USING BIOINFORMATICS TO PREDICT POTENTIAL TARGETS OF MICRORNA-144 IN OSTEOARTHRITIS

HO THI BICH PHUONG, VIEN NGOC THACH
Ho Chi Minh City Open University, Vietnam

LUONG HOANG NGAN, LE THI TRUC LINH
Ho Chi Minh City Open University, Vietnam – linh.ltt@ou.edu.vn

ABSTRACT

MicroRNAs are short endogenous non-coding RNA molecules, typically 19-25 nucleotides in length, which negatively regulate gene expression through binding to 3'UTR of target mRNAs, leading to repression of protein translation or target mRNA degradation. MicroRNA-144 (miR-144) was found as an abnormal expression in various diseases, including osteoarthritis (OA). We have identified increased microRNA-144 expression in early phase and end stage of OA. However, the molecular mechanism of this increase has not been yet to be determined yet. Using bioinformatics tools, we found more than 4,000 mRNAs that are predicted to be potential direct targets of miR-144, including mRNAs involved in the critical signaling pathways in OA e.g. TGFβ/Smad2/3 and WNT/β-catenin. Results from this research provide information for future ex periments to validate miR-144 potential targets.

Keywords: Bioinformatics; MicroRNA-144; Osteoarthritis.

1. Introduction

MicroRNAs (miR, miRNA) are an abundant class of short (~20 to 25nt long) RNAs that have emerged as important post-transcriptional regulators of gene expression by binding to specific sequences within a target mRNA (Ambros, 2004; Bartel, 2004). 4,552 miRNAs have been identified in human cells (miRbase: www.mirbase.org, 2014) and each miRNA is predicted to regulate several target genes (Kozomara & Griffiths-Jones, 2011; Lim et al., 2005). Computational predictions indicate that more than 50% of all human protein-coding genes are potentially regulated by miRNAs (Chen et al., 2005; Lewis, Burge, & Bartel, 2005). A significant number of miRNAs are physiologically expressed and regulated in many diseases. Therefore, studying their direct targets and insight into mechanisms help to understand their functions.

It is now considered that OA is a disease of the whole joint as an organ where all majors components of the joint, e.g. the cartilage, the synovium and the underlying bone, are affected (Loeser, Goldring, Scanzello, & Goldring, 2012). The pathologic changes seen in OA include cartilage destruction, hyperplasia and fibrosis of the synovium, osteophyte formation, and subchondral bone thickening (Aigner, Sachse, Gebhard, & Roach, 2006; Loeser et al., 2012). These changes were result of incompletely understood series of functional events.

The involvement of miRNA expression and OA was evident through studies determining the involvement of miR-140 in OA. MiR-140 expression in the cartilage of patients with OA was significantly lower than in normal cartilage (Miyaki et al., 2009; Tardif, Hum, Pelletier, Duval, & Martel-Pelletier, 2009); also, decreased miR-140 expression was reported in OA chondrocytes (Tardif et al., 2009). Deletion of miR-140 in mice predisposed them to the development of age-related OA-like changes and increased cartilage destruction in surgically-induced OA (Miyaki et al., 2010). Conversely, in an antigen-induced arthritis model, transgenic over-expression of miR-140 in chondrocytes

protected against cartilage damage (Nakamura, Inloes, Katagiri, & Kobayashi, 2011). These data implicate that for cartilage homeostasis, the expression of miR-140 must be strictly controlled.

The identification of miRNAs' target mRNAs is a key factor in elucidating the function of miRNAs, yet this step has proven computationally difficult due to the complexity of the miRNA–target interactions. Several target prediction programs have been developed, but the overlap between sets of predicted target genes for a given miRNA by different programs is surprisingly low (Sethupathy, Megraw, & Hatzigeorgiou, 2006), suggesting that there are a number of false positive predictions. Genes that are targeted by miRNAs but not predicted by the programs are false negatives, and it is not known how many there could be. Nevertheless, these target prediction programs are very useful to define potential targets that can be validated experimentally, often using luciferase reporter assays (Lewis et al., 2003).

Another strategy for identifying targets is based on experimental data and usually involves the manipulation of miRNA activity. The first such study showed that overexpression of miR-1 and miR-124 in cells where these miRNAs are not normally expressed led to repression of many mRNAs (Lim et al., 2005). The opposite has also been demonstrated, in which silencing of miRNA activity resulted in depression of many messages (Krützfeldt et al., 2005; Elmén et al., 2008a,b). Recently, two studies reported biochemical purification of RISC complexes using antibody against the Argonaute-2 protein. miRNA targets were identified by extracting mRNA from the purified complexes and identified by microarray hybridization or sequencing (Karginov et al., 2007; Beitzinger et al., 2007). However, introducing a miRNA into cells where it is normally not expressed and identifying mRNAs targeted by the miRNA has limited

biological relevance.

2. Materials and Methods
Sequence data

The miR-144-3p mature sequence was retrieved from miRbase database (http://www.mirbase.org/). 3'UTR sequences were downloaded from UCSC (https://genome.ucsc.edu/) and Ensembl (http://www.ensembl.org/index.html). RefSeq IDs was used to map probe sets to UCSC database and Ensembl Gene IDs were used to map probesets to the Ensembl database.

MicroRNA-144-3p target prediction

Three types of seed matches in the 3'UTR were considered when predicting direct miRNA-144 targets e.g. 6-mer seed match which is 6nt in length and was complementary to nucleotides 2 to 7 in the miR-29 family; 7-mer seed match which is 7nt length and is complementary to nucleotides 1–7 in the miRNA or nucleotides 2–7 in the miRNA with "A" at the first position; and 8-mer seed match which is 8nt length, and matched nucleotides 1–8 in the miRNA or nucleotides 2–8 in the miRNA with an "A" at the first position. For searching these seed matches in the 3'UTR, 3'UTR sequences were imported and read in R studios using the *"readDNAStringSet"* function in Biostring package. Also, miR-144-3p seed matches were searched using *"vcountPattern"* function.

In line with using R studios, some miRNA target prediction programs available were also used to predict targets for miR-144 including TargetScan (http://www.targetscan.org/), miRNA body map (http://www.mirnabodymap.org/), miRDB (http://mirdb.org/miRDB/), DIANA (http://diana.cslab.ece.ntua.gr/), Pictar (http://pictar.mdc-berlin.de/), miRbase (http://www.mirbase.org/).

Functional pathway analysis

DAVID (Database for Annotation, Visualization and Integrated Discovery) functional annotation tool (http://david.abcc.

ncifcrf.gov/) was used to perform functional analysis for particular gene groups.

3. Results

miR-144 mature sequence

In human genome, microRNA 144 (miR-144) is located in negative strand in intergenic region of chromosome 17 at 17q11. MiR-144 is cluttered with miR-451 and miR-4732, which are hosted in regions around 94-bp miR-144 downstream and 37-bp, respectively (miRbase, www.mirbase.org) (Figure 1). In mouse, however, the intergenic miRNA-144 is encoded in positive strand of chromosome 11 and clustered with only miR-451 (miRbase, www.mirbase.org).

Figure 1. MicroRNA 144 genome

(A) Human miRNA-144 genome region. MicroRNA 144 is located in chromosome 17 together with miR-451; (B) miR-144 precursor in which mature sequences are in red

MicroRNA 144 mature conserves across species

When comparing the human miR-144 pre-mature sequence to its homologs in the genomes of 31 annotated species, both mature miR-144 sequences, termed miR-144-3p (or miR-144) and miR-144-5p (or miR-144*), are conserved in many species (Figure 2).

Figure 2. MicroRNA 144 is conserved across species

MicroRNA 144 precursor sequences of 31 species were downloaded and compared. The star (*) represents nucleotides are similar in all species.

Potential targets of miRNA 144

All 42580 3'UTR of mRNA are downloaded from Ensemble. Three types of seed matches, including 6-mer, 7-mer, 8-mer seed matches, were searched in all 3'UTR regions. Percentage of all mRNAs containing 6-nt seed matches, 7-nt seed matches, 8-nt seed matches are 57% (42385), 21% (8551), 6% (2412) respectively (Figure 3).
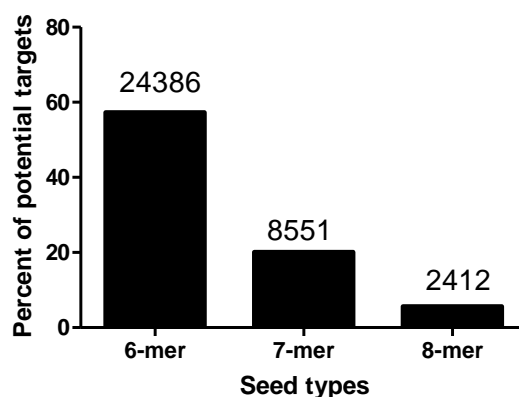
**Figure 3.** Percentage of mRNA containing seed matches of miR-144

Three types of seed matches e.g 6-mer, 7-mer, 8-mer seed matches were searched in 42580 sequences of 3'UTR. All counts for a specific seed match in a 3'UTR was attributed to 1. Number of 3'UTR sequences containing each type of seed matches are in each column.

Interestingly, when counting number of seed matches in each 3'UTR sequence, we found that 3'UTRs containing just only one seed match which accounted for the most and followed by ones containing two and three seed matches. This is true for all three seed match types (Figure 4). There are several 3'UTRs with high number of seed matches. For 6-mer match, gene *SEMA4F, SECISBP2, SLC22A2, SDF4, PURA, STEAP4, ZDHHC21, ATP8B2, NAGA, QK, ZFP26, PPP2CA, ATXN1, MYO9A* having at least 5 6-mer matches. For 7-mer match, *SECISBP2, MDP1 GNB4, PURA* having at least 5 7-mer matches. For 8-mer match, gene *SECISBP2* has 15 8-mer match.

We compared the list of miR-144 candidate mRNAs with miR-144 targets predicted by TargetScan, MiRBase, and PicTar. Surprisingly, a significant number of miR-144 mRNAs were not predicted by any of the programs, suggesting that these programs generate a significant number of false negatives.
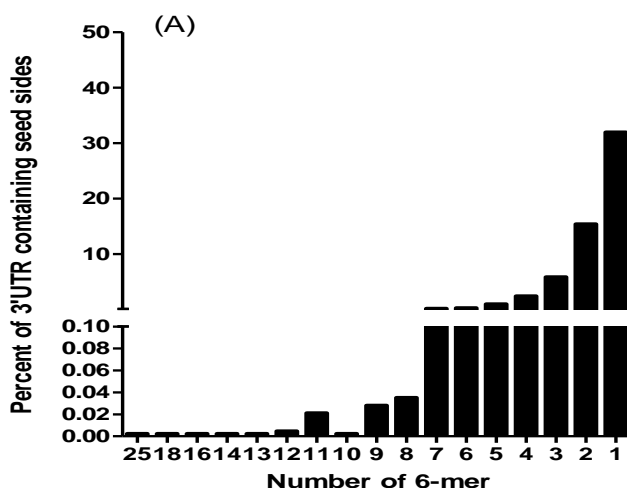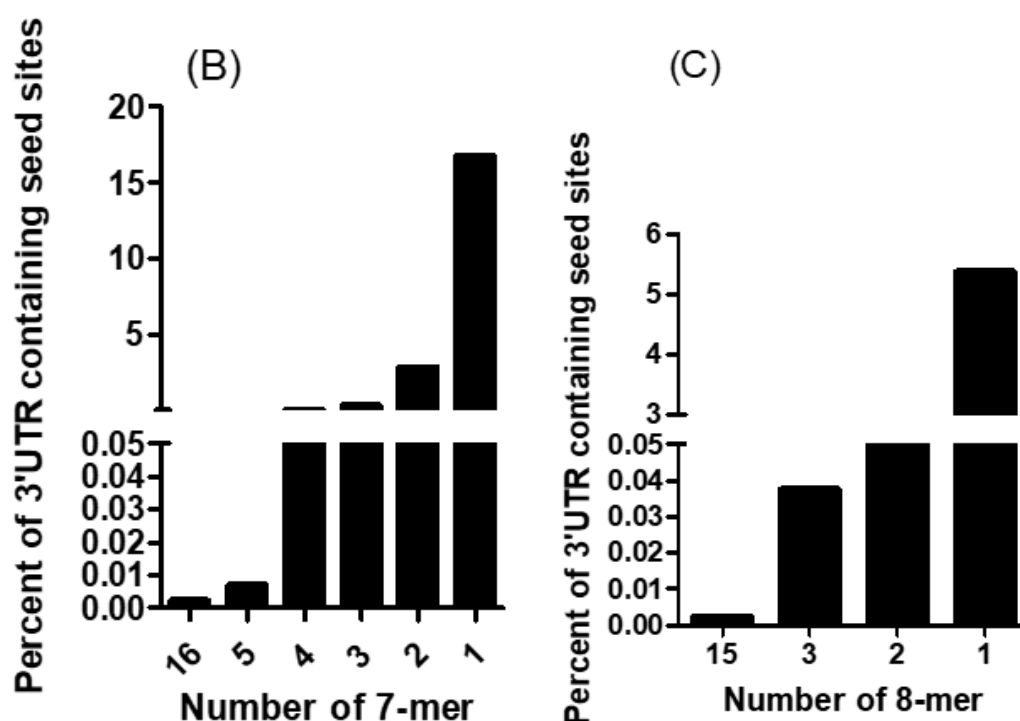
Figure 4. Number of each type of seed match of miR-144 in 3'UTR

Functions of microRNA 144 in Osteoarthritis

In order to evaluate the global effect of miR-144 on the whole genome, miR-144 potential targets, which were identified above, were put in DAVID functional annotation software to identify the most represented gene ontology (GO) categories. We found the enrichment for the Wnt signaling pathway together with MAPK kinase signaling pathway and TGFβ pathway (Figure 5, 6, 7).

Data from DAVID functional annotation tool showed that miR-144 interacted with both canonical and non canonical Wnt signaling pathways at least partly through following potential targets: *APC, CTBP2, CXXC4, FBXW11, ROCK2, VANGL1, AXIN2, CAMK2D, SFRP1, CSNK1A1, DVL1, FZD4, GSK3B, NLK, PRICKLE1, PPP3R1,* *TBL1XR1* (Figure 5).

Similarly, the interaction between miR-144 and both canonical and non canonical TGFβ signalling pathways have been highlighted (Figure 6). Following is miR-144 TGFβ- related potential targets: *CRK, FOS, RAP1A, RAPGEF2, RASGRF1, SOS1, TAOK1, TAB2, CACNA2D1, CACNB2, DUSP1, EGFR, FGF16, FGF7, MAPK1, MAP3K2, MAP3K4, MAP3K8, NLK, PLA2G4A, PPP3R1, PTPN5, STMN1, TGFBR2.*

Another signaling pathway regulated by miR-144 was MAPK signalling (Figure 7). MicroRNA 144 might interact with this signaling partly through controlling expression of *ROCK1, SMAD1, SMAD5, TGIF1, TGIF2, ACVR2B, BMPR1A, BMPR1B, FST, MAPK1, NOG, THBS1, TGFBR2.*
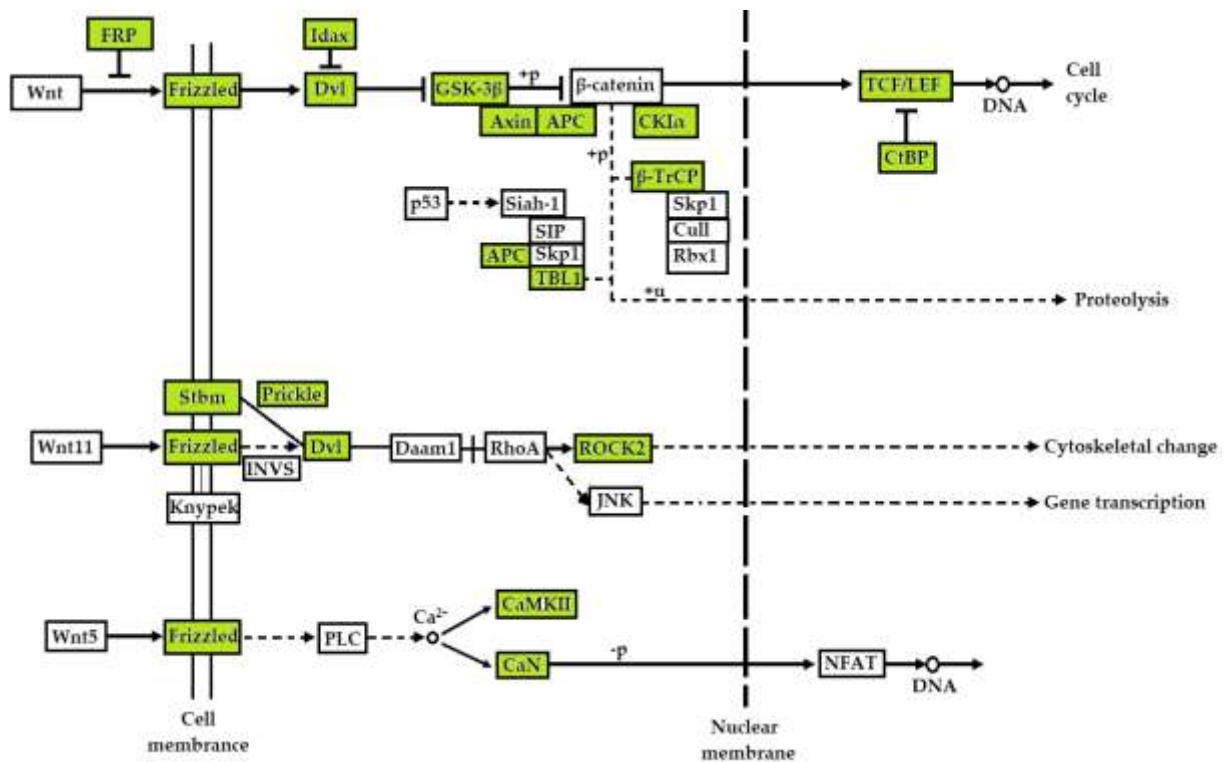
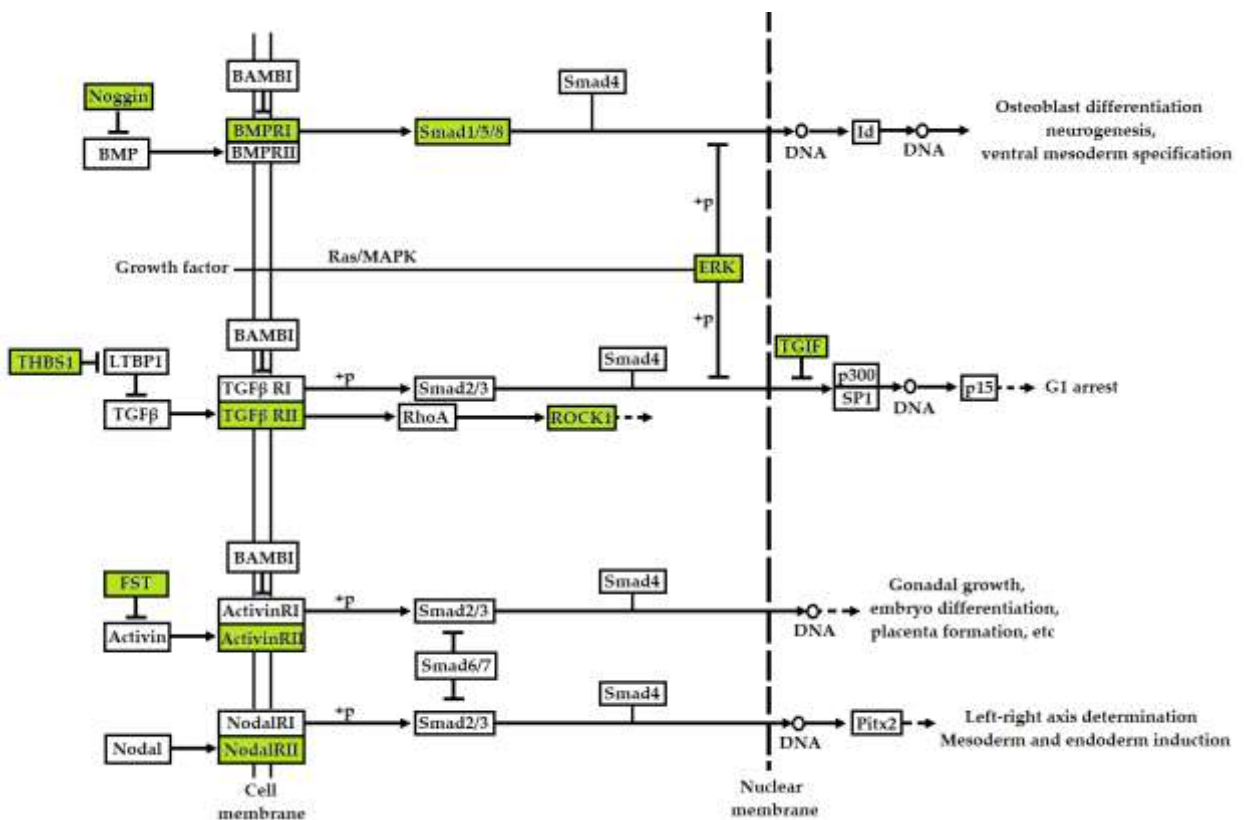**Figure 5.** Wnt signaling pathway. Green boxes indicate the targets of miR-144

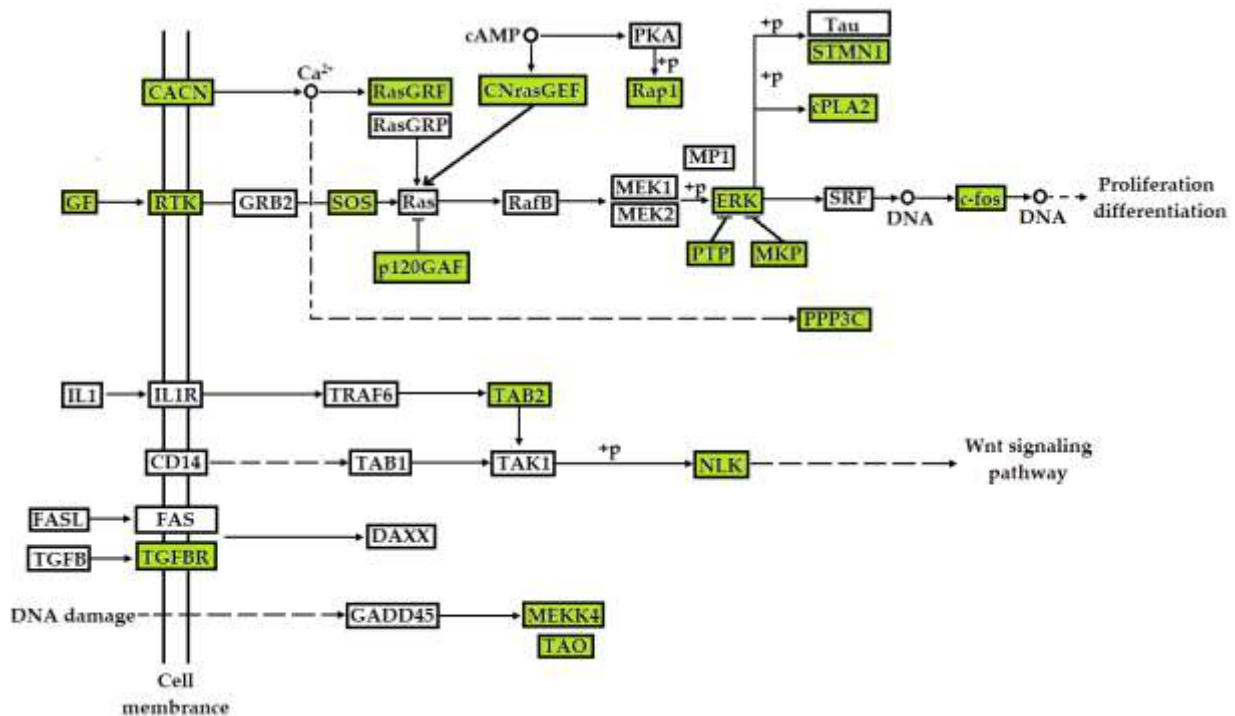**Figure 6.** TGFβ signaling pathway. Green boxes indicate the targets of miR-144

Figure 7. MAPK signaling pathway. Green boxes indicate the targets of miR-144

4. Discussion

The ability of a single miRNA to target multiple mRNAs especially those that function in the same intracellular pathways and/or diseases, adds an additional layer of regulation to gene expression. Moreover, for any potential miRNA therapeutic application, identifying novel miRNA targets and the cell signaling pathways and networks by which miRNAs exert their functions on disease phenotype are therefore of particular importance both to have an insight into OA pathogenesis and also to ensure the specificity in any miRNA-based drug delivery method.

Identifying target mRNAs of miRNAs is an important step, yet this step has proven to be difficult due to the complexity of the miRNA–target interactions. We present here a computational approach to identify miRNA targets that relies on screening for miR seed match on mRNA 3'UTR. This approach reveals a huge number of potential targets but it does not necessary mean that miR-144 will target all of these mRNAs in a specific cell type or tissue at the same time.

In this study, it was found that the miR-144 might have an important role in OA since it could directly regulate signaling cascade of Wnt, TGFβ and MAPK signaling pathway.

The Wnt signaling pathway might be regulated by the miR-144 in which some Wnt signaling related genes were direct targets of the miRNA. Both over-activation and inhibition of Wnt signaling can lead to skeletal deformities and an early onset OA (Zhu et al., 2008; Zhu et al., 2009), illustrating that Wnt signaling needs to be tightly regulated in cartilage homeostasis. However, whether the decreasing of these direct targets is the mechanism for inhibition of the Wnt/β-catenin signaling pathway has not been confirmed in this study.

In line with this, in this study, miR-144 was also found to interact with the TGFβ and MAPK signaling pathway. A number of TGFβ and MAPK signaling pathway components that was miR-144 potential direct targets had been shown. The TGFβ family, consisting of over 35 members including

TGFβ and BMPs, has been widely known to play a crucial role in the development and homeostasis of various tissues. Members of the TGFβ family are considered potent mediators of cartilage matrix synthesis, in which they up-regulate the expression of several types of collagens and proteoglycans but down-regulate cartilage degrading enzymes (Verrecchia, Chu, & Mauviel, 2001; Verrecchia & Mauviel, 2002).

Taken together, all of these data provide good evidence that miR-144 can interact with Wnt, TGFβ and MAPK signaling, at least in part, via repression of these targets. To test this hypothesis, it may be best to perform miR-144 gain-and loss-of function experiment together with whole genome array for all mRNAs. It may also be instructive to perform array experiments in the presence or absence of TGFβ itself.

5. Conclusion

Both mature miR-144 sequences are conserved among various species. Additionally, bioinformatics data shows that the predicted targets of miR-144 are involved in many important signaling pathway in OA e.g. Wnt, TGFβ and MAPK signaling. It suggests that miR-144 not only plays a critical role in cells, but also adheres with process of OA. When bioinformatics-based experiments are carried out, the roles of miR-144 in OA will be clearly demonstrate∎

References

Aigner, T., Sachse, A., Gebhard, P. M., & Roach, H. I. (2006). Osteoarthritis: pathobiology-targets and ways for therapeutic intervention. *Adv Drug Deliv Rev, 58*(2), 128-149. doi: 10.1016/j.addr.2006.01.020

Ambros, V. (2004). The functions of animal microRNAs. *Nature, 431*(7006), 350-355. doi: citeulike-article-id:781405

Bartel, D. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297. doi: citeulike-article-id:504198

Chen, C., Ridzon, D., Broomer, A., Zhou, Z., Lee, D., Nguyen, J., . . . Guegler, K. (2005). Real-time quantification of microRNAs by stem–loop RT–PCR. *Nucleic Acids Research, 33*(20), e179-e179. doi: citeulike-article-id:436781

Kozomara, A., & Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Research, 39*(Database issue), D152-D157. doi: citeulike-article-id:8184232

Lewis, B., Burge, C., & Bartel, D. (2005). Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell, 120*(1), 15-20. doi: citeulike-article-id:455483

Lim, L., Lau, N., Garrett-Engele, P., Grimson, A., Schelter, J., Castle, J., . . . Johnson, J. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature, 433*(7027), 769-773. doi: citeulike-article-id:86487

Loeser, R. F., Goldring, S. R., Scanzello, C. R., & Goldring, M. B. (2012). Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum, 64*(6), 1697-1707. doi: 10.1002/art.34453

Miyaki, S., Nakasa, T., Otsuki, S., Grogan, S., Higashiyama, R., Inoue, A., . . . Asahara, H. (2009). MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses. *Arthritis and rheumatism, 60*(9), 2723-2730. doi: citeulike-article-id:5859549

Miyaki, S., Sato, T., Inoue, A., Otsuki, S., Ito, Y., Yokoyama, S., . . . Asahara, H. (2010). MicroRNA-140 plays dual roles in both cartilage development and homeostasis. *Genes & Development, 24*(11), 1173-1185. doi: citeulike-article-id:7290730

Nakamura, Y., Inloes, J., Katagiri, T., & Kobayashi, T. (2011). Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling. *Molecular and Cellular Biology, 31*(14), 3019-3028. doi: citeulike-article-id:10545632

Sethupathy, P., Megraw, M., & Hatzigeorgiou, A. G. (2006). A guide through present computational approaches for the identification of mammalian microRNA targets. *Nature methods, 3*(11), 881-886.

Tardif, G., Hum, D., Pelletier, J.-P., Duval, N., & Martel-Pelletier, J. (2009). Regulation of the IGFBP-5 and MMP-13 genes by the microRNAs miR-140 and miR-27a in human osteoarthritic chondrocytes. *BMC Musculoskeletal Disorders, 10*, 148. doi: citeulike-article-id:6302126

Verrecchia, F., Chu, M. L., & Mauviel, A. (2001). Identification of novel TGF-beta /Smad gene targets in dermal fibroblasts using a combined cDNA microarray/promoter transactivation approach. *J Biol Chem, 276*(20), 17058-17062. doi: 10.1074/jbc.M100754200

Verrecchia, F., & Mauviel, A. (2002). Transforming growth factor-beta signaling through the Smad pathway: role in extracellular matrix gene expression and regulation. *J Invest Dermatol, 118*(2), 211-215. doi: 10.1046/j.1523-1747.2002.01641.x

Zhu, M., Chen, M., Zuscik, M., Wu, Q., Wang, Y. J., Rosier, R. N., . . . Chen, D. (2008). Inhibition of beta-catenin signaling in articular chondrocytes results in articular cartilage destruction. *Arthritis Rheum, 58*(7), 2053-2064. doi: 10.1002/art.23614

Zhu, M., Tang, D., Wu, Q., Hao, S., Chen, M., Xie, C., . . . Chen, D. (2009). Activation of beta-catenin signaling in articular chondrocytes leads to osteoarthritis-like phenotype in adult beta-catenin conditional activation mice. *J Bone Miner Res, 24*(1), 12-21. doi: 10.1359/jbmr.080901