miR-16 is highly expressed in Paget's associated osteosarcoma

Darrell Green[1], Irina Mohorianu[2], Iain McNamara[3], Tamas Dalmay[2], William D Fraser[1, 4]

1. Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
2. School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
3. Department of Orthopaedics and Trauma, Norfolk and Norwich University Hospital, Norwich Research Park, NR4 7UY, United Kingdom
4. Department of Endocrinology, Norfolk and Norwich University Hospital, Norwich Research Park, NR4 7UY, United Kingdom

Corresponding author: Professor William Fraser, Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK. Tel: 01603 597174 Email: W.Fraser@uea.ac.uk

Conflict of interest: No conflict to declare.

Running title: miR-16 Paget

Dear Editor,

Bone is an endocrine organ that produces key hormones and cytokines (Green et al., 2015b). Paget's Disease of Bone (PDB) is a polygenic disorder of bone turnover first described by Sir James Paget in 1876. PDB is characterised by hyper differentiation and hyper activity of osteoclast cells which induces increased bone remodelling by osteoblasts. The resultant mosaic of bone is structurally weaker, larger, more vascular and porous with an increased susceptibility to fracture. On histology the osteoclasts are increased in size, population and number of nuclei, expressing a "pagetic phenotype" that distinguishes them from normal osteoclasts. Malignant transformation is a rare complication of PDB reported to arise in <1% of PDB patients. Paget's associated osteosarcoma (PDB-OS) consistently arises in sites of pagetic bone and may present with multifocal lesions (Hansen et al., 2006). On histology the lesions are osteoblastic and characterised as an exaggerated form of the accelerated bone remodelling that manifests in PDB. Median survival at diagnosis is 21 months for those treated with curative intent and 7 months for those treated palliatively (Shaylor et al., 1999).

PDB has a strong genetic component. A number of loci have been linked to the disorder with sequestosome 1 (*SQSTM1*) variants associated with more severe symptoms, polyostotic foci and heritable transmission (Hansen et al., 2006). SQSTM1 is multifunctional protein which serves as a signalling hub for diverse cellular events including activation of nuclear factor kappa B (NFKB) and tumour necrosis factor superfamily member 11 (TNFSF11). SQSTM1 also serves as an autophagy receptor for degradation of ubiquitinated molecules via its ubiquitin binding domain (Katsuragi et al., 2015). *SQSTM1* variants associated with PDB are typically located within the coding region of the ubiquitin binding domain. Impairment of autophagy is accompanied by massive accumulation of SQSTM1 and formation of SQSTM1-positive aggregate structures (Katsuragi et al., 2015). As the role of SQSTM1 has not been fully elucidated in transformation and there is no transcriptomic analysis of this cancer, we

55 took a next generation sequencing approach to evaluate the expression of small RNAs such
56 as microRNAs (miRNAs) in PDB and PDB-OS.
57
58 MiRNAs are key regulators of gene expression through gene silencing. MiRNAs can also be
59 used as biomarkers to classify poorly differentiated cancers and cancer tissue origin (Green
60 et al., 2015a). We extracted RNA using the miRCURY RNA isolation kit (Exiqon) from two
61 tissue specimens of PDB-OS with a proven histological diagnosis of osteosarcoma in Paget's
62 affected bone (ages 73 and 81, 2 men). We extracted RNA from four *SQSTM1*-positive PDB
63 tissue specimens taken from affected trabecular bone (ages 79-87, 2 women and 2 men). We
64 extracted RNA from five control bone tissue specimens taken from the femoral heads of
65 trauma patients (ages 68-86, 3 women and 2 men). Tissue samples were collected and
66 preserved at -20 °C. RNA was stored at -80 °C. We generated small RNA libraries using high
67 definition (HD) adapters as previously described (Xu et al., 2015). HD adapters increase the
68 annealing efficiency between small RNAs and adapters. An increased annealing efficiency
69 significantly reduces the RNA ligase-dependent ligation bias in next generation sequencing
70 studies (Xu et al., 2015). We performed sequencing on the HiSeq 2500 Ultra-High-Throughput
71 Sequencing System (Illumina) at the Earlham Institute, Norwich Research Park.
72
73 Raw fastq files were converted to fasta format. The HD signatures of the sequencing reads
74 were trimmed. Quality checking was performed using The UEA Small RNA Workbench
75 (www.srna-workbench.cmp.uea.ac.uk). Reads were mapped with no gaps allowed to the
76 human genome v38 using PatMaN. Small RNA expression levels were normalised using a
77 scaling approach, reads per total, to a fixed total of 10 million reads (Mohorianu et al., 2011).
78 Comparison of the samples was conducted using scatter plots, size-split boxplot of the
79 replicate-to-replicate differential expression, intersection and Jaccard similarity analyses
80 (Mohorianu et al., 2011). Differentially expressed reads between the control, PDB and PDB-
81 OS samples were identified using both an expression interval approach and pairwise
82 comparison using offset fold change (Mohorianu et al., 2011; Mohorianu et al., 2013). Analysis

was conducted using custom made Perl (5.24.0.1) and R (3.2.2) scripts. We observed a high proportion of reads matching to the reference genome (Table 1). The size class distributions were bimodal with peaks at 22 and 32 nt (Figure 1). These peaks correspond to miRNAs and tRNA fragments which was confirmed using annotations. The small number of unique sequences with high abundance is also visible in the complexity distributions (Figure 1). The lower complexities correspond to 22 and 32 nt which indicate a small number of highly abundant sequences (Figure 1). As with all human studies there was variability in the number of reads assigned to either miRNAs or tRNA fragments in the control samples (when compared to inbred genetic models such as mice and fruit flies where variability is minimal). We identified a low expression of miR-16 in PDB compared to controls and a high expression of miR-16 in PDB-OS compared to controls and PDB. We also identified a downregulation of miR-144 and upregulation of miR-21 in PDB-OS compared to controls. Downregulation of miR-144 is in line with previous data in osteosarcoma cells. Upregulation of miR-21 is associated with many types of cancer. In this study we explored miR-16 further.

Total RNA used for next generation sequencing was quantified by density measurement after separation by agarose gel electrophoresis with ethidium bromide staining. Equal amounts of RNA across the sample pools were reverse transcribed using the TaqMan advanced miRNA cDNA synthesis kit (Thermo Fisher Scientific). Differential expression of miR-16 was validated three times in triplicate using TaqMan miRNA advanced assays (Thermo Fisher Scientific). Digital PCR was performed on the QuantStudio 3D Digital PCR System using the GeneAmp PCR System 9700 (Thermo Fisher Scientific). After PCR the chips were imaged on the QuantStudio 3D Instrument which assesses raw data and calculates the concentration of the cDNA sequence targeted by FAM and VIC labelled probes by Poisson distribution (Fazekas de St, 1982). For more in depth analysis the QuantStudio 3D AnalysisSuite was used to report the data as copies/µL. Probe sequence used was miR-16-5p (5'-UAGCAGCACGUAAAUAUUGGCG-3'). We confirmed the downregulation of miR-16 in PDB

compared to controls and upregulation of miR-16 in PDB-OS compared to controls and PDB (Figure 1).

To elucidate a mechanistic role of miR-16 in PDB-OS we used TargetScan v7.1 to search for its messenger RNA targets. *SQSTM1* was identified as a candidate which was supported by performing a second search and identification on miRSearch v3. Target identification of miRNAs is computationally difficult due to the relatively low homology between miRNAs and their targets. We searched the literature and identified *SQSTM1* as an experimentally confirmed target of miR-16, i.e. levels of *SQSTM1* are reduced by miR-16 overexpression (Selbach et al., 2008). We find that miR-16 is downregulated in PDB which means there is a loss of *SQSTM1* negative regulation. In line with previous data *SQSTM1* variants in PDB are activating mutations. SQSTM1 is a signalling hub for the activation of TNFSF11. TNFSF11 is the primary ligand for stimulating the differentiation of monocyte progenitors to osteoclasts through interaction with tumour necrosis factor receptor superfamily member 11a (TNFRSF11A). TNFSF11 is produced by osteoblasts and downstream repercussions of the TNFSF11-TNFRSF11 interaction is the increase of bone resorption. Bone resorption achieved by various osteoclast-produced cytokines is set within a positive feedback loop which encourages osteoblast proliferation and activity. Increased proliferation of osteoblast cells containing *SQSTM1* activating variants repeats the bone remodelling cycle each time increasing the number of *SQSTM1* positive osteoblasts. PDB does not typically present until patients are >55-years-old. In patients harbouring *SQSTM1* variants, presentation can appear earlier. Presentation may be due to the speed of the bone remodelling cycle or the gradual loss of miR-16 mediated silencing of *SQSTM1*.

We find a high expression of miR-16 in PDB-OS. The effect of upregulated miR-16 means there is increased negative regulation of genes including *SQSTM1*. Loss of SQSTM1 has little effect on autophagy (Katsuragi et al., 2015). The lack of phenotypic impact might be explained by the presence of other autophagy regulators including optineurin (OPTN) (Katsuragi et al.,

2015). As a signalling hub SQSTM1 is a partner of the Kelch like ECH associated protein 1 (KEAP1)-nuclear factor erythroid 2 like 2 (NFE2L2) pathway, a major cellular defence mechanism against oxidative stress (Katsuragi et al., 2015). In normal conditions NFE2L2 is constitutively degraded by the ubiquitin-proteasome system because its binding partner KEAP1 is an adapter of ubiquitin ligases. Upon exposure to oxidative stress KEAP1 is structurally modified which releases NFE2L2 to translocate to the nucleus. NFE2L2 is a transcription factor which then induces a multitude of antioxidant and anti-inflammatory genes. SQSTM1 is able to abrogate the interaction between KEAP1 and NFE2L2 leading to protection against oxidative stress (Katsuragi et al., 2015). We propose miR-16 mediated silencing of *SQSTM1* incapacitates the cell's ability to protect itself against oxidative stress-induced oncogenesis. This proposal warrants further molecular investigation to show an inverse correlation between the expression of miR-16 and *SQSTM1*.

Molecular testing may help provide a robust diagnosis and is particularly useful in rare cancers. This is especially true in PDB where transformation to osteosarcoma is often missed until late stage. Radiograph morphology is subtly different between PDB and PDB-OS. Symptoms of PDB-OS can be similar to the day-to-day symptoms experienced by a PDB patient. Biomarkers such as serum total alkaline phosphatase used in the diagnosis of PDB are the same biomarkers used to aid the diagnosis of PDB-OS. These features combined make it difficult to distinguish between a chronic non-life threatening disorder and a fatal disease which require very different treatment strategies. Quantitative PCR or next generation sequencing performed on RNA extracted from a biopsy of a presenting lesion is precise for detecting a reduced expression of miR-16 in 'normal PDB' and high expression of miR-16 in PDB-OS. The switch in miR-16 expression could alert physicians to the change to an osteosarcoma phenotype.

A limitation of this study is the size of the cohorts studied. PDB-OS arises in <1% of PDB patients and donation to tissue banks is scarce. We are in the process of collecting a larger

cohort of tissue specimens to validate the findings from this study. Our data highlights the value of being able to provide a robust tissue diagnosis in addition to identifying regulatory transcriptomic molecules that could be exploited for targeted therapy. Expression of miR-16 could also be investigated in paediatric osteosarcoma to evaluate similarities and/or differences between the two age-related incidence peaks of osteosarcoma.

DATA AVAILABILITY

The data presented in this study is publicly available on Gene Expression Omnibus under the accession numbers GSE85809 (GSM2284729 to GSM228473 are control samples; GSM2284736 and GSM2284736 are PDB-OS samples) and GSE87018 (GSM2318966 to GSM2318970 are PDB samples).

FUNDING

This study was supported by Big C and Paget's Association.

ACKNOWLEDGEMENTS

Approval to obtain and study human tissue from the Norwich Biorepository was granted by the Faculty of Medicine and Health Sciences Research Ethics Committee (2013/14 – 22 HT). The authors thank Dr Michael Davie and Sally Evans from the Robert Jones and Agnes Hunt Orthopaedic Hospital for contributing samples to the biorepository. We are indebted to the study participants. We dedicate this study to the memory of Allan Reid.

REFERENCES

1. Fazekas de St, G. 1982 The evaluation of limiting dilution assays. *Journal of Immunological Methods* **49** R11-23.

2. Green D, Dalmay T & Chapman T 2015a Microguards and micromessengers of the genome. *Heredity* **116** 125-134.

3. Green D, Dalmay T & Fraser WD 2015b Role of miR-140 in embryonic bone development and cancer. *Clinical Science* **129** 863-873

4. Hansen MF, Seton M & Merchant A 2006 Osteosarcoma in Paget's disease of bone. *Journal of Bone and Mineral Research* **21** 58-63.

5. Katsuragi Y, Ichimura Y & Komatsu M 2015 p62/SQSTM1 functions as a signaling hub and an autophagy adaptor. *FEBS Journal* **282** 4672-4678.

6. Mohorianu I, Schwach F, Jing R, Lopez-Gomollon S, Moxon S, Szittya G, Sorefan K, Moulton V & Dalmay T 2011 Profiling of short RNAs during fleshy fruit development reveals stage-specific sRNAome expression patterns. *The Plant Journal* **67** 232-246.

7. Mohorianu I, Stocks MB, Wood J, Dalmay T & Moulton V 2013 CoLIde: a bioinformatics tool for CO-expression-based small RNA Loci Identification using high-throughput sequencing data. *RNA Biology* **10** 1221-1230.

8. Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R & Rajewsky N 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455** 58-63.

9. Shaylor PJ, Peake D, Grimer RJ, Carter SR, Tillman RM & Spooner D 1999 Paget's Osteosarcoma - no cure in sight. *Sarcoma* **3** 191-192.

10. Xu P, Billmeier M, Mohorianu I, Green D, Fraser WD & Dalmay T 2015 An improved protocol for small RNA library construction using high definition adapters. *Methods in Next Generation Sequencing* **2** 1-10.