# Supplementary data

## Downregulated MYL1 emerges as a promising diagnostic biomarker for rheumatoid arthritis

Chaoquan Yang[1*]
Zhiling Huang[1*]
Xifan Zheng[1*]
Haojun Tang[1]
Wenpeng Qin[1]
Yicheng Liang[1]
Haidong Chen[1]
Wenjun Hao[1]
Dan Yi[2]
William W Lu[3]
Yan Chen[1]
[1] Department of Bone and Joint Surgery, the First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China
[2] Faculty of Pharmaceutical Sciences, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, People's Republic of China
[3] Department of Orthopedics and Traumatology, the University of Hong Kong, Hong Kong, People's Republic of China

*These authors contributed equally to this work

Correspondence: Dr. Yan Chen
Department of Bone and Joint Surgery, the First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China
Tel (optional): +8607715319091
Fax (optional): +8607715319091
Email (mandatory): cy003@connect.hku.hk

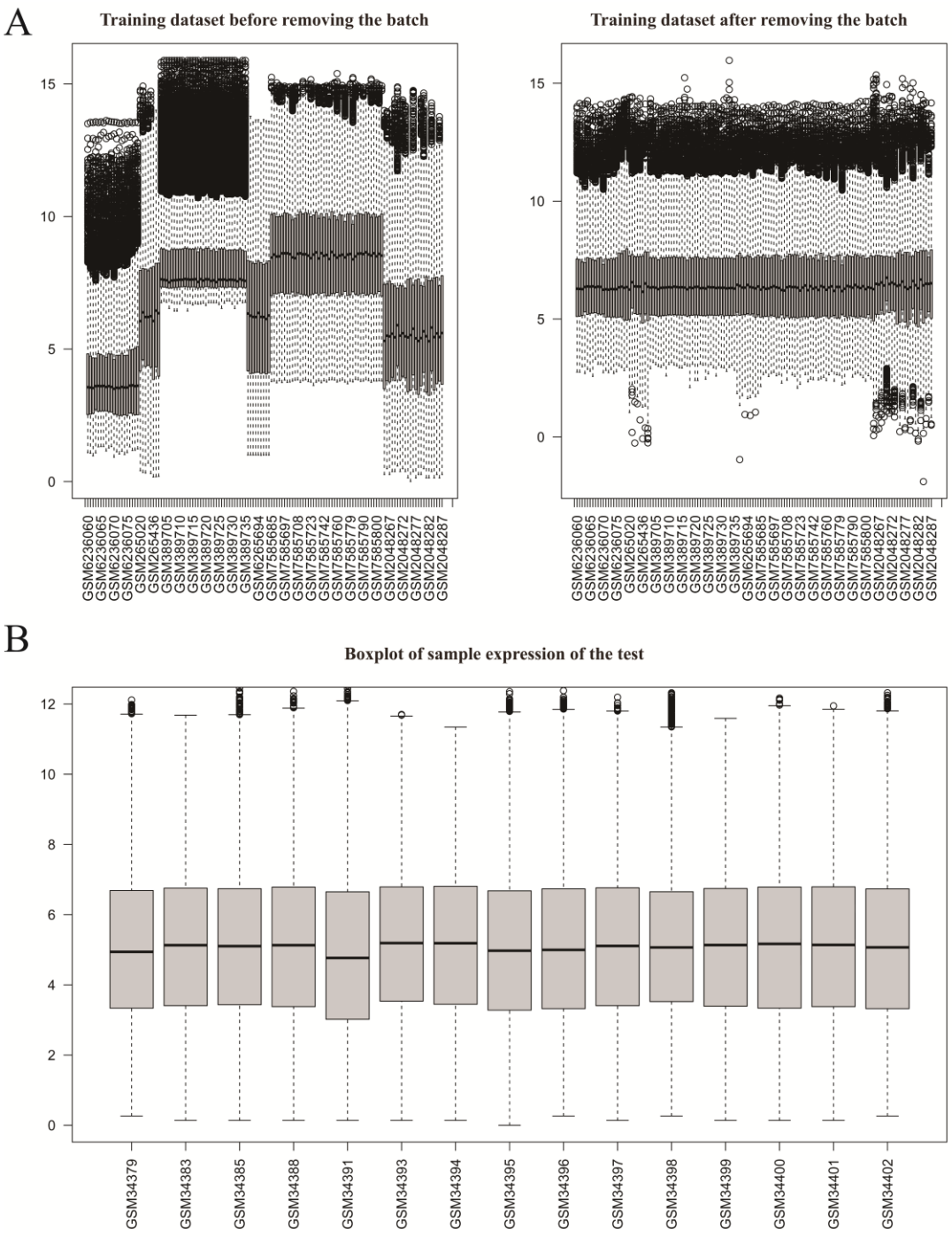


**Figure S1. A.** The boxplot of the range of relative logarithmic expression for each sample in training database before and after batch removal. **B.** Boxplot of the relative logarithmic expression for each sample in the test dataset.

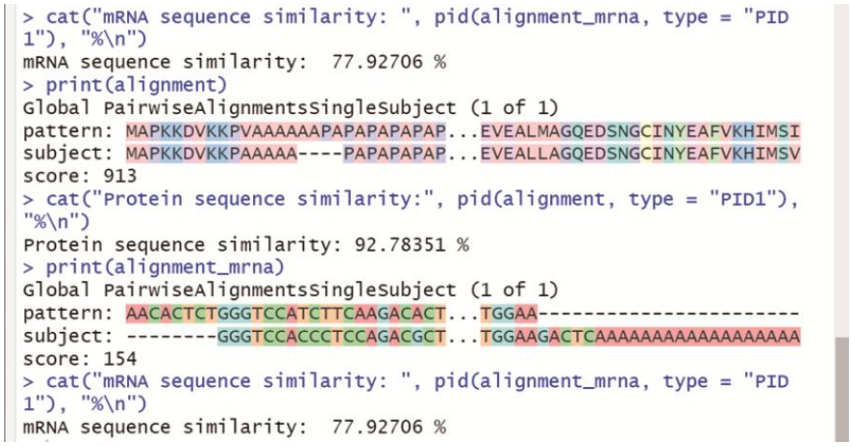


**Figure S2. Alignment results of human and rat MYL1 protein and mRNA sequences.**

Global pairwise alignments were performed for both mRNA and protein sequences. Alignment results show matched and mismatched regions between the target and reference sequences. Sequence similarity (PID1) was computed using the pid() function. The mRNA sequence similarity was 77.93%, while the protein sequence similarity was 92.78%. Highlighted regions indicate conserved amino acid or nucleotide segments.

**Supplementary Methods: Sequence Alignment Procedure for Human and Rat MYL1**

To assess the similarities and differences between human and rat MYL1 protein and mRNA sequences, we performed sequence alignment analysis using a global alignment approach (Needleman-Wunsch algorithm). The specific steps were as follows:

**Sequence data preparation**

The target protein and mRNA sequences for both humans and rats were downloaded from the National Center for Biotechnology Information (NCBI). The protein sequence files were named human_protein.fasta (https://www.ncbi.nlm.nih.gov/protein/NP_524144.1?report=fasta) and rat_protein.fasta (https://www.ncbi.nlm.nih.gov/protein/NP_001071124.1?report=fasta), respectively. The mRNA sequence files were named human_mrna.fasta (https://www.ncbi.nlm.nih.gov/nuccore/NM_079420.3?report=fasta) and rat_mrna.fasta (https://www.ncbi.nlm.nih.gov/nuccore/NM_001077656.1?report=fasta), respectively. All sequence files were stored in FASTA format to ensure the accuracy and integrity of the sequence data.

**Protein Sequence Alignment**

Global alignment of the target human and rat protein sequences was

performed using the pairwiseAlignment function from the Biostrings package. The BLOSUM62 matrix was used as the amino acid substitution matrix during the alignment process. The gap opening penalty was set to 10 and the gap extension penalty to 0.5, effectively balancing the stringency and flexibility of the alignment to yield reliable and biologically meaningful results.

**mRNA Sequence Alignment**

For global alignment of the target human and rat mRNA sequences, a custom nucleotide substitution matrix was employed, with a match score of 1 and a mismatch penalty of -3. Similarly, the gap opening penalty was set to 10 and the gap extension penalty to 0.5. This configuration effectively distinguishes between matched and mismatched positions while avoiding excessive error introduced by gaps.

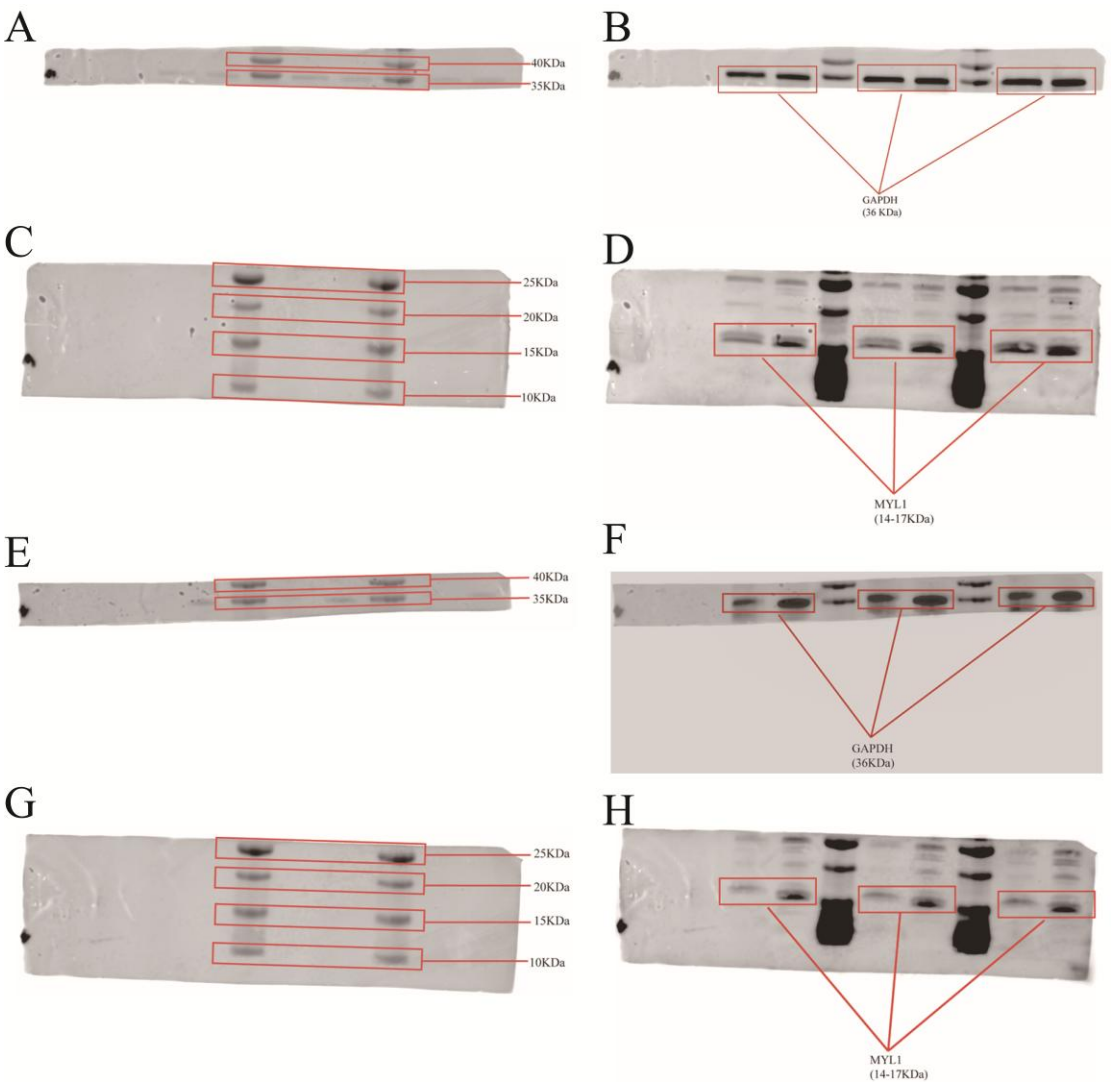


**Figure S3. Original Western blot images of GAPDH and MYL1.**
Western blot analyses were performed using a molecular weight marker (WB1902S, Botex, Beijing, China) and the MYL1 antibody (15814-1-AP, Proteintech, Wuhan, China).
**(A, E)** Original GAPDH blot images showing molecular weight markers (ladder). The red boxes indicate the regions containing the ladder bands.
**(B, F)** Original GAPDH loading control blots used for protein normalization; the red boxes mark the protein band areas.
**(C, G)** Original MYL1 blot images showing molecular weight markers (ladder). The red boxes indicate the regions containing the ladder bands.
**(D, H)** Original MYL1 blots detected using the anti-MYL1 antibody (15814-1-AP, Proteintech, Wuhan, China); the red boxes mark the MYL1 band areas.
Panels E–H represent repeated experiments.