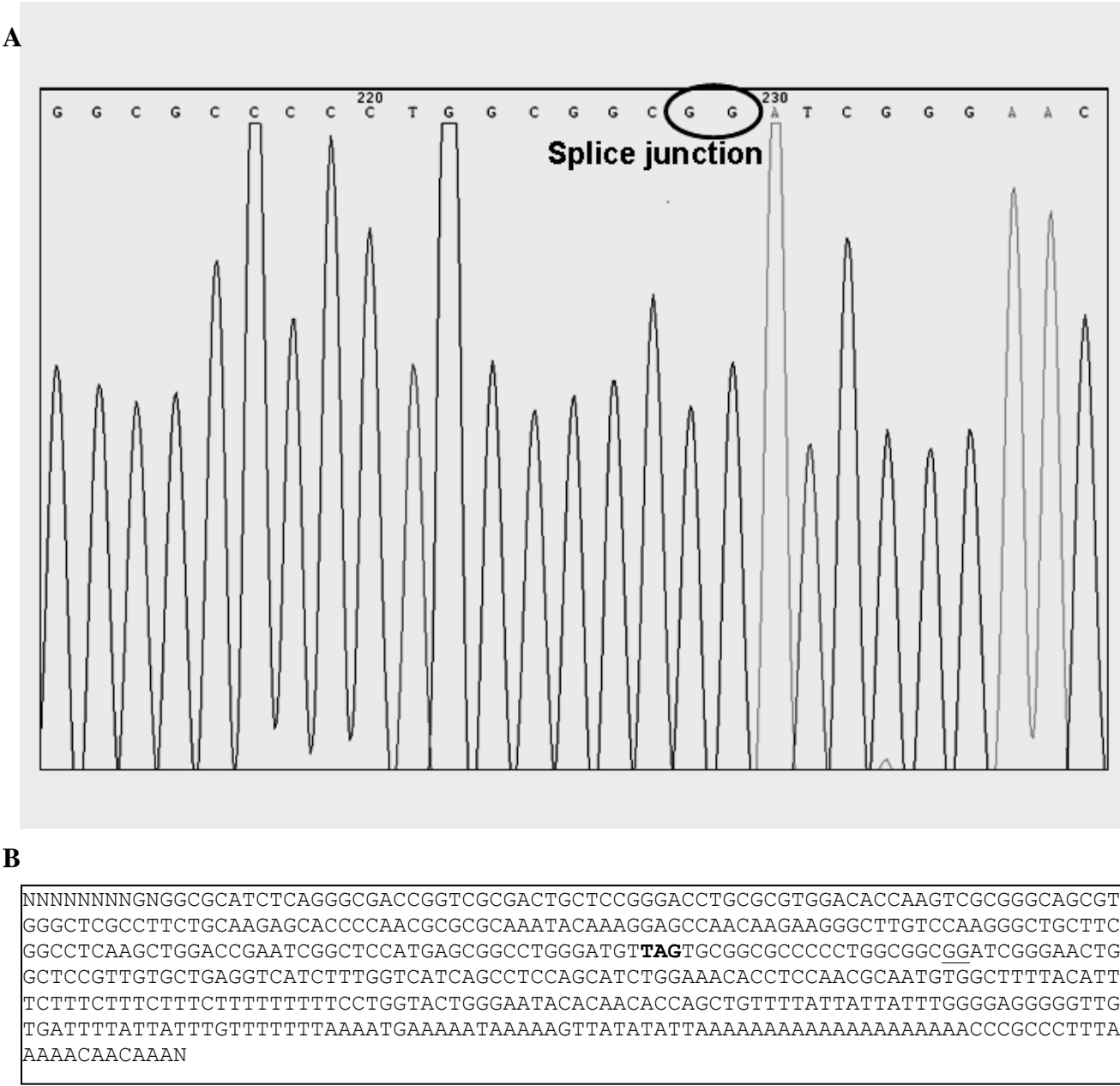
A

B

```
NNNNNNNNNGNGGCGCATCTCAGGGCGACCGGTCGCGACTGCTCCGGGACCTGCGCGTGGACACCAAGTCGCGGGCAGCGT
GGGCTCGCCTTCTGCAAGAGCACCCCAACGCGCGCAAATACAAAGGAGCCAACAAGAAGGGCTTGTCCAAGGGCTGCTTC
GGCCTCAAGCTGGACCGAATCGGCTCCATGAGCGGCCTGGGATGTTAGTGCGGCGCCCCCTGGCGGCGGATCGGGAACTG
GCTCCGTTGTGCTGAGGTCATCTTTGGTCATCAGCCTCCAGCATCTGGAAACACCTCCAACGCAATGTGGCTTTTACATT
TCTTTCTTTCTTTCTTTTTTTTCCTGGTACTGGGAATACACAACACCAGCTGTTTTATTATTATTTGGGGAGGGGGTTG
TGATTTTATTATTTGTTTTTTTAAAATGAAAAATAAAAGTTATATATTAAAAAAAAAAAAAAAAAAAAACCCGCCCTTTA
AAAACAACAAAN
```

Figure 3: A) *NPPC* **gene structure from Genome Browser. B) The conservation score for the L-3'UTR. C) The sequence of the most conserved region within the L-3'UTR from chr2:232786879 to chr2:232786909.**
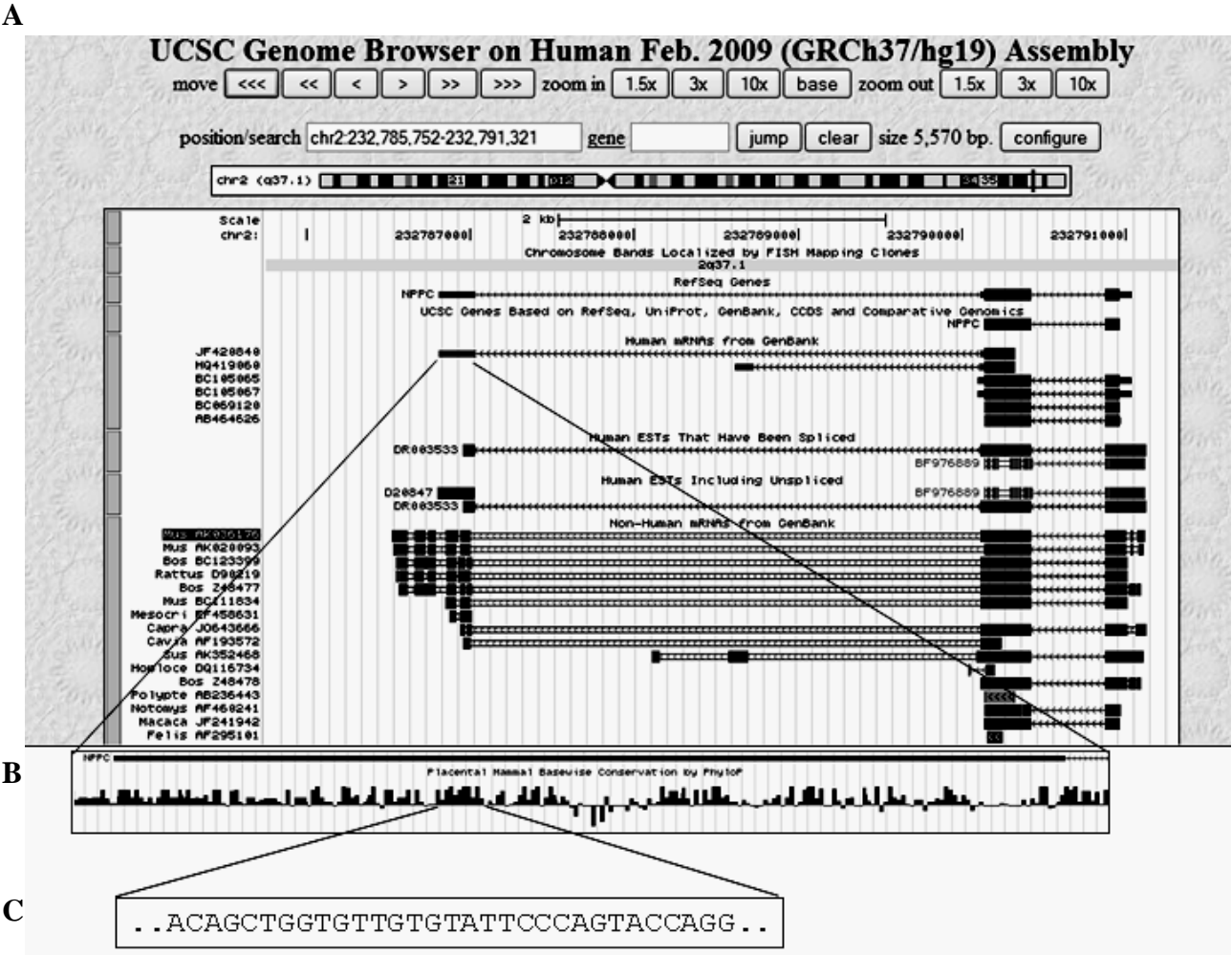
A

B

C

`..ACAGCTGGTGTTGTGTATTCCCAGTACCAGG..`

Table 1: primer sequences for PCR reactions.

Primer Id	Sequence
RACE-1	5'-CTCGCGAGCGCGTTTAAACGCGCACGCGTTTTTTTTTTTTTTTTTTVN-3'
RACE-2	5'-GCTCGCGAGCGCGTTTAAAC-3'
RACE-3	5'-GCGTTTAAACGCGCACGCGT- 3'
RT-LF	5'-CGCCGCCGAAGGTCCCG-3'
RT-LF'	5'-GTCAGAAGAAGGGCGACAAG-3'
SR	5'-ACTTGGGACGAAACTTACAC-3'
LF	5'-CCGAACAGACTCGTGGAAAT-3'
LR	5'-CCAGGTACGAACCAATGGAG-3'
HRM-F	5'-TGTTGGCTCCTTTGTATTTGC-3'
HRM-R	5'-GTCAGAAGAAGGGCGACAAG-3'

Table 2: List of predicted miRNAs targeting the highly conserved sequence within L-3'UTR depicted in figure 3 and their corresponding ΔG binding free energy.

	1st binding site			2nd binding site			3rd binding site				
	pairing [1]			pairing [1]			pairing [1]				
miRNA ID	ΔG (kcal/mol)	from	to	ΔG (kcal/mol)	from	to	ΔG (kcal/mol)	from	to	Perfect pairing [2]	Total binding free energy (kcal/mol)
hsa-let-7a-2-3p	-20.08	**159**	**175**	-17.01	4	28				11	-37.09
hsa-let-7e-3p	-19.20	**159**	**176**							10	-19.20
hsa-miR-26b-5p	-15.18	**140**	**162**							10	-15.18
hsa-miR-103b	-16.94	**149**	**175**	-18.16	11	39				10	-35.10
hsa-miR-1236	-21.00	**142**	**169**							10	-21.00
hsa-miR-130b-5p	-16.12	**151**	**170**							11	-16.12
hsa-miR-488-5p	-17.53	2	29	-15.78	**140**	**162**				11	-33.31
hsa-miR-2114-5p	-21.76	**152**	**172**							11	-21.76
hsa-miR-2355-5p	-17.18	**143**	**164**	-18.99	10	33				10	-36.17
hsa-miR-4465	-16.42	**138**	**162**	-16.01	107	127	-18.80	29	60	10	-51.23
hsa-miR-4659b-3p	-16.45	**140**	**168**							10	-16.45
hsa-miR-4685-3p	-18.31	**150**	**169**							10	-18.31
hsa-miR-5006-3p	-18.59	**142**	**169**							10	-18.59

[1] Position of the miRNA-target duplex. Numbering is related to the L- 3'UTR, where position 1 is the chromosomal position chr2: 232787025 and position 232790137 is the stop codon.

[2] Maximal number of contiguous bases perfectly matching the target (the bold characters correspond to the starting/stop sites of the perfect pairing)

References

Babiarz, J.E., Ravon, M., Sridhar, S., Ravindran, P., Swanson, B., Bitter, H., Weiser, T., Chiao, E., Certa, U., Kolaja, K.L. 2012. Determination of the Human Cardiomyocyte mRNA and miRNA Differentiation Network by Fine-Scale Profiling. Stem Cells Dev.

Barr, C.S., Rhodes, P., Struthers, A.D. 1996. C-type natriuretic peptide. Peptides 17:1243-1251.

Boettger, T., Beetz, N., Kostin, S., Schneider, J., Krüger, M., Hein, L., Braun, T. 2009. Acquisition of the contractile phenotype by murine arterial smooth muscle cells depends on the Mir143/145 gene cluster. J Clin Invest 119:2634-2647.

Burnett, Jr, J., Kao, P.C., Hu, D.C., Heser, D.W., Heublein, D., Granger, J.P., Opgenorth, T.J., Reeder, G.S. 1986. Atrial natriuretic peptide elevation in congestive heart failure in the human. Science 231:1145-1147.

Clerico, A., Fontana, M., Ripoli, A., Emdin, M. 2009. Clinical relevance of BNP measurement in the follow-up of patients with chronic heart failure. Adv Clin Chem 48:163-179.

de Bold, A.J., Borenstein, H.B., Veress, A.T., Sonnenberg, H. 1981. A rapid and potent natriuretic response to intravenous injection of atrial myocardial extract in rats. Life Sci 28:89-94.

Del Ry, S., Cabiati, M., Stefano, T., Catapano, G., Caselli, C., Prescimone, T., Passino, C., Emdin, M., Giannessi, D. 2011b. Comparison of NT-proCNP and CNP plasma levels in heart failure, diabetes and cirrhosis patients. Regul Pept 166:15-20.

Del Ry, S., Cabiati, M., Vozzi, F., Battolla, B., Caselli, C., Forini, F., Segnani, C., Prescimone, T., Giannessi, D., Mattii, L. 2011a. Expression of C-type natriuretic peptide and its receptor NPR-B in cardiomyocytes. Peptides 32:1713-1718.

Del Ry, S., Giannessi, D., Maltinti, M., Prontera, C., Iervasi, A., Colotti, C., Emdin, M., L'Abbate, A., Neglia, D. 2007. Increased levels of C-type natriuretic peptide in patients with idiopathic left ventricular dysfunction. Peptides 28:1068-1073.

Del Ry, S., Maltinti, M., Cabiati, M., Emdin, M., Giannessi, D., Morales, M.A. 2008. C-type natriuretic peptide and its relation to non-invasive indices of left ventricular function in patients with chronic heart failure. Peptides 29:79-82.

Del Ry, S., Passino, C., Emdin, M., Giannessi, D. 2006. C-type natriuretic peptide and heart failure. Pharmacol Res 54:326-333.

Del Ry, S., Passino, C., Maltinti, M., Emdin, M., Giannessi, D. 2005. C-type natriuretic peptide plasma levels increase in patients with chronic heart failure as a function of clinical severity. Eur J Heart Fail 7:1145-1148.

Esau, C., Kang, X., Peralta, E., Hanson, E., Marcusson, E.G., Ravichandran, L.V., Sun, Y., Koo, S., Perera, R.J., Jain, R., Dean, N.M., Freier, S.M., Bennett, C.F., Lollo, B., Griffey, R. 2004. MicroRNA-143 regulates adipocyte differentiation. J Biol Chem 279:52361-52365.

Ikeda, S., Pu, W.T. 2010. Expression and function of microRNAs in heart disease. Curr Drug Targets 11:913-925.

Kalra, P.R., Clague, J.R., Coats, A.J., Anker, S.D., Poole-Wilson, P.A., Struthers, A.D. 2010. C-type natriuretic peptide production by the human kidney is blunted in chronic heart failure. Clin Sci (Lond) 118:71-77.

Kalra, P.R., Clague, J.R., Bolger, A.P., Anker, S.D., Poole-Wilson, P.A., Struthers, A.D., Coats, A.J. 2003. Myocardial production of C-type natriuretic peptide in chronic heart failure. Circulation 107:571-573.

Kalra, P.R., Anker, S.D., Struthers, A.D., Coats, A.J. 2001. The role of C-type natriuretic peptide in cardiovascular medicine. Eur Heart J 22:997-1007.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T. 2001. Identification of novel genes coding for small expressed RNAs. Science 294:853-858.

Lee, Y.S., Dutta, A. 2007. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes Dev 21:1025-1030.

Liu, J., Hao, D.-D., Zhang, J.-S., Zhu, Y.-C. 2011. Hydrogen sulphide inhibits cardiomyocyte hypertrophy by up-regulating miR-133a. Biochem Biophys Res Commun 413:342-347.

Liu, N., Bezprozvannaya, S., Williams, A.H., Qi, X., Richardson, J.A., Bassel-Duby, R., Olson, E.N. 2008. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. Genes Dev 22:3242-3254.

Najafi-Shoushtari, S.H., Kristo, F., Li, Y., Shioda, T., Cohen, D.E., Gerszten, R.E., Näär, A.M. 2010. MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis. Science 328:1566-1569.

Ogawa, Y., Nakao, K., Mukoyama, M., Hosoda, K., Shirakami, G., Arai, H., Saito, Y., Suga, S., Jougasaki, M., Imura, H. 1991. Natriuretic peptides as cardiac hormones in normotensive and spontaneously hypertensive rats. The ventricle is a major site of synthesis and secretion of brain natriuretic peptide. Circ Res 69:491-500.

Ono, K., Mannami, T., Baba, S., Tomoike, H., Suga, S.-i., Iwai, N. 2002. A single-nucleotide polymorphism in C-type natriuretic peptide gene may be associated with hypertension. Hypertens Res 25:727-730.

Passino, C., Del Ry, S., Severino, S., Gabutti, A., Prontera, C., Clerico, A., Giannessi, D., Emdin, M. 2008. C-type natriuretic peptide expression in patients with chronic heart failure: effects of aerobic training. Eur J Cardiovasc Prev Rehabil 15:168-172.

Schulz-Knappe, P., Forssmann, K., Herbst, F., Hock, D., Pipkorn, R., Forssmann, W.G. 1988. Isolation and structural analysis of "urodilatin", a new peptide of the cardiodilatin-(ANP)-family, extracted from human urine. Klin Wochenschr 66:752-759.

Schweitz, H., Vigne, P., Moinier, D., Frelin, C., Lazdunski, M. 1992. A new member of the natriuretic peptide family is present in the venom of the green mamba (Dendroaspis angusticeps). J Biol Chem 267:13928-13932.

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S.N., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J.M., Marra, M.A., , M.G.C.P.T. 2002. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. Proc Natl Acad Sci U S A 99:16899-16903.

Sudoh, T., Minamino, N., Kangawa, K., Matsuo, H. 1990. C-type natriuretic peptide (CNP): a new member of natriuretic peptide family identified in porcine brain. Biochem Biophys Res Commun 168:863-870.

Sudoh, T., Kangawa, K., Minamino, N., Matsuo, H. 1988. A new natriuretic peptide in porcine brain. Nature 332:78-81.

Sun, Y., Koo, S., White, N., Peralta, E., Esau, C., Dean, N.M., Perera, R.J. 2004. Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. Nucleic Acids Res 32:e188.

Thum, T., Gross, C., Fiedler, J., Fischer, T., Kissler, S., Bussen, M., Galuppo, P., Just, S., Rottbauer, W., Frantz, S., Castoldi, M., Soutschek, J., Koteliansky, V., Rosenwald, A., Basson, M.A., Licht, J.D., Pena, J.T.R., Rouhanifard, S.H., Muckenthaler, M.U., Tuschl, T., Martin, G.R., Bauersachs, J., Engelhardt, S. 2008. MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts. Nature 456:980-984.

Ucar, A., Vafaizadeh, V., Jarry, H., Fiedler, J., Klemmt, P.A.B., Thum, T., Groner, B., Chowdhury, K. 2010. miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development. Nat Genet 42:1101-1108.

van Rooij, E. 2011. The art of microRNA research. Circ Res 108:219-234.

van Rooij, E., Quiat, D., Johnson, B.A., Sutherland, L.B., Qi, X., Richardson, J.A., Kelm, Jr, R.J., Olson, E.N. 2009. A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. Dev Cell 17:662-673.

van Rooij, E., Sutherland, L.B., Qi, X., Richardson, J.A., Hill, J., Olson, E.N. 2007. Control of stress-dependent cardiac growth and gene expression by a microRNA. Science 316:575-579.

van Rooij, E., Sutherland, L.B., Liu, N., Williams, A.H., McAnally, J., Gerard, R.D., Richardson, J.A., Olson, E.N. 2006. A signature pattern of stress-responsive microRNAs that can evoke cardiac hypertrophy and heart failure. Proc Natl Acad Sci U S A 103:18255-18260.

Zhou, B., Wang, S., Mayr, C., Bartel, D.P., Lodish, H.F. 2007. miR-150, a microRNA expressed in mature B and T cells, blocks early B cell development when expressed prematurely. Proc Natl Acad Sci USA 104:7080-7085.