

# MicroRNA memory II: A Novel Scoring Integration Model for Prediction of Human Diseases by MicroRNA/MicroRNA Quantum Multi-Interaction

## ABSTRACT

**Aims:** Further, to investigate quantum characters of the microRNA (miRNA) gene as the disease memory device, we described neo-score tool for miRNA/miRNA multi-interaction. Since the potential miRNA/target interaction is not one-to-one correspondence; therefore, the network of miRNA/protein gene is too busy to achieve the goal of prognosis and diagnosis on human diseases.

**Methodology:** Neo-score tool based on quantum and electrodynamics for miRNA/miRNA multi-interaction, which are Dynamic Nexus Score (DNS) and Electric Field Tangent score (EFTS) is compared with Context+ Score Change (CSC) on PolymiRTS database, an integrated platform of the functional impact of genetic polymorphisms in miRNA seed regions and miRNA target sites within 150 of human diseases.

**Results:** the DNS was correlated with CSC. Since the EFTS was mathematically functioned into the DNS, the function of the DNS to the EFTS were integrated together disease prediction on CSC. Further, a possibility was suggested in this context that miRNA/miRNA multi-interaction on the algorithmic function could be applied for specific discernment of diseases without miRNA/target interaction complex.

**Conclusion:** Thus, miRNA/miRNA multi-interaction may have an important role for a *priori* event to control human health.

MicroRNA (miRNA) is the gene. The miRNA gene is the key tuner to control almost all of processes events in cells of lives including humans. The tuning mechanisms of miRNAs are dominantly implicated in direct or indirect transcriptional or post-transcriptional regulation. In the case of post-transcriptional suppression, a short seed region of miRNA can be recruited to the 3'UTR of messenger RNA (mRNA) as the target with Argonaute (Ago). For the prediction or validation of functional miRNAs, a large amount of computational complex task has been preceded with the miRNA target prediction tools before *in vitro* and *in vivo* experiments. The potential miRNA/target interaction is not one-to-one correspondence; therefore, the network of miRNA/protein gene is too busy to achieve the goal of prognosis and diagnosis on human diseases. On the other hand, miRNA/miRNA interaction is not so clear. Beyond this difficult state on the status quo miRNA/miRNA interaction was integrated to improve the prediction accuracy and simplicity. Here, we described neo-score tool based on quantum and electrodynamics for miRNA/miRNA multi-interaction, which are Dynamic Nexus Score (DNS) and Electric Field Tangent score (EFTS), respectively. As the results, the DNS was correlated with Context+ Score Change (CSC) on PolymiRTS database, an integrated platform of the functional impact of genetic polymorphisms in miRNA seed regions and miRNA target sites within 150 of human diseases. Since the EFTS was mathematically functioned into the DNS, the function of the DNS to the EFTS were integrated together disease prediction on CSC. Further, a possibility was suggested in this context that miRNA/miRNA multi-interaction on the algorithmic function could be applied for specific discernment of diseases without miRNA/target interaction complex. Thus, miRNA/miRNA multi-interaction may have an important role for a *priori* event to control human health.

**Keyword:** human disease, miRNA, quantum, mRNA, PolymiRTS

# 1. INTRODUCTION

Recently, the function of non-coding RNAs (ncRNAs) has been investigated intensely, because ncRNAs have a key role in almost all cellular processes[1–3]. The ncRNAs are classified according to their length and function. The operation of long ncRNAs, whose length is more than 200 nucleotides, remains unclear. However, one type of small ncRNAs, the microRNAs (miRNAs), have an important direct or indirect role in the transcriptional or post-transcriptional regulation of gene expression and are highly conserved in various species. The miRNAs are approximately 22 nucleotides long, and their function is predicted on the basis of their seed region, which consists of nucleotides 2–8 at the 5' end. The seed region mediates binding to the target genes and to the 3' untranslated region (3' UTR) of the messenger RNAs (mRNAs) and the miRNA/mRNA interaction; thus, it downregulates gene expression[4–6]. Competitive binding of the seed, target and competing endogenous RNAs regulates the protein expression level[7–9].

Various databases and miRNA target prediction tools are available currently. For instance, TargetScan identifies candidate miRNAs based on the strength of interaction between the seed region and target site of miRNAs. TargetScan introduced the Context+ Score value, which enables the selection of the most favorable target sites for miRNAs. Context+ Score value evaluates miRNA binding in the context of the entire 3' UTR of the mRNA by summing contributions made by individual sites that are perfectly complementary to the miRNA seed[10]. Context+ Score value consists of the following six contributions: site-type, 3' pairing, local AU nucleotides, position, target site abundance (TA) and seed-pairing stability (SPS)[10]. This value is calculated for one miRNA and one target region. Another database, miRBase, can be used to obtain mature miRNA sequences[11], and dbSNP collects data about human single nucleotide polymorphisms (SNPs)[12]. Because TargetScan, miRBase and dbSNP are only available independently to investigate miRNA/mRNA interactions and SNPs, a combined prediction tool was lacking for more complex prediction tasks. In 2014, Bhattacharya et al. integrated these three databases creating the PolymiRTS combined prediction tool[13–15]. PolymiRTS database 3.0 (PolymiRTS) collects SNPs located in miRNAs and miRNA target sites that are verified experimentally. In addition, PolymiRTS contains numerous miRNA targets, which are implicated in human diseases, and data about the interaction between miRNA seed regions and 3' UTRs of mRNAs. Context+ Score change (CSC) value is the Context+ Score value corrected with the effect of polymorphisms in the miRNA target sites and miRNA seed regions. Compared with references, a more negative CSC value indicates an increase in the binding strength of miRNA and mRNA. A specific mRNA is often targeted by the same miRNA at multiple sites. CSC represents the change in binding strength between multiple miRNAs and SNPs within the mRNAs.

RNA Wave 2000 Model is the model interpreted RNA information genes (Rigs)[16]. The RNA Wave 2000 Model has four pillars; 1) the Rigs as a mobile genetic element induce transcriptional and post-transcriptional silencing via networking-architecture; 2) the Rigs expand into the environmental cycle of life; 3) the Rigs can self-proliferate; 4) the Rigs have two types of information as resident and genomic ones. The RNA Wave 2000 Model describes a novel mechanism of miRNA regulation, which can be affected by environmental factors[16]. The coherent on the quantum direction in the high dimensions has been shown in the matrix with the miRNA qubit code; however, the coherent was unavailable to apply for challenging against miRNA/miRNA multi-interaction.

~~Because CSC cannot evaluate miRNA/miRNA interactions,~~ Previously, we developed two new scores to represent multiple miRNA/miRNA interactions based on the RNA Wave 2000 Model (MicroRNA memory I). One is based on quantum[17] and the other is based on electrodynamics[18]. Quantum based score, Dynamic Nexus Score (DNS), and the electrodynamics based score, Electric Field Tangent score (EFTS), were compared with

CSC because CSC cannot evaluate miRNA/miRNA interactions. CSC was linearly correlated with DNS. DNS and EFTS were non-linearly associated with human diseases independent of CSC. After gaining an insight into miRNA/miRNA multiple interactions using network prediction tools, we might use this information for tailored medical care.

74

75

## 2. METHODOLOGYS

77

### 2.1 MiRNA and mRNA Databases

All CSC data placed in Diseases and Traits were prepared from PolymiRTS (<http://compbio.uthsc.edu/miR SNP/>). In PolymiRTS, the cut off value of <50 of miRNA in one target mRNA was used. The mature.fa file in miRBase (<http://mirbase.org/>) was downloaded in Release 21: June 2014, miRNA counts: 28,645 entries from a downloading page. All 2,588 of human miRNAs were extracted. CSCs related with 2,297 mRNA of 150 diseases are used in this study (Table 1).

84

Number of Significant MiRNA Target Loci

| Phenotype                           | Number of diseases | Number of targets | Number of significant correlation* | Phenotype                    | Number of diseases | Number of targets | Number of significant correlation* |
|-------------------------------------|--------------------|-------------------|------------------------------------|------------------------------|--------------------|-------------------|------------------------------------|
| <b>Tumor</b>                        | 39                 | 350               | 125                                | <b>Immune disease</b>        | 8                  | 164               | 69                                 |
| prostate cancer                     |                    | 64                | 20                                 | systemic lupus erythematosus |                    | 45                | 16                                 |
| breast cancer                       |                    | 39                | 13                                 | celiac disease               |                    | 41                | 20                                 |
| others                              |                    | 247               | 92                                 | others                       |                    | 78                | 33                                 |
| <b>Neurodegeneration</b>            | 24                 | 375               | 126                                | <b>Infection</b>             | 6                  | 37                | 13                                 |
| bipolar disorder                    |                    | 69                | 18                                 | leishmaniasis                |                    | 14                | 6                                  |
| schizophrenia                       |                    | 60                | 22                                 | AIDS progression             |                    | 9                 | 4                                  |
| others                              |                    | 246               | 86                                 | others                       |                    | 14                | 3                                  |
| <b>Bone and Muscle</b>              | 15                 | 150               | 65                                 | <b>Skin</b>                  | 6                  | 90                | 37                                 |
| rheumatoid arthritis                |                    | 42                | 18                                 | vitiligo                     |                    | 49                | 19                                 |
| amyotrophic lateral sclerosis       |                    | 23                | 6                                  | psoriasis                    |                    | 22                | 13                                 |
| others                              |                    | 85                | 41                                 | others                       |                    | 19                | 5                                  |
| <b>Cardio disease</b>               | 15                 | 170               | 61                                 | <b>Renal dysfunction</b>     | 5                  | 54                | 20                                 |
| coronary heart disease              |                    | 57                | 24                                 | chronic kidney disease       |                    | 36                | 13                                 |
| hypertension                        |                    | 24                | 6                                  | nephropathy                  |                    | 11                | 3                                  |
| others                              |                    | 89                | 31                                 | others                       |                    | 7                 | 4                                  |
| <b>Gastrointestinal disturbance</b> | 12                 | 367               | 147                                | <b>Others</b>                | 19                 | 540               | 222                                |
| crohn's disease                     |                    | 137               | 43                                 |                              |                    |                   |                                    |
| inflammatory bowel disease          |                    | 133               | 55                                 |                              |                    |                   |                                    |
| others                              |                    | 97                | 49                                 |                              |                    |                   |                                    |
|                                     |                    |                   |                                    | <b>Total</b>                 | 149                | 2297              | 885                                |

\*Significance of DNS and CSC correlation.  $p < 0.05$

state the abbreviations

85

86

87

88

Table 1. The Number of MiRNA's Target in PolymiRTS Used in This Study

DNSC stand for what? u explained DNS before but not DNSC

### 2.2 The New Score of DNSC

The miR-ket code and single miRNA/miRNA superposition in matrix were described previously (Fujii, 2013b). The DNS is equal to the sum of each mathematical clause of the direct product when two miRNA sequences were diagnosed as a vector. Therefore, the miRNA group was defined as  $A$ , each element of  $A$  as  $A_i \{i = 0, 1, 2, \dots, N - 1\}$ , and sequence of each element as  $A_{ij} \{j = 0, 1, 2, \dots, 23\}$ .

94

stands for what

what 23 stands for? chromosomes??

$$DNS = \sum_{j=0}^{23} A_{0j} \times \sum_{j=0}^{23} A_{1j} \times \dots \times \sum_{j=0}^{23} A_{(N-1)j} \quad (1, 0)$$

The DNS (1, 0) is shown as synergy between among miRNAs (Text S1) in our previous report (MicroRNA memory I).

Next,  $\delta$  DNS were calculated as a change of the DNS. When miRNA group element shown as  $A_x$  and sequence of  $A_x$  as  $A_{xj}\{j = 0, 1, 2, \dots, 23\}$ ,  $\delta$  DNS was expressed as follows,

$$\delta DNS = DNS - \frac{DNS}{\sum_{j=0}^{23} A_{xj}} \quad (1, 1)$$

This equation was rearranged as follows (1, 1).

$$\delta DNS = \frac{DNS(\sum_{j=0}^{23} A_{xj} - 1)}{\sum_{j=0}^{23} A_{xj}} \quad (1, 2)$$

Furthermore, to simplify this calculation (1, 2) was defined as DNSC.

$$DNSC = \frac{\ln(\delta DNS)}{N} = \frac{1}{N} \left( \sum_{i=0}^{N-1} \left( \ln \left( \sum_{j=0}^{23} A_{ij} \right) \right) + \ln \left( \sum_{j=0}^{23} A_{xj} - 1 \right) - \ln \left( \sum_{j=0}^{23} A_{xj} \right) \right) \quad (1, 3)$$

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

## 2.3 The New Score of EFTS

Although the DNSC is the score that quantized the quantum energy in each base of miRNA, the new score based on electrodynamics was further developed. The electric field vector as  $E1$  in any coordinate (x, y) to receive from electric charge Q in coordinate (X, Y) was expressed as follows.

$$\vec{E1} = k \frac{Q(x-X)}{\{(x-X)^2 + (y-Y)^2\}^{\frac{3}{2}}} \vec{x} + k \frac{Q(y-Y)}{\{(x-X)^2 + (y-Y)^2\}^{\frac{3}{2}}} \vec{y} \quad (2, 0)$$

120

121

122

123

124

125

126

127

128

129

When the number of electric charge increases, the electric field vector is demanded in the sum of the vector like the following expression as  $E2$ . The total number as N, each electric charge as  $Q_i$ , each coordinate as  $(X_i, Y_i)\{i = 0, 1, 2, \dots, N - 1\}$  was used as follows.

$$\vec{E2} = k \sum_{i=0}^{N-1} \frac{Q_i(x-X_i)}{\{(x-X_i)^2 + (y-Y_i)^2\}^{\frac{3}{2}}} \vec{x} + k \sum_{i=0}^{N-1} \frac{Q_i(y-Y_i)}{\{(x-X_i)^2 + (y-Y_i)^2\}^{\frac{3}{2}}} \vec{y} \quad (2, 1)$$

126

127

128

129

The hydrogen atom was defined as +1 in a nucleic acid (Figure 1, red circle). The nitrogen atom and the oxygen atom were defined as -1 in a nucleic acid (Figure 1, blue circle). The hydrogen atoms red-circled are related with hydrogen bonds.

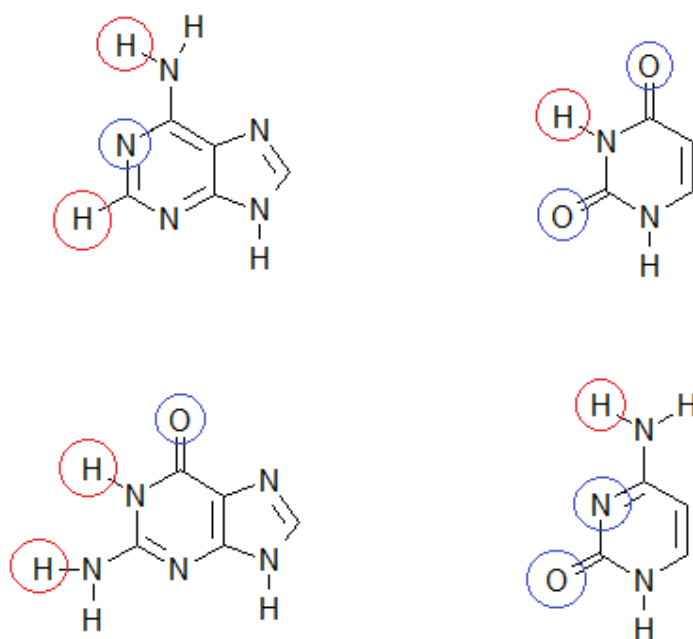


**Figure 1. EFTS Related Atoms in the Nucleotide**

Further, (2, 1) applied for miRNA. MiRNA sequence was defined as  $A_i \{i = 0, 1, 2, \dots, M\}$  and charge of each  $A_i$  as  $A_{ij} \{j = 0, 1, 2\}$ . Coordinates of  $A_{ij}$  was shown as  $(X_{ij}, Y_{ij})$ , electric field vector  $E_3$  was expressed as follows.

$$\vec{E_3} = k \sum_{i=0}^M \sum_{j=0}^2 \frac{A_{ij}(x - X_{ij})}{\{(x - X_{ij})^2 + (y - Y_{ij})^2\}^{\frac{3}{2}}} \vec{x} + k \sum_{i=0}^M \sum_{j=0}^2 \frac{A_{ij}(y - Y_{ij})}{\{(x - X_{ij})^2 + (y - Y_{ij})^2\}^{\frac{3}{2}}} \vec{y} \quad (2, 2)$$

In this formula, miRNA sequence was circularized in this model and  $(x, y)$  was any optional point.

The X ingredient ( $E_x$ ) and the Y ( $E_y$ ) ingredient of the electric field vector were separated as follows.

$$\vec{E_x} = k \sum_{i=0}^M \sum_{j=0}^2 \frac{A_{ij}(x - X_{ij})}{\{(x - X_{ij})^2 + (y - Y_{ij})^2\}^{\frac{3}{2}}} \vec{x} \quad (2, 3)$$

$$\vec{E_y} = k \sum_{i=0}^M \sum_{j=0}^2 \frac{A_{ij}(y - Y_{ij})}{\{(x - X_{ij})^2 + (y - Y_{ij})^2\}^{\frac{3}{2}}} \vec{y} \quad (2, 4)$$

The miRNA group as  $A$ , each element of  $\vec{E_x}$  and  $\vec{E_y}$  of  $A$  as  $\vec{E_{ix}}$  and  $\vec{E_{iy}}$   $\{i = 0, 1, 2, \dots, N - 1\}$ , respectively. Total electric field vector  $E_4$  was expressed as follows.

$$\vec{E_4} = \sum_{i=0}^{N-1} \vec{E_{ix}} + \sum_{i=0}^{N-1} \vec{E_{iy}} \quad (2, 5)$$

151  
152  
153  
154  
155  
156  
157  
158

The radius of miRNA was defined as 50 units and one unit was corresponding to a lattice. The electric field was expressed at lattice points. The coordinate of lattice point was defined as  $(x_{kl}, y_{kl}) \{k = -50, \dots, 0, \dots, 50; l = -50, \dots, 0, \dots, 50\}$  and each ingredient of the electric field vector as  $\vec{E}_{klx}$  and  $\vec{E}_{kly}$ . Tangential value when the electric field vector in each coordinates was defined as Electric Field Tangent score (EFTS). The formula is as follows.

$$159 \quad EFTS_1 = \frac{\sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{kly}}{\sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{klx}} \quad (2, 6)$$

160  
161  
162  
163

Further,  $EFTS_1$  (2, 6) in a miRNA were changed to  $EFTS_2$  in multiple miRNAs and the number of miRNA was shown as N.

$$164 \quad EFTS_2 = \frac{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nkly}}{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nklx}} \quad (2, 7)$$

165  
166  
167  
168  
169

In (2, 7), a component of electric field vector in each miRNA was shown as

$\vec{E}_{nklx}, \vec{E}_{nkly} \{n = 0, 1, \dots, N-1\}$ , EFTSC was derived from  $EFTS_2$  (2, 7). Next, EFTSC were calculated as a change of the EFTS.

$$EFTSC = \frac{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nkly}}{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nklx}} - \frac{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nkly} - \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{kly}}{\sum_{n=0}^{N-1} \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{nklx} - \sum_{k=-50}^{50} \sum_{l=-50}^{50} \vec{E}_{klx}} \quad (2, 8)$$

170  
171

## 2.4 Mathematical Functioning EFTSC into DNSC

172  
173  
174  
175  
176  
177

From (1, 0), the scalar product is corresponding to number of G, therefore, DNSC can be replaced for number of G. EFTSC was calculated by Fourier transformation as analogy of quantum computation. Mean of EFTSC, y and number of G (DNSC), x were functioned as follows,

$$178 \quad y = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left( \cos\left(\frac{2\pi ij}{N}\right) \cos(ix) - \frac{\sin\left(\frac{2\pi ij}{N}\right)}{N} \sin(ix) \right) \quad (3, 0)$$

179  
180  
181  
182

## 3. RESULTS

183

### 3.1 Molecular Modeling of miRNA/miRNA Multi-Interaction

184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201

DNS and CSC analysis were performed to evaluate miRNA/miRNA multi-interaction in aging, bone and muscle diseases, cardio diseases, endocrine, eye diseases, genital, gastrointestinal disturbance, immune diseases, infection, lifestyle-related diseases, neurodegeneration, renal dysfunction, respiratory diseases, skin diseases, tooth and tumors (Table S1). The number of significant correlation between DNS and CSC was shown in Table 1. In cancer, chronic lymphocytic leukemia, CSCs of miRNAs in SNPs of RFX7 mRNA were plotted to DNSCs of same miRNAs (Figure 2A). CSCs in the same DNSCs average was calculated and plotted to DNSCs in Figure 2A (Figure 2B).  $R^2$  was contained as 0.8488, therefore the average of CSCs was highly correlated to DNSCs. Pearson correlation coefficient between mean of CSCs and DNSCs was calculated in all diseases, and p-values were plotted in (Figure 2C). Gaussian distribution was observed. Under less than 1% of p-value were 391 (17%) of target, under less than 5% of p-value were 885 (39%) of target and under less than 10% of p-value were 1190 (52%) of target (Figure 2D). Top 5 of linear relation ( $p < 0.01$ ) between DNSC and CSC were shown in Table 2. In age-related macular degeneration, SNP allele alteration of 3'UTR target of GLI3 (NM\_000168), MBL2 (NM\_000242) and ARMS2 (NM\_001099667) were strongly correlated with DNSC. Significant correlations against truncating variants in SNPs of

TXNDC16 tv2 (NM\_001160047), GRM7 tv1 (NM\_000844) and GREM1 tv1 (NM\_013372) were also observed in orofacial cleft. About coronary heart disease, three SNPs in HLA-DPA1 tv3 (NM\_001242525), SLC22A3 (NM\_021977) and LAMC2 (NM\_005562) had the high relationship to DNSC. Crohn's disease was clearly implicated in SNPs of EHB1 tv3 (NM\_001142615) and DNSC or PDLIM4 tv2 (NM\_001131027) and DNSC. In the case of type 2 diabetes mellitus, PAX4 (NM\_006193), PLS1 tv2 (NM\_002670), MAP3K1 (NM\_005921) and CDC123 (NM\_006023) showed profound nexuses of DNSC. Further, DNS was allele-specifically correlated with SNPs of ST8SIA2 (NM\_006011), CREB5 tv3 (NM\_182899) and SRP72 (NM\_006947) in airflow obstruction, and CHRNA4 (NM\_000744) and IREB2 (NM\_004136) in chronic obstructive pulmonary disease. SNPs of PCLO tv1 (NM\_014510), ESRRG tv4 (NM\_001134285), STMN2 tv1 (NM\_001199214) and GRM7 tv1 (NM\_000844) were significantly related with DNSC in neurodegeneration of the major category of human diseases. These results indicate that miRNA/miRNA interaction computed by DNSs provides different functions for browsing and searching data about SNPs and human disease and would be important candidates for disease phenotype prediction and complex trait studies.

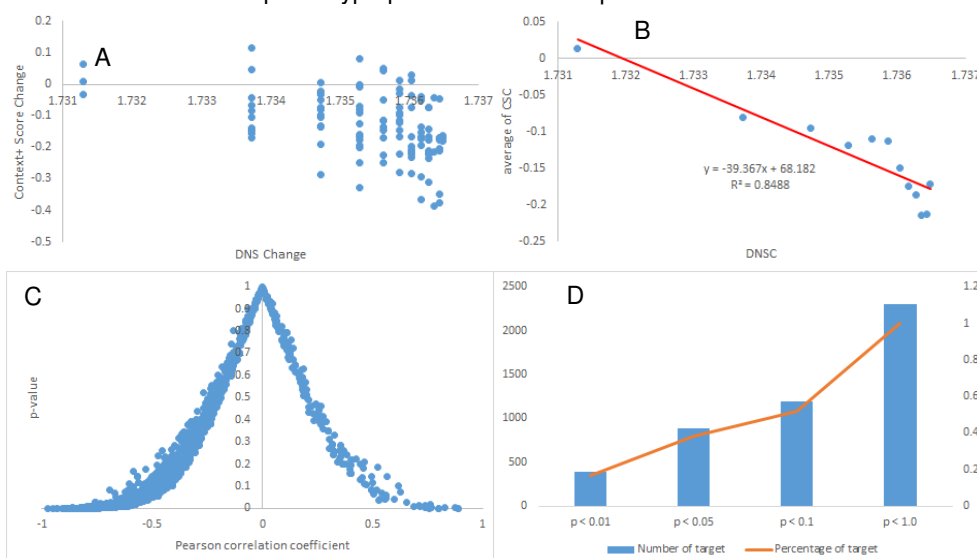


Figure 2. Correlation between Quantum Score and the Target Seed.

| Top 5 of linear relation between DNSC and CSC |                          |             |                    |                                  |              |                            |                          |             |                    |                                       |              |
|-----------------------------------------------|--------------------------|-------------|--------------------|----------------------------------|--------------|----------------------------|--------------------------|-------------|--------------------|---------------------------------------|--------------|
| Phenotype                                     | Top 5 of linear-relation | p-value     | Morphogenic Target | Disease                          | RefSeqID     | Phenotype                  | Top 5 of linear-relation | p-value     | Morphogenic Target | Disease                               | RefSeqID     |
| Aging                                         | -0.845290358             | 0.000140632 | GLI3               | age-related macular degeneration | NM_000168    | Lifestyle-related diseases | -0.907718358             | 4.50643E-05 | PAX4               | Type 2 diabetes                       | NM_006193    |
|                                               | -0.834991958             | 0.000382272 | MBL2               | age-related macular degeneration | NM_000242    |                            | -0.890754612             | 0.000234895 | PLS1 tv2           | Type 2 diabetes                       | NM_002670    |
|                                               | -0.80529956              | 0.001572887 | ARMS2              | age-related macular degeneration | NM_001099667 |                            | -0.863868578             | 0.005680496 | MAP3K1             | Type 2 diabetes                       | NM_005921    |
|                                               | -0.791520692             | 0.000438571 | ATCAY              | aging                            | NM_033064    |                            | -0.854462938             | 0.000199328 | DYRK1A tv5         | Metabolic syndrome                    | NM_130438    |
|                                               | -0.782353553             | 0.000566449 | SH3BGRL2           | aging                            | NM_174963    |                            | -0.848247823             | 0.000247746 | CDC123             | Type 2 diabetes                       | NM_006023    |
| Bone and Muscle                               | -0.898195673             | 1.28649E-05 | TXNDC16 tv2        | Orofacial clefts                 | NM_001180047 | Neurodegeneration          | -0.883022757             | 2.86199E-05 | PCLO tv1           | Major depressive disorder             | NM_014510    |
|                                               | 0.889635021              | 2.04874E-05 | POU3F1             | Rheumatoid arthritis             | NM_002699    |                            | 0.873516273              | 0.00459101  | ARHGAP22           | Conduct disorder                      | NM_001256026 |
|                                               | -0.859538242             | 0.000338505 | IL23R              | Ankylosing spondylitis           | NM_144701    |                            | -0.854547347             | 0.000198727 | ESRRG tv4          | Major depressive disorder             | NM_001134285 |
|                                               | -0.847744619             | 0.000496137 | GRM7 tv1           | Orofacial clefts                 | NM_000844    |                            | -0.853508761             | 0.000831565 | STMN2 tv1          | Creutzfeldt-Jakob disease             | NM_001199214 |
|                                               | -0.842426081             | 0.000301169 | GREM1 tv1          | Orofacial clefts                 | NM_013372    |                            | -0.847744619             | 0.000496137 | GRM7 tv1           | Parkinson's disease                   | NM_000844    |
| Cardio disease                                | -0.824819815             | 0.000520313 | HLA-DPA1 tv3       | Coronary heart disease           | NM_001242525 | Renal dysfunction          | -0.928109182             | 1.70478E-06 | CST3               | Chronic kidney disease                | NM_000099    |
|                                               | -0.822069621             | 0.001033163 | SLC22A3            | Coronary heart disease           | NM_021977    |                            | -0.888962965             | 0.000110021 | CST9               | Chronic kidney disease                | NM_001008693 |
|                                               | -0.799898989             | 0.000591578 | LAMC2 tv1          | Coronary heart disease           | NM_005562    |                            | -0.824819815             | 0.000520313 | HLA-DPA1 tv3       | Nephropathy                           | NM_001242525 |
|                                               | -0.794191926             | 0.000692815 | ABO                | Venous thromboembolism           | NM_020469    |                            | -0.80415005              | 0.000526544 | AQP1 tv1           | Nephrolithiasis                       | NM_198098    |
|                                               | -0.794191926             | 0.000692815 | ABO                | Coronary heart disease           | NM_020469    |                            | -0.782871671             | 0.000558494 | CYB561D1 tv5       | Chronic kidney disease                | NM_001134404 |
| gastrointestinal disturbance                  | -0.974117294             | 1.89276E-08 | SPIB tv1           | Primary biliary cirrhosis        | NM_003121    | Respiratory                | -0.828077145             | 0.000137626 | CHRNA4             | Chronic obstructive pulmonary disease | NM_000744    |
|                                               | -0.901172526             | 2.60356E-05 | EHBP1 tv3          | Crohn's disease                  | NM_001142615 |                            | -0.777122836             | 0.000652041 | ST8SIA2            | Airflow obstruction                   | NM_006011    |
|                                               | -0.880688071             | 3.20543E-05 | PDLIM4 tv2         | Inflammatory bowel disease       | NM_001131027 |                            | -0.773219987             | 0.000722518 | CREB5 tv3          | Airflow obstruction                   | NM_182899    |
|                                               | -0.880688071             | 3.20543E-05 | PDLIM4 tv2         | Crohn's disease                  | NM_001131027 |                            | -0.714685681             | 0.0134544   | IREB2              | Chronic obstructive pulmonary disease | NM_004136    |
|                                               | -0.859538242             | 0.000338505 | IL23R              | Ulcerative colitis               | NM_144701    |                            | -0.677364791             | 0.022028364 | SRP72              | Airflow obstruction                   | NM_006947    |
| Immune disease                                | -0.931790666             | 3.59782E-06 | GHR tv12           | Systemic lupus erythematosus     | NM_001242462 | Skin                       | -0.859538242             | 0.000338505 | IL23R              | Psoriasis                             | NM_144701    |
|                                               | -0.878514336             | 7.74575E-05 | TSHR tv1           | Graves' disease                  | NM_000369    |                            | -0.824819815             | 0.000520313 | HLA-DPA1 tv3       | Vitiligo                              | NM_001242525 |
|                                               | -0.871092268             | 0.000480311 | SLC15A4            | Systemic lupus erythematosus     | NM_145648    |                            | -0.813951                | 0.002295363 | IL12B              | Psoriasis                             | NM_002187    |
|                                               | -0.859538242             | 0.000338505 | IL23R              | Behcet's disease                 | NM_144701    |                            | -0.793935896             | 0.003524671 | CDH23 tv2          | Vitiligo                              | NM_052836    |
|                                               | -0.845290358             | 0.000140632 | GLI3               | Allergic rhinitis                | NM_000168    |                            | -0.790907983             | 0.00128436  | BNC2               | Freckles                              | NM_017637    |
| Infection                                     | -0.880778738             | 7.01592E-05 | RNF39 tv1          | AIDS progression                 | NM_025236    | Tumor                      | -0.921297968             | 2.08135E-05 | RFX7               | Chronic lymphocytic leukemia          | NM_022841    |
|                                               | -0.859538242             | 0.000338505 | IL23R              | Leprosy                          | NM_144701    |                            | -0.901172526             | 2.60356E-05 | EHBP1 tv3          | Prostate cancer                       | NM_001142615 |
|                                               | -0.824819815             | 0.000520313 | HLA-DPA1 tv3       | Leishmaniasis                    | NM_001242525 |                            | -0.896331064             | 7.90568E-05 | CLDN11 tv1         | Prostate cancer                       | NM_005602    |
|                                               | -0.794191926             | 0.000692815 | ABO                | Malaria                          | NM_020469    |                            | -0.865311942             | 0.000133047 | TAP2 tv2           | Lymphoma                              | NM_018833    |
|                                               | -0.758981635             | 0.004203325 | SCO1               | Malaria                          | NM_004589    |                            | -0.863868578             | 0.005680496 | MAP3K1             | Breast cancer                         | NM_005921    |

Table 2. Top 5 of Correlation between CSC and DNSC in Each Phenotype

3.2 Relations among EFTSs, DNSCs and CSCs

When the electric field vector in EFTS was applied for nucleotides of miRNA, E3 vectors of each nucleotide as torus showed specific characters among miRNAs as example, miR-10a-5p, miR-27a-5p, miR-191-5p, miR-200a-5p, miR-573 and miR-592 (Figure S2A). Further, under mathematically simulated condition, coherence state of two miRNA, such as miR-27a-5p (stayed) and miR-573 (moving) was made as two scalarsand E4 was

presented as a torus in Hilbert space (Figure S2B). Superposing state of two miRNA/miRNA interaction was identical among every prohibiting set of miRNA/miRNA interaction, which also be corresponding to identical points of miRNAs between the CSC in the x-axis (red), DNSC in the y-axis (green) and the EFTSC in the z-axis (blue) on Figure 3A. Therefore, these scalar values were transformed for multi miRNA/miRNA coherence as binary annotation in consideration of Watson-Crick and wobble base pairings, and then EFTSC score was computed. To further improve the accuracy of miRNA/miRNA multi-interaction EFTSC was integrated into the relation between CSC and DNSC in chronic lymphocytic leukemia, RFX7 target (Figure 3A and Figure S3A). These data showed that CSCs have linear relation with DNSCs (See Pearson correlation coefficient in Table S1) but all CSCs were not related with EFTSCs. Increasing of target further computation was performed. A figure of three-dimensional plot of mRNA related to chronic lymphocytic leukemia, 11 targets were visualized in web (Figure S3B). Next, in bladder cancer, breast cancer, chronic lymphocytic leukemia and colorectal cancer, the DNSC plotting was performed and DNSCs in each disease were shown in Figure 3B. In three-dimensional plot of mRNA related with Figure 3B (Figure S3C), CSCs, DNSCs and EFTSCs were visualized. In the case of another diseases, atopic dermatitis, bladder cancer, biliary obstruction and chronic pulmonary disease, the DNSCs were processed at the same way (Figure 3C and Figure S3D). All CSCs, all DNSCs and all EFTSCs did not show significant relation among them under summing up of targets or diseases. Since we did not obtain the possible formula to make impact of miRNA/miRNA multi-interaction, DNSC was compared to EFTSC in chronic lymphocytic leukemia, RFX7 target (Figure 3D). The number of G-base was plotted to EFTSC, non-linear regression analysis was performed (blue line) and non-linear relation was observed. By definition of DNSC, DNSC is uniquely determined number of G. Therefore, we confirmed relation between EFTSC and DNS (by number of G).

Explain more what is number of G

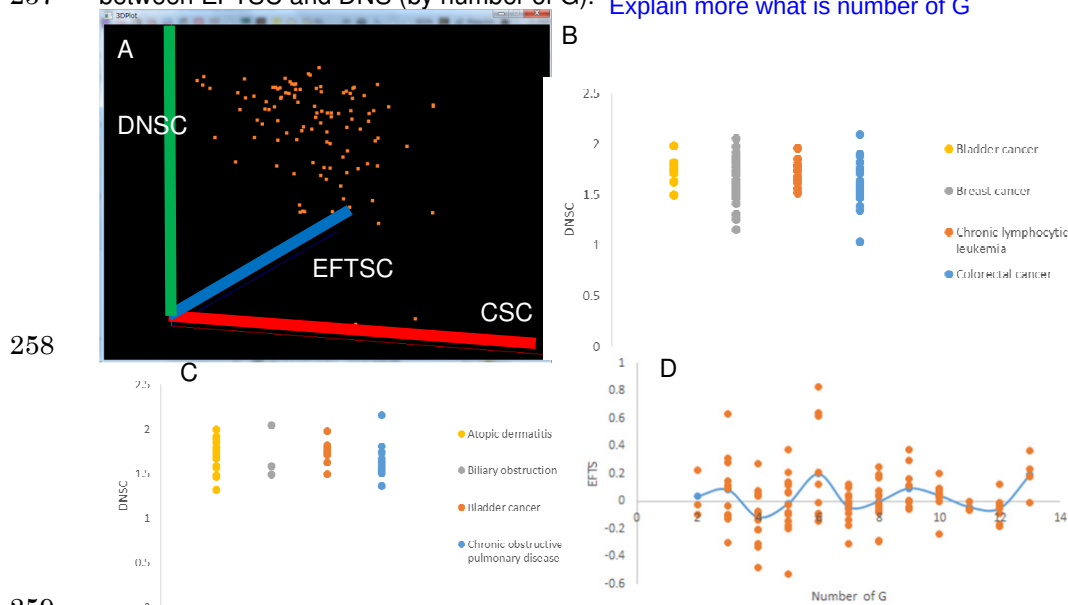


Figure 3. Integration of EFTSC into Relation between DNSC and CSC.

### 3.3 Function of EFTSCs by DNSCs

As by non-linear regression analysis with Fourier transformation, number of G (x) and EFTSC (y) had some formula, a formula was prepared to be functioned by DNSCs. About top 5 of linear relation between DNSC and CSC (Table 2), 4, 5 and 6 of G (x) were substitute for the formula (Table 3). In each number of G in a phenotype, the maximum (red) and the minimum (blue) values of average of EFTSCs were represented. Although Pearson correlation efficient did not identify the value of average of EFTSCs, non-linear

curve was computed on the targets, APOBEC3A tv1, TMEM129 tv2, PSCA tv1, CCNE1, FGFR3 tv3, SLC14A1 tv1, TP63 tv5 and TP63 tv3 in the bladder cancer (Figure 4A). In Table 3, the maximum and the minimum values were tremendously different among  $x=4$ , 5 and 6. Therefore, each value was adjusted from -100 to 100 to make it easy to compare. Next, the average of each non-linear data was calculated on the same target as described above in the bladder cancer (Figure 4B). Similar processes were performed in bladder cancer, breast cancer, chronic lymphocytic leukemia and colorectal cancer (Figure 4C). Further, in atopic dermatitis, biliary obstruction, bladder cancer and chronic obstructive pulmonary diseases, non-linear analysis was done (Figure 4D). These results indicate that function of EFTSC by DNSC specifically showed disease phenotype modes.

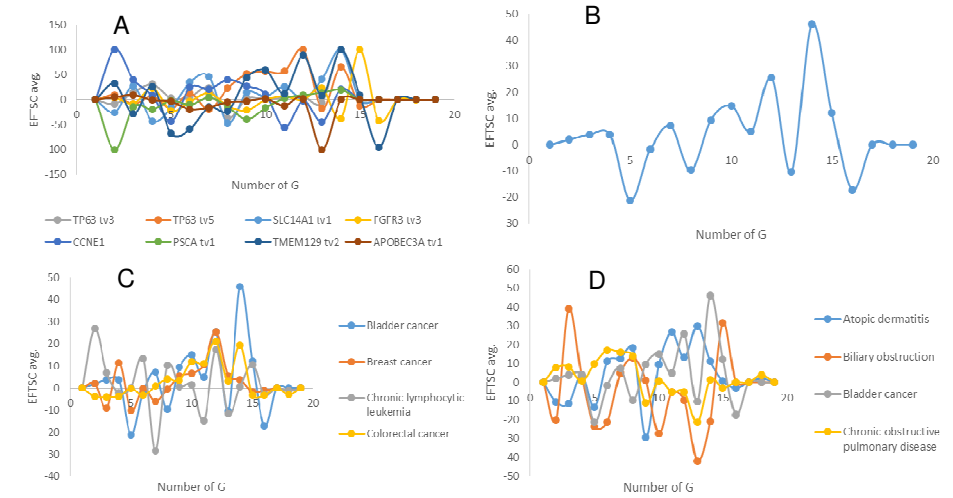
Genome-wide association studies (GWAS) have identified miRNA and SNPs of target implicated in complex human traits and diseases, therefore the same non-linear regression analysis as above was performed in cardio disease Top 5 target SNPs, HLA-DPA1 tv3, SLC22A3, LAMC2 tv1, ABO (Venous thromboembolism; yellow) and ABO (Coronary heart disease; blue) (Figure 4E). Other 11 phenotypes, aging (A), bone and muscle (B), gastrointestinal disturbance (C), immune diseases (D), infection (E), lifestyle-related diseases (F), neurodegeneration (G), renal dysfunction (H), respiratory diseases (I), skin diseases (J) and tumors (K), were also shown in Figure S4(A-K). In Figure 4F, the average of each phenotype (above 12 types) was calculated. These data shows that EFTSCs was related with DNSCs in non-linear regression and miRNA/miRNA multi-interaction on this algorithmic function could specifically discern among each disease phenotype from one to another without CSCs.

| Non-linear relation between EFTSC and DNSC |                          |              |              |              |                            |                          |              |              |              |
|--------------------------------------------|--------------------------|--------------|--------------|--------------|----------------------------|--------------------------|--------------|--------------|--------------|
| Phenotype                                  | Top 5 of linear-relation | x=4          | x=5          | x=6          | Phenotype                  | Top 5 of linear-relation | x=4          | x=5          | x=6          |
| Aging                                      | NM_000168                | 0.012058168  | -0.016470248 | 0.007452843  | Lifestyle-related diseases | NM_006193                | -0.015756313 | -0.014113215 | 0.018592863  |
|                                            | NM_000242                | 0.001314196  | -0.022492015 | -0.015682442 |                            | NM_002670                | -0.104688423 | 0.27070867   | 0.528960572  |
|                                            | NM_001099667             | -0.098859574 | 0.021129359  | -0.138909618 |                            | NM_005921                | -0.635503095 | -23.81224702 | -21.57072988 |
|                                            | NM_033064                | -0.002926991 | 0.013081728  | -0.000290015 |                            | NM_130438                | 0.795029009  | 0.069444085  | 0.178751835  |
|                                            | NM_174963                | -0.00299472  | 0.031559759  | 0.001035805  |                            | NM_006023                | -0.222866552 | -17.48658595 | -6.032304852 |
| Bone and Muscle                            | NM_001160047             | -1.836936013 | -17.82369986 | 0.694816778  | Neurodegeneration          | NM_014510                | 0.044623465  | 0.076935488  | 0.034141391  |
|                                            | NM_002699                | -1.18033182  | -1.967665348 | 83.22938139  |                            | NM_001256026             | 0            | -0.024120577 | -0.029340082 |
|                                            | NM_144701                | -6.253676602 | -9.691228596 | -13.89262309 |                            | NM_001134285             | 0.033145629  | -0.166328428 | 0.014796891  |
|                                            | NM_000844                | -0.04491237  | -0.127543572 | 0.061402831  |                            | NM_001199214             | 0.033730152  | -0.024916436 | 0.020727983  |
|                                            | NM_013372                | -0.042339429 | -0.020033218 | -0.029586741 |                            | NM_000844                | -0.04491237  | -0.127543572 | 0.061402831  |
| Cardio disease                             | NM_001242525             | 0.138769773  | 0.230265412  | 0.210840898  | Renal dysfunction          | NM_000099                | 0.002745566  | 0.012823723  | -0.001536632 |
|                                            | NM_021977                | -0.058052944 | -0.050262614 | -0.055614456 |                            | NM_001008693             | 0.010906843  | 0.014083217  | 0.040142249  |
|                                            | NM_005562                | -13.81923458 | -47.35860291 | -3.500206743 |                            | NM_001242525             | -0.023421473 | -0.019529767 | 0.011764977  |
|                                            | NM_020469                | -0.00297803  | -0.001560835 | 0.001826529  |                            | NM_198098                | 0.047943309  | -0.07178538  | -0.332776355 |
|                                            | NM_020469                | -0.002849478 | -0.00119117  | 0.00207463   |                            | NM_001134404             | 0.002085085  | 0.006446119  | 0.010796594  |
| gastrointestinal disturbance               | NM_003121                | -0.06990877  | -0.03076629  | -0.010936752 | Respiratory                | NM_000744                | -0.002082209 | 0.001906862  | -0.001311065 |
|                                            | NM_001142615             | -0.13318713  | -0.005935339 | -0.019596216 |                            | NM_006011                | -0.000873471 | -0.006073873 | -0.003134019 |
|                                            | NM_001131027             | 0.009354864  | -0.010909834 | 0.003936925  |                            | NM_182899                | 3.53877E-05  | -0.001849748 | 0.001995308  |
|                                            | NM_001131027             | 0.009354864  | -0.010909834 | 0.003936925  |                            | NM_004136                | -0.065039109 | 0.10865928   | 0.223369783  |
|                                            | NM_144701                | -92.93760687 | 204.0269602  | 64.58128544  |                            | NM_006947                | 0.011580779  | 0.010562821  | 0.010230871  |
| Immune disease                             | NM_001242462             | -0.031505294 | -0.008264728 | -0.033645037 | Skin                       | NM_144701                | -92.93760687 | 204.0269602  | 64.58128544  |
|                                            | NM_000369                | 0.005137557  | 0.002821046  | -0.005253159 |                            | NM_001242525             | -0.023421473 | -0.019529767 | 0.011764977  |
|                                            | NM_145648                | -0.004850696 | -0.005382945 | -0.005782403 |                            | NM_002187                | -36.21218608 | -22.48218207 | -10.73120793 |
|                                            | NM_144701                | -6.253676602 | -9.691228596 | -13.89262309 |                            | NM_052836                | -0.170209037 | -0.084559706 | 0.049496433  |
|                                            | NM_000168                | 0.012058168  | -0.016470248 | 0.007452843  |                            | NM_017637                | 0.001071423  | -0.009816768 | -0.009148302 |
| Infection                                  | NM_025236                | 0.001476273  | -0.007097089 | 0.005720006  | Tumor                      | NM_022841                | -0.115799996 | -0.016097043 | 0.198224634  |
|                                            | NM_144701                | -92.93760687 | 204.0269602  | 64.58128544  |                            | NM_001142615             | -0.13318713  | -0.005935339 | -0.019596216 |
|                                            | NM_001242525             | -0.023421473 | -0.019529767 | 0.011764977  |                            | NM_005602                | -12.37588879 | -3.373459446 | -1.087459373 |
|                                            | NM_020469                | -0.00297803  | -0.001560835 | 0.001826529  |                            | NM_018833                | -0.010531302 | 0.007146861  | -0.004657371 |
|                                            | NM_004589                | 0.022963287  | -0.047169311 | 0.015660523  |                            | NM_005921                | -7.164133989 | 8.330129711  | 6.649582876  |

derived  
x shows number of G derived from DNS

the maximum (red) and the minimum (blue) values of average of EFTSCs were represented.

Table 3. EFTSC of Top 5 Correlation between CSC and DNSC in Each Phenotype



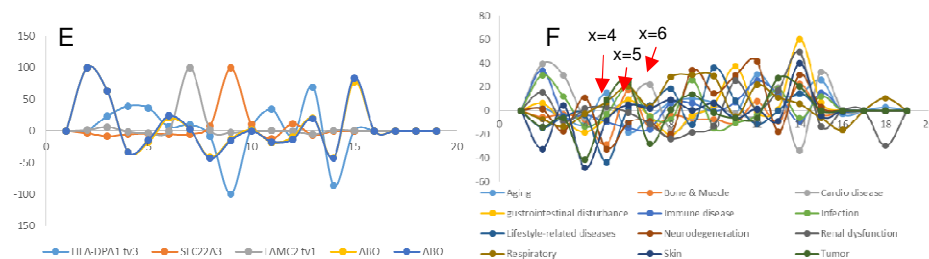


**Figure 4. Fourier Transformation in EFTSC**

Each figure should include explanation, details, and explanations if read solely without reference to the text.

#### 4. DISCUSSION

Recently, miRNA regulatory networks and miRNA-related SNPs have been studied intensely to understand their role in human diseases[19, 20]. A GWAS was conducted to reveal associations between miRNAs and diseases[21]. In this study, we used PolymiRTS, which contains information about miRNAs and polymorphisms in their targets. PolymiRTS connects numerous diseases with SNPs in pre-miRNAs, pri-miRNAs, miRNA promoters and miRNA targets[13]. SNPs in the miRNA genes and in their 3' UTR mRNA targets are largely diverse among individuals, and cause various phenotypes and diseases. In our model experiments, we used the PolymiRTS database, because according to GWAS, SNPs in target sites are associated with diseases. Role of these genetic traits has been confirmed using luciferase-based gain- and loss-of-function assays *in vitro*[19, 22], by comparative analysis of the human genome, and a network-based experiment has recently revealed the target miRNAs (miR-943 and miR-571) for the therapy of hepatocellular carcinoma[23]. Furthermore, based on a functional assay using a miRNA lentivector library, the tamoxifen responsive miRNA has been identified as target in breast cancer using the genome-integrated pre-miRNA forced expression assay[24]. This suggests that the role of SNPs in the 3' UTR are identified with gain- and loss-of-function models in humans, and multiple parameters of miRNAs and their targets are essential in the diagnosis, treatment and prognosis of diseases. In addition, SNPs in miRNA targets might predict miRNA function; therefore, pharmacogenomics and molecular epidemiology can utilize these results. GWAS revealed disease-related genes, which are related or not related to various biological pathways and mechanisms. We hypothesized that the latter case could be explained with miRNA/miRNA interactions, because almost of all biological events are controlled by miRNAs. For instance, there is a strong association between DNS and CSC values of GRM7 of Parkinson's disease. GRM7 can modulate neurotransmitter release and neuronal excitability. Furthermore, GRM7 controls neural differentiation via the phosphorylation of CREB and regulation of YAP expression[25]. Therefore, according to GWAS, GRM7 is linked to bipolar[26], depression[27], hyperactive disorder[28], schizophrenia[29] and age-related hearing loss[30]. On the contrary, ESRRG is related to neurodegeneration; however, according to the literature, ESRRG is mainly implicated in the energy-balance of the whole body[31], and function of ESRRG in neurotransmission remains unknown. In the case of this GWAS, our data shows that ESRRG might function via miRNA/miRNA multiple interactions. Thus, our findings suggest that miRNA/miRNA multiple interactions are crucial in human diseases and sufficient for the prediction of diagnosis and prognosis (see Figure 6).

Torus RNA could be advantageous for biological prevalence because it can serve as template for rolling circle replication and be stable due to be resistant to exonucleases[32]. In our experiments, it can be simulated that torus model of miRNA can transmit genetic information among miRNAs therefore, the circular type of RNA might have an important role for its function to architect the mechanisms of tuning of miRNA. Quite recently, it has been reported that natural circular RNA were identified in human cells and the circular RNA (1,400 nts) can bind to miR-7 as the sponge[33]. Further Y RNA is known as a family of non-coding and a natural circular RNA[34], and human Y3 and Y5 (25 nts

each) were annotated as miR-1979 and miR-1975 in miRBase, respectively. Artificially, the repetitive circular RNAs (approximate 14 nts) are catalyzed by polymerase from ribozymes and 34 nts RNA were made as circular RNA by using a template-directed non-enzymatic ligation[35]. These results indicated that natural mobile torus miRNA could be bio-generated and have a function to transmit gene information. Further, circular RNA has been related with cancer[36]. Intriguingly, above miR-1979 and miR-1975 are circulating miRNA, and as biomarkers, miR-1979 is associated with multiple sclerosis[37] and chronic congestive heart failure[38]. In addition, miR-1975 is implicated in liver cancer[39], suggesting that the circulating circular miRNA genes would have some function to induce human diseases, therefore, DNSC and EFTS scorings under tours condition are suitable for human disease prediction to be relevant to GWAS. We classified human diseases with algorithms based on electrodynamics and quantum energy of miRNA nucleotides. EFTS based on electrodynamics corresponds to nucleotide changes in miRNAs, such as keto-enol tautomerism[40], changes induced by environmental stress, including ultraviolet rays[41] and A/I RNA editing[42]. On the contrary, DNS corresponds to the presence of G bases in miRNAs but not tautomerism. Guanine generate G-quartets[43] or scale of the fish caused by changes in the magnetic field[44]. In addition, miRNA SNPs changed to guanine are more effective[45]. Accordingly, G-dependent miRNA/miRNA interactions, based on the miR-ket code, are important factors in the association of miRNAs and human diseases[46]. While CSC and EFTS were not related, CSC correlated with DNS. Therefore, guanine bases in miRNAs are often associated with disease-related polymorphisms. Furthermore, in the targets, approximately 40% of DNS values have a linear relationship with CSC. Therefore, DNS might be a useful tool to remove false positive data from a large miRNA dataset in targets.

As a miRNA contain approximately 22 nucleotides, the DNS value can only have 276 varieties. However, corresponding to Watson-Crick base pair and wobble base pair, the EFTS value can have  $2^{131} + 2^{65}$  varieties. We found that EFTS convergently has a non-linear relationship with DNS. The non-linear regression analysis with a Fourier transformation showed that each disease can be represented with a specific non-linear curve. Because in classical mechanics all physical events can be described with specific wave properties using quantum mechanics, our results would be rational to physical criteria. Thus, EFTS (related to classical mechanics) and DNS (related to quantum mechanics) had a non-linear relationship, suggesting that each disease is visualized using specific wave properties. In addition, using non-linear curve without CSC, phenotype of human diseases can be identified. Therefore, human diseases may be predicted by the miRNA/miRNA interaction only.

As the calculation of EFTS and DNS values is still problematic, further research is necessary. Firstly, when we set EFTS, atoms that form a Watson-Crick base pair[47] or wobble base pair[48] were selected as active ones. However, nucleic acids can also form a Hoogsteen base pair[49]. Hoogsteen base pairs are related to the G-quartet structure and Fragile X mental retardation protein (FMRP) nucleic acid chaperone has facilitated miRNA assembly involving recognition of G-quartet structures on target mRNAs[50]; therefore, EFTS calculations have to be set up for Hoogsteen base pairs. Secondly, although we presented the relationship between miRNA/miRNA interactions and diseases based on GWAS, the expression levels of miRNAs were not considered, because miRNA-mRNA binding sites were identified with direct mapping experiments such as cross linking, ligation and sequencing of hybrids (CLASH)[51]. Therefore, integration of comparative miRNA qualitative expression levels might improve the performance of the model. Further progression was necessary for precision scoring.

399  
400

## 401 5. CONCLUSION

402

403 MiRNA genes are key regulators of almost all cellular processes in various species and  
404 humans. The miRNAs are implicated in the direct or indirect transcriptional or  
405 post-transcriptional regulation of gene expression. For the post-transcriptional  
406 suppression of gene expression, short seed region of miRNAs is recruited to the 3' UTR

of mRNAs via argonaute (Ago). For the prediction or validation of miRNA binding and function, a complex computational task has been performed using miRNA target prediction tools before validation *in vitro* and *in vivo*. The potential miRNA/target interaction is not a one-to-one correspondence; therefore, translating miRNA/protein gene interactions into disease prognosis and diagnosis is a challenging task. On the other hand, identifying miRNA/miRNA interactions is not an easy task. Beyond this difficult state on the status quo, miRNA/miRNA interactions were integrated to improve prediction accuracy and simplicity. Neo-scores based on the quantum and electrodynamic data of the miRNA/miRNA interaction, DNS and EFTS, respectively, is shown in Figure 5. DNS was correlated with CSC using PolymiRTS data to evaluate the relationship between miRNAs and protein coding genes involved in 150 human diseases. Because EFTS was mathematically functioned into DNS, function of DNS to EFTS was integrated for disease prediction on CSC. Furthermore, in this context, miRNA/miRNA interactions in the algorithmic function might be applied for specific discernment of diseases with the miRNA/target interaction complex. Thus, miRNA/miRNA interactions may have an important role in the *a priori* control of human health. Utility of neo scores is shown in Figure 6. Artificial Intelligence (AI) combined with the EFTS/DNS-based prediction method might help in the diagnosis of diseases. From a technical point of view, our method might improve the prediction of miRNA/miRNA interactions with AI by using more accurate information.

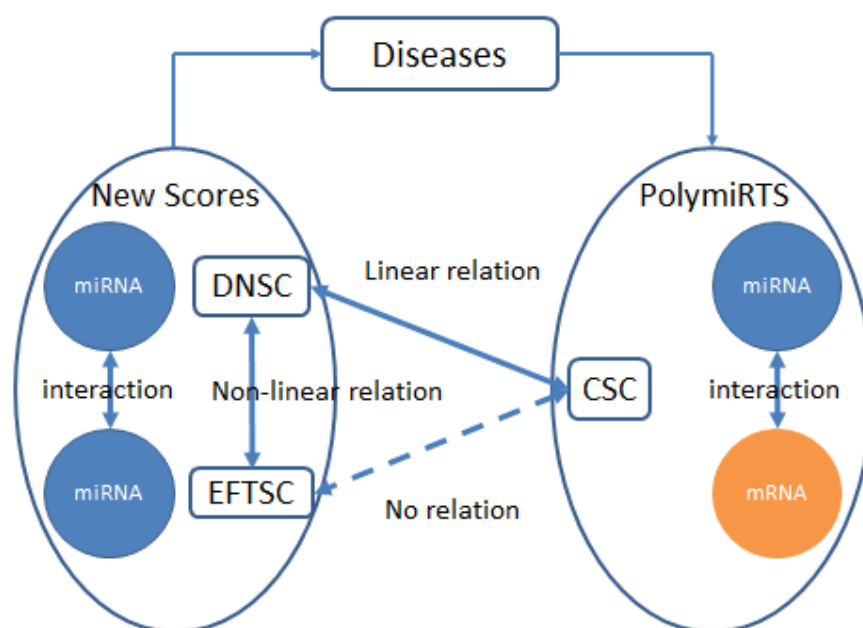


Figure 5. Scheme of Relation between CSC, DNSC and EFTSC

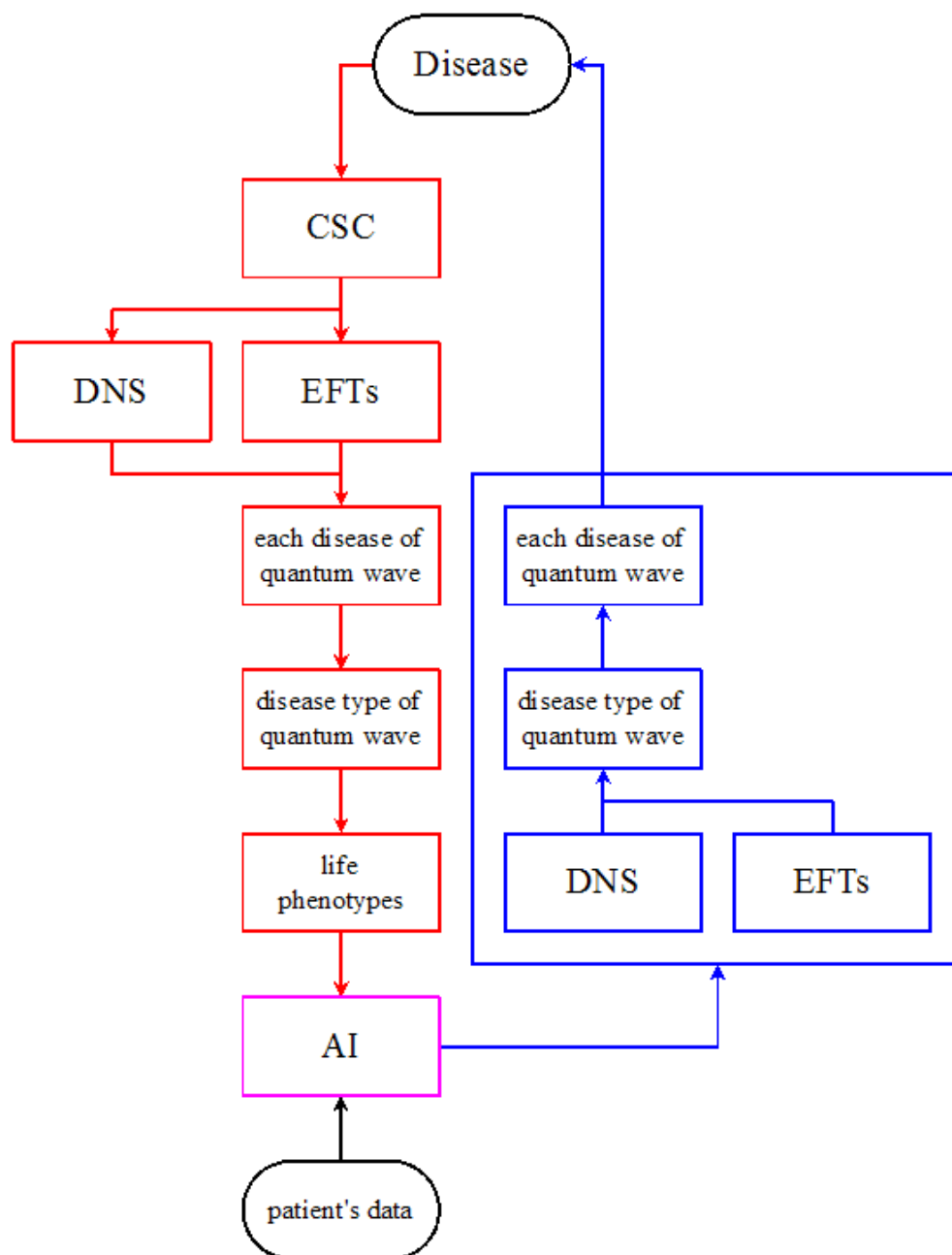


Figure S6. The Future Prospecting Model

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and three table and can be found with this article online at (somewhere; not available now).

## REFERENCES AND NOTES

443

- 444 1. C. Yang et al. Tag SNPs in long non-coding RNA H19 contribute to susceptibility  
445 to gastric cancer in the chinese han population. *Oncotarget*. 2015.
- 446 2. M. S. Ebert, P. a. Sharp Roles for MicroRNAs in conferring robustness to  
447 biological processes. *Cell*. 2012;149:505–524.
- 448 3. V. Ambros, The functions of animal microRNAs. *Nature*. 2004;431:350–355.
- 449 4. I. L. Jung et al., Cesium Toxicity Alters MicroRNA Processing and AGO1  
450 Expressions in *Arabidopsis thaliana*. *PLoS One*. 2015;10:e0125514.
- 451 5. C. G. Hill, N. Jabbari, L. V. Matyunina, J. F. McDonald, Functional and  
452 Evolutionary Significance of Human MicroRNA Seed Region Mutations. *PLoS*  
453 *One*. 2014;9:e115241.
- 454 6. H. Kume, K. Hino, J. Galipon, K. Ui-Tei, A-to-I editing in the miRNA seed region  
455 regulates target mRNA selection and silencing efficiency. *Nucleic Acids*  
456 *Res*.2014;42:3–8.
- 457 7. J. T. Granados-Riveron, G. Aquino-Jarquín, Does the linear Sry transcript function  
458 as a ceRNA for miR-138? The sense of antisense. *F1000Research*.  
459 2014;90:2012–2014.
- 460 8. L. Salmena, L. Poliseno, Y. Tay, L. Kats, P. P. Pandolfi, A ceRNA hypothesis: The  
461 rosetta stone of a hidden RNA language? *Cell*. 2011;146:353–358.
- 462 9. J. Yu et al., The microRNA-520a-3p inhibits proliferation , apoptosis and  
463 metastasis by targeting MAP3K2 in non-small cell lung cancer. *Am J Cancer Res*.  
464 2015;5:802–811.
- 465 10. D. M. Garcia et al., Weak Seed-Pairing Stability and High Target-Site Abundance  
466 Decreases the Proficiency of Isy-6 and other miRNA's. *Nat. Struct. Mol.*  
467 *Biol*.2011;18:1139–1146.
- 468 11. S. Griffiths-Jones, R. J. Grocock, S. van Dongen, A. Bateman, A. J. Enright,  
469 miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids*  
470 *Res*.2006;34:D140–D144.
- 471 12. E. M. Smigielski, K. Sirotkin, M. Ward, S. T. Sherry, dbSNP: a database of single  
472 nucleotide polymorphisms. *Nucleic Acids Res*.2000;28:352–355.
- 473 13. A. Bhattacharya, J. D. Ziebarth, Y. Cui, PolymiRTS Database 3.0: Linking  
474 polymorphisms in microRNAs and their target sites with human diseases and  
475 biological pathways. *Nucleic Acids Res*.2014;42:86–91.
- 476 14. C. Delay et al., MicroRNAs targeting Nicastrin regulate A $\beta$ <sup>2</sup> production and are  
477 affected by target site polymorphisms. *Front. Mol. Neurosci*.2014;7:1–7.
- 478 15. M. Jia et al., Computational analysis of functional single nucleotide polymorphisms  
479 associated with the CYP11B2 gene. *PLoS One*. 2014;9

- 480 16. Y. R. Fujii, RNA Wave for the HIV Therapy : Foods , Stem Cells and the RNA  
481 Information Gene. World J. AIDS. 2013:131–146.
- 482 17. Y. R. Fujii, The RNA Gene Information: Retroelement-MicroRNA Entangling as  
483 the RNA Quantum Code. Methods Mol. Biol.2013:936:47–67.
- 484 18. Y. R. Fujii, RNA Genes: Retroelements and Virally Retroposable microRNAs in  
485 Human Embryonic Stem Cells. Open Virol. J.2010;4:63–75.
- 486 19. L. J. Chin et al., A SNP in a let-7 microRNA complementary site in the KRAS  
487 3'UTR Increases Non-Small Cell Lung Cancer Risk. Cancer Res.  
488 2008;68:8535–8540.
- 489 20. C. L. Plaisier, M. Pan, N. S. Baliga, A miRNA-regulatory network explains how  
490 dysregulated miRNAs perturb oncogenic processes across diverse cancers.  
491 Genome Res.2012;22:2302–2314.
- 492 21. L. F. Goulart et al., MicroRNAs enrichment in GWAS of complex human  
493 phenotypes. BMC Genomics. 2015:16.
- 494 22. J. Gong et al., Genome-wide identification of SNPs in MicroRNA genes and the  
495 SNP effects on MicroRNA target binding and biogenesis. Hum.  
496 Mutat.2012;33:254–263.
- 497 23. F. Yu, X.-Y. Shen, L. Fan, Z.-C. Yu, Genome-wide analysis of genetic variations  
498 assisted by Ingenuity Pathway Analysis to comprehensively investigate potential  
499 genetic targets associated with the progression of hepatocellular carcinoma. Eur.  
500 Rev. Med. Pharmacol. Sci.2014;18:2102–2108.
- 501 24. T. Ujihira et al., MicroRNA-574-3p, identified by microRNA library-based functional  
502 screening, modulates tamoxifen response in breast cancer. Sci. Rep.2015;5:7641.
- 503 25. W. Xia, Y. Liu, J. Jiao, GRM7 Regulates Embryonic Neurogenesis via CREB and  
504 YAP. Stem Cell Reports. 2015;4:1–16.
- 505 26. N. Alliey-Rodriguez et al., Genome-wide association study of personality traits in  
506 bipolar patients. Psychiatr. Genet.2011;21:190–194.
- 507 27. S. I. Shyn, S. P. Hamilton, The Genetics of major depression: Moving beyond the  
508 monoamine hypothesis. Psychiatr. Clin. North Am.2010;33:125–140.
- 509 28. J. Elia et al., Genome-wide copy number variation study associates metabotropic  
510 glutamate receptor gene networks with attention deficit hyperactivity disorder. Nat.  
511 Genet.2012;44:78–84.
- 512 29. K. H. Choi, M. E. Zepp, B. W. Higgs, C. S. Weickert, M. J. Webster, Expression  
513 profiles of schizophrenia susceptibility genes during human prefrontal cortical  
514 development. J. Psychiatry Neurosci.2009;34:450–458.
- 515 30. R. a. Friedman et al., GRM7 variants confer susceptibility to age-related hearing  
516 impairment. Hum. Mol. Genet.2009;18:785–796.

- 517 31. M. S. Byerly et al., Estrogen-related receptor  $\beta$  deletion modulates whole-body  
518 energy balance via estrogen-related receptor  $\gamma$  and attenuates neuropeptide Y  
519 gene expression. *Eur. J. Neurosci.*2013;37:1033–1047.
- 520 32. E. Lasda, R. O. Y. Parker, Circular RNAs : diversity of form and function. *RNA*.  
521 2015;20:1829–1842.
- 522 33. T. B. Hansen, J. Kjems, C. K. Damgaard, Circular RNA and miR-7 in Cancer.  
523 *Cancer Res.*2013;73:5609–5612.
- 524 34. F. E. Nicolas, A. E. Hall, T. Csorba, C. Turnbull, T. Dalmay, Biogenesis of  $\gamma$   
525 RNA-derived small RNAs is independent of the microRNA pathway. *FEBS*  
526 *Lett.*2012;586:1226–1230.
- 527 35. A. M. Diegelman, E. T. Kool, Generation of circular RNAs and trans-cleaving  
528 catalytic RNAs by rolling transcription of circular DNA oligonucleotides encoding  
529 hairpin ribozymes. *Nucleic Acids Res.*1998;26:3235–3241.
- 530 36. T. B. Hansen et al., Natural RNA circles function as efficient microRNA sponges.  
531 *Nature*. 2013;495:384–8.
- 532 37. S. R. Siege, J. MacKenzie, G. Chaplin, N. G. Jablonski, L. Griffiths, Circulating  
533 microRNAs involved in multiple sclerosis. *Mol. Biol. Rep.*2012;39:6219–6225.
- 534 38. H. A. Cakmak et al., The prognostic value of circulating microRNAs in heart  
535 failure : preliminary results from a genome-wide expression study . *J Cardiovasc*  
536 *Med*. 2015;16.
- 537 39. L. Lin, Y. Lin, Y. Jin, C. Zheng, Microarray analysis of microRNA expression in  
538 liver cancer tissues and normal control. *Gene*. 2013;523:158–160.
- 539 40. K. Szczepaniak, M. Szczesniak, Matrix isolation infrared studies of nucleic acid  
540 constituents : Part 4. Guanine and 9-methylguanine monomers and their  
541 keto—enol tautomerism. 1987;156:29–42.
- 542 41. J. Hu, S. Adar, C. P. Selby, J. D. Lieb, A. Sancar, Genome-wide analysis of  
543 human global and transcription-coupled excision repair of UV damage at  
544 single-nucleotide resolution. *GENES Dev.*2015;29:48–96.
- 545 42. N. Paz-Yaacov et al., Adenosine-to-inosine RNA editing shapes transcriptome  
546 diversity in primates. *Proc. Natl. Acad. Sci. U. S. A.*2010;107:12174–12179.
- 547 43. J. R. Williamson, M. K. Raghuraman, T. R. Cech, Monovalent cation-induced  
548 structure of telomeric DNA: The G-quartet model. *Cell*. 1989;59:871–880.
- 549 44. M. Iwasaka, Y. Mizukawa, Light reflection control in biogenic micro-mirror by  
550 diamagnetic orientation. *Langmuir*. 2013;29:4328–4334.
- 551 45. Z. Wang et al., Association between miR-27a genetic variants and susceptibility to  
552 colorectal cancer. *Tumor Biol.*2014;35:2151–2156.
- 553 46. Y. R. Fujii, Formulation of New Algorithmics for miRNAs. *Open Virol*.  
554 *J.*2008;2:37–43.

555 47. J. D. Watson, F. H. C. Crick, Molecular structure of nucleic acids. *Nature*.  
556 1953;171:737–738.

557 48. F. H. Crick, Codon--anticodon pairing: the wobble hypothesis. *J. Mol.*  
558 *Biol.*1966;19:548–555.

559 49. K. Hoogsteen, The crystal and molecular structure of a hydrogen-bonded complex  
560 between 1-methylthymine and 9-methyladenine. *Acta Cryst.*1963;16, 907–916.

561 50. I. Plante, P. Provost, Hypothesis: A role for fragile X mental retardation protein in  
562 mediating and relieving microRNA-guided translational repression? *J. Biomed.*  
563 *Biotechnol.*2006;2006:1–7.

564 51. G. Kudla, S. Granneman, D. Hahn, J. D. Beggs, D. Tollervey, Cross-linking,  
565 ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc.*  
566 *Natl. Acad. Sci. U. S. A.*2011;108:10010–10015.

567