Identification of microRNA-16-5p and microRNA-21-5p in feces as potential noninvasive biomarkers for inflammatory bowel disease

Rui Zhou^{1,2,*}, Peishan Qiu^{1,2,*}, Haizhou Wang^{1,2}, Huijie Yang^{1,2}, Xueying Yang³, Mingliang Ye^{1,2}, Fan Wang^{1,2}, Qiu Zhao^{1,2}

¹Department of Gastroenterology, Zhongnan Hospital of Wuhan University, Wuhan 430071, China ²Hubei Clinical Center and Key Laboratory of Intestinal and Colorectal Diseases, Wuhan 430071, China ³Department of Medical Records, The Central Hospital of Enshi Autonomous Prefecture, Enshi 445000, China *Equal contribution

Correspondence to: Fan Wang, Qiu Zhao; email: <u>fanndywang@foxmail.com</u>, <u>https://orcid.org/0000-0003-3510-3010</u>; <u>giuzhao@whu.edu.cn</u> Keywords: microRNA, feces, IBD, ulcerative colitis, Crohn's disease

Received: July 13, 2020 Accepted: December 10, 2020

Published: February 1, 2021

Copyright: © 2021 Zhou et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution</u> <u>License</u> (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Inflammatory bowel disease (IBD) is a chronic idiopathic gastrointestinal disease. Increasing evidence suggests that microRNAs (miRNAs) may participate in the pathophysiology of IBD.

Methods: A miRCURY[™] LNA Array and *in situ* hybridization were employed to screen for differentially expressed miRNAs (DEMs) in fecal specimens from 41 IBD patients (22 ulcerative colitis (UC), 19 Crohn's disease (CD)) and 23 healthy controls (HC). RT-qPCR was performed to confirm the findings. The DEMs target genes and corresponding biological functions were predicted by bioinformatics analysis.

Results: Compared with HC, miR-16-5p in the feces was up-regulated both in UC and CD patients (p < 0.01), while miR-21-5p was up-regulated only in UC patients (p < 0.01). TargetScan 7.2, miRWalk, and miRDB were used to predict 216 public target genes of miR-16-5p and miR-21-5p, and six hub genes (PIK3R1, GRB2, SUZ12, NTRK2, Smurf2, and WWP1) were analyzed using the STRING database and Cytoscape. All the hub genes promote the occurrence and development of IBD-related colorectal cancer.

Conclusions: The elevated levels of miR-16-5p and miR-21-5p in feces of IBD patients have to guide significance for the noninvasive clinical diagnosis of IBD and have a warning effect on the occurrence of IBD-related colorectal cancer.

INTRODUCTION

Inflammatory bowel disease (IBD) is a cluster of chronic idiopathic, immune, relieving, and recurrent gastrointestinal diseases that occur when genetically susceptible populations are exposed to environmental risk factors. IBD includes Crohn's disease (CD) and ulcerative colitis (UC) based on differences in clinical phenotypes [1]. The etiology of IBD is considered as complex interactions among genes, immune responses, and environmental factors [2]. The medical and surgical treatment of IBD is primarily driven by an accurate assessment of disease activity. The mucosal inspection via endoscopy remains the golden standard for estimating disease activity in IBD, and the use of crosssectional imaging is increasing [3]. However, most of these monitoring methods are not suitable and optimal for routine clinical applications due to their disadvantages of intrusiveness, time-consuming, and expensive. It is urgently needed to develop better tools to overcome this dilemma in screening and disease activity assessment of IBD.

In recent years, studies on noninvasive measurable biomarkers for the diagnosis of IBD have gradually increased [4, 5]. Due to the influence of bewildering

factors such as age, sex, or body mass index, conventional inflammatory marks (such as C-reactive protein (CRP) and erythrocyte sedimentation rate) do not accurately reflect disease activity [6]. Most importantly, both of them are not specific to IBD [7]. Due to the complexity of detection technology, the detection results lack clinical reliability, lack of assessment of responsiveness to changes in disease states, and cannot replace the performance of endoscopy assessment. Fecal levels of calprotectin and lactoferrin have not been identified as biomarkers for diagnosing IBD disease activity [8]. Fecal molecular aberration detection is also a promising noninvasive method for IBD screening, among which DNA testing is the most established test [9-11]. Johnson et al. reported significantly higher methylation rates of bone morphogenic protein 3 (BMP3) or N-Myc downstreamregulated gene 4 (NDRG4) in IBD lesions than matched controls, which revealed that specific DNA markers that were present in advanced IBD neoplasia could be detected in the tissues and feces when small adenomas occur in IBD patients [9].

Epigenetic factors can mediate interactions between the environment and the genome [12]. As a primary epigenetic mechanism, RNA interference delivered by microRNA (miRNA) may have a significant effect on the pathogenesis of IBD and other diseases [13–16]. Koukos et al. found that miR-4284 was the most down-regulated microRNA significantly in the intestinal mucosa of pediatric patients with UC compared with non-IBD controls. What's more, the lower the expression level of miR-4284 predicted the higher the disease activity of IBD [13]. MiRNAs, a group of short single-stranded RNA molecules, could reduce gene expression through the deterioration of target mRNAs or blockage of translation [17], thereby altering the export of many protein-coding genes and related pivotal cellular biological functions [18]. MiRNAs were reported to be associated with the regulation of autophagy, inflammation, and fibrosis associated with IBD [14-16]. Besides, the normal intestinal development of mice is inseparable from the role of miRNAs produced by intestinal epithelial cells [19]. The most widely used method for screening miRNAs in whole blood, serum specimens, and intestinal mucosa as biomarkers of IBD disease activity are microarray approaches, which are rarely used in feces [20, 21]. Fecal miRNAs are acquired in a relatively minimally invasive manner and can be briskly quantified by quantitative polymerase chain reaction (qPCR) or microarrays. It follows that fecal miRNAs are alluring noninvasive biomarkers for the diagnosis of IBD. In this study, the fecal miRNA expression profile from IBD patients and healthy controls (HC) were compared to recognize novel potential miRNA biomarkers with higher sensitivity involved in IBD in feces. Our study will help to avoid invasive radiological or endoscopic investigation in IBD patients and provide a direction for the exploration of new treatment strategies.

RESULTS

MiRNA expression profiling in IBD feces

To explore the underlying role of fecal miRNAs in diagnosing IBD, fecal miRNA expression profiles of IBD patients and healthy volunteers were evaluated by microarray analysis. Microarray analysis revealed 3100 miRNAs were detected in total in the fecal specimens collected from 41 IBD patients (22 UC and 19 CD) and 23 HC. Of seven DEMs identified in the feces of IBD patients, two miRNAs (let-7i-3p and miR-326) exhibited decreased expression, and five miRNAs (miR-15a-5p, miR-16-5p, miR-21-5p, miR-338-5p, miR-483-5p) demonstrated increased expression compared with that in the feces of HC (fold change > 7, P < 0.05), as shown in Table 1.

Validation of miRNA candidates in IBD feces

The expression levels of DEMs in the feces of IBD patients and healthy volunteers were verified by RTqPCR. Compared with HC, the expression of miR-16-5p was significantly increased both in the fecal specimens of UC and CD patients (Figure 1A). However, the increase of miR-21-5p was only observed in the feces of UC patients (Figure 1B), and there was no difference between CD patients and HC. It was worth noting that the expression levels of the other five miRNAs had negligible differences in fecal specimens from IBD patients and healthy volunteers (Figure 1C). Based on all these findings, the follow-up analysis of this study focused on miR-16-5p and miR-21-5p.

Validation of DEMs

In the GSE101817 test set, miR-21-5p expression was increased in the intestinal mucosa of CD patients than in the normal intestinal mucosa (Figure 1D). Thus, in IBD patients, miR-21-5p was distinctly expressed not only in the feces but also in the intestinal mucosa.

MiR-21-5p and miR-16-5p are potential noninvasive feces markers for IBD

The diagnostic capability of miR-21-5p and miR-16-5p for IBD was verified by Receiver operating characteristics (ROC) curves analysis. It was surprising to find that miR-16-5p showed a great power in diagnosing CD (area under curve (AUC) = 0.868, p <

Upregulated miRNA	Downregulated miRNA		
miR-15a-5p	let-7i-3p		
miR-16-5p	miR-326		
miR-21-5p			
miR-338-5p			
miR-483-5p			

Table 1. Differentially expressed miRNAs in the feces of IBD patients detected by miRCURY[™] LNA Array (fold change > 7, P < 0.05).

0.001) and UC (AUC = 0.853, p < 0.001). Besides, the ROC curve analysis of miR-21-5p for the diagnostic accuracy of UC showed an AUC of 0.810 (p < 0.001) (Figure 1E).

Prediction of the target genes of miR-21-5p and miR-16-5p

The potential target genes of miR-16-5p and miR-21-5p were obtained from three online databases (TargetScan 7.2 [22], miRWalk [23], and miRDB [24]), respectively. After duplicates removal, there were 1925 and 954 target genes of miR-16-5p and miR-21-5p, respectively. A total of 216 genes were recognized as common targets of both miR-16-5p and miR-21-5p by overlapping analysis, and the result was shown using a Venn diagram (Figure 2A).

PPI network construction and module analysis

The Search Tool for the Retrieval of Interacting Genes (STRING) database [25] was performed to construct the protein-protein interaction (PPI) network and analyze the interaction between target genes (Figure 2B). Cytohubba, a plug-in of the STRING database, was used to find the top 15 hub genes [26] (Figure 2C). Furthermore, with degree ≥ 10 as the threshold, six genes were selected for further analysis, including phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1), growth factor receptor-bound protein 2 (GRB2), SUZ12 polycomb repressive complex 2 subunit (SUZ12), neurotrophic receptor tyrosine kinase 2 (NTRK2), SMAD specific E3 ubiquitin-protein ligase 2 (Smurf2), and WW domain-containing E3 ubiquitinprotein ligase 1 (WWP1). MCODE was used to select three remarkable modules from the PPI network complex on the basis of the degree of importance [27] (Figure 2D).

Function and pathway enrichment analysis

The function analysis of target genes in the modules was predicted using the STRING database. Kyoto Encyclopedia of Genes and Genomes (KEGG) [28] and Gene Ontology (GO) enrichment analysis of hub genes in the most notable modules showed that the crucial enrichment pathways included signal transduction, cellular response to endogenous stimulus, cellular response to growth factor stimulus, enzyme-linked receptor protein signaling pathway, cell surface receptor signaling pathway, regulation of growth, regulation of transferase activity, and regulation of protein modification process (Figure 2E). Most of these pathways involved epithelial regeneration and epithelial-mesenchymal transition (EMT) associated with IBD [29, 30].

DISCUSSION

The etiology of IBD, a chronic inflammatory gastrointestinal disease, is unknown [31]. Over the past few decades, the incidence of IBD has increased significantly in many countries, which has placed broad social and economic pressures on public health structures and systems [32, 33]. Many studies have shown that miRNAs, as gene expression regulators, are associated with various inflammatory states [34-36]. Of course, IBD is no exception. In recent years, the key role of miRNAs in regulating the pathological progression of IBD has been gradually reported [37-40]. It has been reported that the serum miRNAs expression profiles in CD and UC patients were distinct from that of normal controls [41, 42]. Chen et al. found a positive correlation between serum miR-146b-5p expression and IBD disease activity. Furthermore, miR-146b-5p was more specific than CRP [43], which was the presently available succedaneous biomarker for inflammation in IBD [44]. Besides, Ge et al. reported that the expression of miR-125a in the inflamed intestinal mucosa of IBD patients was lower than that in healthy volunteers, and miR-125a could protect intestinal mucosa from inflammatory injury [45]. Compared with mucosal miRNA, detecting the differential expression of fecal miRNAs can more accurately assess the disease activity or mucosal healing of gastrointestinal diseases. Since differentially expressed miRNAs in feces are another manifestation of local variation, such as exfoliation of intestinal



Figure 1. MiR-21-5p and miR-16-5p are potential noninvasive feces markers for IBD. (A) Expression of miR-16-5p in fecal specimens from HC, UC, and CD. (B) Expression of miR-21-5p in fecal specimens from HC, UC, and CD. (C) Expression of seven candidate DEMs (let-7i-3p, miR-15a-5p, miR-16-5p, miR-21-5p, miR-326, miR-338-5p and miR-483-5p) in fecal specimens from HC, UC and CD. (D) Compared with normal intestinal mucosa tissue based on GSE101817, the expression of miR-21-5p in the intestinal mucosa of CD patients was up-regulated. (E) Diagnostic value of miR-16-5p and miR-21-5p in fecal specimens of IBD patients was evaluated by ROC curves. *p < 0.05, **p < 0.01, ***p < 0.001, ***p < 0.001; HC, healthy control; UC, ulcerative colitis; CD, Crohn's disease; ns, not statistically significant; ROC, receiver operating characteristic.



Figure 2. Prediction of miR-16-5p and miR-21-5p target genes and their biological functions. (A) Venn Diagram of the common target genes of miR-16-5p and miR-21-5p predicted by TargetScan, miRWalk and miRDB. (B) Protein-protein interaction (PPI) network of miR-21-5p and miR-16-5p target genes. (C) Top 15 hub target genes identified by the Cytohubba plug-in of Cytoscape. (D) Three key modules of the PPI network identified by the MCODE plug-in of Cytoscape. (E) Function analysis of all target genes in three modules. FDR, false discovery rate.

epithelial cells, exosome production, and mucosal barrier changes [40, 46]. Moreover, fecal miRNAs are relatively stable and resistant to harsh conditions, making it clinically possible as a noninvasive biomarker [46]. However, there have been relatively few studies on miRNAs screening in feces of IBD patients [47]. Duran-Sanchon et al. believed that the elevated levels of miR421, miR27a-3p, and hemoglobin in feces could identify patients with advanced adenoma or colorectal cancer (CRC) more accurately than the concentration of hemoglobin in feces alone [48].

In our study, we observed differences in fecal miRNA between IBD patients and healthy subjects through microarray approaches. The results showed that 34 miRNAs increased and nine miRNAs decreased in both UC and CD patients. Subsequently, combined with previous reports [47], we identified seven miRNAs for later studies. However, qPCR analysis of fecal specimens confirmed that miR-16-5p was increased in both UC and CD, while miR-21-5p was only increased in UC patients' feces. Also, miR-16-5p expression was down-regulated in CRC tissues compared with the normal intestinal mucosa [49], and miR-16-5p showed sharp tumor-suppressive roles in CRC [50]. It is worth mentioning that colitis-associated CRC in IBD patients is substantially clinically different from the fragmented CRC observed in the common population. Occult evolution occurs long before clinically detectable neoplasms develop to IBD-associated carcinogenesis. [51].

MiR-21-5p was greatly obviously expressed in colorectal cancer tissues and could target tumor suppressor genes through an epigenetic modification to promote survival and resist apoptosis. Therefore, miR-21-5p was considered to be a carcinogenic miRNA [30, 52]. Moreover, miR-21-5p had higher accuracy in distinguishing superficial and deep tumors in oral squamous cell carcinoma [53]. In patients with colorectal cancer, serum miR-21-5p was higher before the surgery but down-regulated after the surgery, which revealed that the decline of miR-21-5p might support a better overall survival in colorectal cancer patients [54]. In our study, miR-21-5p was expressed remarkably only in fecal specimens of patients with UC. The important thing was that there was an inextricable relationship between UC and CRC [55]. Therefore, the increase of miR-21-5p promoted the revolution from UC to CRC.

Three online websites (TargetScan 7.2, miRWalk, and miRDB) were used to predict the potential target genes of miR-16-5p and miR-21-5p, and co-expression network construction and module analysis were carried out in Cytoscape, and six hub genes were finally screened. Among them, PIK3R1 was reported to be

involved in immune-related mechanisms in the progression of UC [56]. Early embryos of mice with complete deletion of GRB2 cannot survive. Members of the GRB2 protein family are essential for the initiation and development of a variety of developmental and disease-related signaling complexes. Besides, GRB2 makes a great difference in the progression of autoimmune diseases due to its involvement in T cells' development [57]. GRB2 associated binding protein 2/3 (Gab2/3) double knockout mice could develop spontaneous colitis with rectal prolapse and diarrhea, mainly involving macrophages and CD8⁺ T cells, which was due to the role of Gab2/3 in suppressing the inactivation of immune cells in the process of inflammation [58]. SUZ12 was reported to be an oncogene in colorectal cancer, which plays a cancerpromoting role by methylating histone H3 [59]. Compared with sporadic CRC, the expression of NTRK2 in UC-associated CRCs was significantly different [60]. The expression of Smurf2 in CRC tissues was markedly increased than the corresponding healthy intestinal mucosa, and the level of Smurf2 in microsatellite instability (MSI) CRC was also dramatically higher than that in microsatellite stability (MSS) CRC [61]. More importantly, the high expression of Smurf2 in tumors indicated a poor prognosis. Besides, the overall survival and disease-free survival of CRC patients with high WWP1 expressions were worse than those with low WWP1 expressions [62], which was consistent with the tumorigenic effect of Smurf2. In short, most hub target genes of miR-16-5p and miR-21-5p are essential for the pathomechanism and development of CRC. Patients with UC or CD have a significantly increased risk of gastrointestinal and extra-intestinal malignancies, with CRC being the most common cancer associated with IBD [63]. Mortality from cardiovascular disease, infection, and cancer in IBD patients increases year by year [64]. Active cancers are the second leading cause of mortality after cardiovascular disease in IBD patients [65].

In conclusion, fecal specimens were collected from clinical IBD patients, and differentially expressed miRNAs associated with the occurrence of IBD were screened. The elevated levels of miR-16-5p and miR-21-5p in feces of IBD patients are of guiding significance for the noninvasive clinical diagnosis of IBD and have a warning effect on the occurrence of IBD-related CRC in IBD patients.

MATERIALS AND METHODS

Patient recruitment

Fecal specimens were collected from 41 IBD patients and 23 healthy subjects. Table 2 listed the basic

Category	UC	CD	НС	P value
No. of cases	22	19	23	-
Age at enrollment, y				
Mean±SD	36.22±10.11	34.19±11.76	29.71±10.08	ns
Gender, number (%)				
Female	5(22.7)	12(63.2)	7(30.4)	
Disease duration, mo	41.11±43.51	55.64±71.63		ns
Disease location, n (%)				
L1	-	9(47)		
L2	-	7(37)		
L3	-	3(16)		
Disease extent, n (%)				
Proctosigmoiditis	5(23)	-		
Left-sided colitis	6(27)	-		
Pancolitis	11(50)	-		
CRP (mg/L)	6.05 ± 7.46	17.57±18.13	-	ns
ESR (mm/h)	14.67±11.12	30.33 ± 28.52	-	ns
ALB (g/L)	41.94 ± 3.78	38.64 ± 4.45	-	ns
WBC (×10 ⁹ /L)	6.82±2.72	6.09 ± 2.71	-	ns
HGB (g/L)	128.33±17.96	102.52±14.23	-	ns
RBC (×10 ¹² /L)	4.11±0.58	3.73±0.76	-	ns
HCT (%)	35.41±9.96	30.69±4.17	-	ns
PLT(×10 ⁹ /L)	222.71±37.90	288.14±109.08	-	****
FOBT (+), n (%)	11(50.0)	7(36.8)	-	-

Table 2. Clinical characteristics of the study subjects.

UC, ulcerative colitis; CD, Crohn's disease; HC, healthy control; SD, standard deviation; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; ALB, albumin; WBC, white blood cell; HGB, hemoglobin; RBC, red blood cell; HCT, hematocrit; PLT, platelet count; FOBT, fecal occult blood test; *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001; ns, not statistically significant.

characteristics of all subjects. Before specimen collection, written informed consent was acquired from all subjects, and the study protocol was accepted by the Ethics Review Committee of Zhongnan Hospital of Wuhan University (Hubei, China). All authors and researchers involved in the study read and strictly adhered to the ethics guidelines of the World Medical Association (Declaration of Helsinki). The healthy controls had no abnormal symptoms and a negative colonoscopy, excluding those with other gastrointestinal diseases or taking antibiotics. The subjects were between the ages of 16 and 58.

Fecal specimen collection and fecal miRNA extraction

Fresh human fecal specimens (20-30 g) were collected with a 50-ml specimen cup and stored at -80° C for a long time. All samples were tested together after collection to reduce the batch effect. 10 mg feces (wet weight) from each specimen was added into a 1.5 ml RNase-free microcentrifuge tube containing 1 ml of Trizol LS reagent (Invitrogen, Carlsbad, CA, USA). Use an RNase-free grinder to homogenize the feces mixture completely. 200 µl of chloroform was added to the microcentrifuge tube for RNA precipitation. According to the manufacturer's recommendation, QIAgen miRNeasy Mini Kit (Qiagen, Valencia, CA, USA) was used for total RNA extraction. Total RNA was eluted in RNase-free water. Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA) was used for the assessment of total RNA concentration and standardized the total RNA concentration of all fecal samples to 1ng/ul.

MiRNA microarray screening

Kang Chen (Shanghai, China) performed the miRNA microarray screening after RNA isolation from the specimens. Total miRNAs from fecal specimens were labeled with miRCURY[™] Hy3[™]/Hy5[™] fluorescent dye using the miRCURY[™] Power Labelling Kit (Exiqon, Vedbaek, Denmark) following the manufacturer's guidelines, and then hybridized to the miRCURY[™] LNA miRNA Array (version 18.0, Exiqon, Vedbaek, Denmark). The Axon GenePix 4000B microarray scanner (Molecular Devices, USA) was used to scan the slides. The Axon GenePix 4000B microarray scanner (Molecular Devices, USA) was used to scan the slides. The data were extracted with Agilent Feature Extraction v10.7 software and analyzed with GeneSpring GX (version 13.0, Agilent Technologies, California, USA). After normalization, the absolute value of the fold change of gene expression was more than seven as a significant difference, and this method was used to screen DEMs.

Quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR) detection

The miRNA expression level was assessed utilizing qRT-PCR, which was performed using the SYBR green method on Bio-Rad according to the manufacturer's specification. The RNAs were subjected to reverse transcribed with TaqMan miRNA Reverse Transcription kit (Applied Biosystems, Foster City, CA, USA) following the protocol, obtaining cDNA product. All amplifications were performed in triplicate, and the quantification of the results was analyzed by the comparative Ct $(2^{-\Delta\Delta}Ct)$ method. U6 was used to standardize miRNA expressions.

Validation of DEMs

The GSE101817 downloaded from the GEO database (<u>http://www.ncbi.nlm.nih.gov/geo/</u>) was used as a test set to confirm the relationship between the previously selected miRNAs and IBD.

Predicting the target gene of DEMs

The possible target genes of miRNAs were predicted using TargetScan 7.2 [22] (<u>http://www.targetscan.org/</u><u>vert 72/</u>), miRWalk [23] (<u>http://mirwalk.umm.uniheidelberg.de/</u>), and miRDB [24] (<u>http://www. mirdb.org/</u>). These databases contain experimentally validated miRNA targets and provide the most recently updated targets. The anticipated target genes of each differentially expressed miRNA from three different sites were combined. After removing the duplicates, the Venn diagram was used for overlap analysis.

Protein-protein interaction (PPI) network and module analysis

The PPI network analysis of the overlapping target genes was performed by the STRING (<u>http://string-db.org</u>, version 11.0) [25] online database. The minimum required interaction score of > 0.4 was defined as statistically significant. Cytoscape (version 3.7.2) enabled the visualization of complex protein interaction networks [66]. The Molecular Complex Detection (MCODE) (version 1.5.1) was an application for performing module analysis in Cytoscape [27]. The PPI network was built in Cytoscape and used MCODE to identify the most significant modules. Set the thresholds as follows: degree cut-off = 2, node score cut-off = 0.2, Max. depth = 100 and k-score = 2. Further hub gene screening was performed using the cytoHubba (version 0.1) plug-in of Cytoscape [26].

Function and pathway enrichment analysis

Gene Ontology (GO) contains three aspects of functional information, including the biological processes involved in genes, the location of cells, and the molecular functions they play and organizes these functional concepts into a DAG (Directed Acyclic Graph) structure [67]. KEGG (https://www.kegg.jp/) is an exhaustive database of functional interpretation and practical application of genomic information and integrates macromolecular datasets from genome sequencing and other high-throughput experimental techniques [28]. KEGG and GO functional enrichment analysis for target genes was performed using the STRING database [25]. False discovery rate (FDR) < 0.05 was defined to be statistically significant.

Statistical analysis

Statistical analyses were achieved by GraphPad Prism 5.0 (GraphPad Software, Inc.). The mean and standard deviation (SD) or standard deviation (SE) are used to describe the values obtained by each continuous variable in the study. For categorical variables, absolute and relative frequencies were used. The nonparametric correlation between the two variables was evaluated by the Spearman rank correlation coefficient (r). The diagnostic effect of fecal miRNAs compared to the control group for IBD were analyzed by ROC curve and estimated by AUC. P < 0.05 was considered statistically significant.

AUTHOR CONTRIBUTIONS

RZ, HW, and FW conceived the study and edited the manuscript. PQ and MY were involved in the acquisition, analysis, interpretation of data, and composed the manuscript. HY, XY, WL, and ML performed the statistical analysis. QZ revised the manuscript. All authors approved the final version for publication.

CONFLICTS OF INTEREST

All authors declare that there are no conflicts of interest.

FUNDING

This project was supported by the Program of Excellent Doctoral (Postdoctoral) of Zhongnan Hospital of Wuhan University (Grant No. ZNYB2019003); National Science Foundation of China (No: 81870390); Hubei Province Key Laboratory Project (2015CFA153).

REFERENCES

- Yilmaz B, Juillerat P, Øyås O, Ramon C, Bravo FD, Franc Y, Fournier N, Michetti P, Mueller C, Geuking M, Pittet VEH, Maillard MH, Rogler G, and Swiss IBD Cohort Investigators, Wiest R, Stelling J, Macpherson AJ. Microbial network disturbances in relapsing refractory Crohn's disease. Nat Med. 2019; 25:323–36. <u>https://doi.org/10.1038/s41591-018-0308-z</u> PMID:<u>30664783</u>
- Goodman WA, Erkkila IP, Pizarro TT. Sex matters: impact on pathogenesis, presentation and treatment of inflammatory bowel disease. Nat Rev Gastroenterol Hepatol. 2020; 17:740–54. <u>https://doi.org/10.1038/s41575-020-0354-0</u> PMID:<u>32901108</u>
- Allocca M, Danese S, Laurent V, Peyrin-Biroulet L. Use of cross-sectional imaging for tight monitoring of inflammatory bowel diseases. Clin Gastroenterol Hepatol. 2020; 18:1309–23.e4. <u>https://doi.org/10.1016/j.cgh.2019.11.052</u> PMID:31812657
- Kane SV, Sandborn WJ, Rufo PA, Zholudev A, Boone J, Lyerly D, Camilleri M, Hanauer SB. Fecal lactoferrin is a sensitive and specific marker in identifying intestinal inflammation. Am J Gastroenterol. 2003; 98:1309–14. <u>https://doi.org/10.1111/j.1572-0241.2003.07458.x</u> PMID:<u>12818275</u>
- Tøn H, Brandsnes Ø, Dale S, Holtlund J, Skuibina E, Schjønsby H, Johne B. Improved assay for fecal calprotectin. Clin Chim Acta. 2000; 292:41–54. <u>https://doi.org/10.1016/s0009-8981(99)00206-5</u> PMID:<u>10686275</u>
- Panes J, Jairath V, Levesque BG. Advances in Use of Endoscopy, Radiology, and Biomarkers to Monitor Inflammatory Bowel Diseases. Gastroenterology. 2017; 152:362–73.e3. <u>https://doi.org/10.1053/j.gastro.2016.10.005</u> PMID:27751880
- Walsham NE, Sherwood RA. Fecal calprotectin in inflammatory bowel disease. Clin Exp Gastroenterol. 2016; 9:21–29. <u>https://doi.org/10.2147/CEG.S51902</u> PMID:26869808
- Dulai PS, Peyrin-Biroulet L, Danese S, Sands BE, Dignass A, Turner D, Mantzaris G, Schölmerich J, Mary JY, Reinisch W, Sandborn WJ. Approaches to integrating biomarkers into clinical trials and care pathways as

targets for the treatment of inflammatory bowel diseases. Gastroenterology. 2019; 157:1032–43.e1. https://doi.org/10.1053/j.gastro.2019.06.018 PMID:31228441

- Johnson DH, Taylor WR, Aboelsoud MM, Foote PH, Yab TC, Cao X, Smyrk TC, Loftus EV Jr, Mahoney DW, Ahlquist DA, Kisiel JB. DNA methylation and mutation of small colonic neoplasms in ulcerative colitis and Crohn's colitis: implications for surveillance. Inflamm Bowel Dis. 2016; 22:1559–67. <u>https://doi.org/10.1097/MIB.000000000000795</u> PMID:<u>27104828</u>
- Ghavami SB, Rostami E, Sephay AA, Shahrokh S, Balaii H, Aghdaei HA, Zali MR. Alterations of the human gut Methanobrevibacter smithii as a biomarker for inflammatory bowel diseases. Microb Pathog. 2018; 117:285–89. <u>https://doi.org/10.1016/j.micpath.2018.01.029</u>
 - PMID:29477743
- Klepp P, Kisiel JB, Småstuen MC, Røseth A, Andersen SN, Vatn MH, Ahlquist DA, Moum BA, Brackmann S. Multi-target stool DNA test in the surveillance of inflammatory bowel disease: a cross-sectional cohort study. Scand J Gastroenterol. 2018; 53:273–78. <u>https://doi.org/10.1080/00365521.2018.1424935</u> PMID:<u>29313389</u>
- Schaefer JS. MicroRNAs: how many in inflammatory bowel disease? Curr Opin Gastroenterol. 2016; 32:258–66. <u>https://doi.org/10.1097/MOG.00000000000284</u> PMID:<u>27138057</u>
- Koukos G, Polytarchou C, Kaplan JL, Oikonomopoulos A, Ziring D, Hommes DW, Wahed R, Kokkotou E, Pothoulakis C, Winter HS, Iliopoulos D. A microRNA signature in pediatric ulcerative colitis: deregulation of the miR-4284/CXCL5 pathway in the intestinal epithelium. Inflamm Bowel Dis. 2015; 21:996–1005. <u>https://doi.org/10.1097/MIB.00000000000339</u> PMID:25738378
- Chapman CG, Pekow J. The emerging role of miRNAs in inflammatory bowel disease: a review. Therap Adv Gastroenterol. 2015; 8:4–22. <u>https://doi.org/10.1177/1756283X14547360</u> PMID:<u>25553076</u>
- Wu F, Zikusoka M, Trindade A, Dassopoulos T, Harris ML, Bayless TM, Brant SR, Chakravarti S, Kwon JH. MicroRNAs are differentially expressed in ulcerative colitis and alter expression of macrophage inflammatory peptide-2 alpha. Gastroenterology. 2008; 135:1624–35.e24.

https://doi.org/10.1053/j.gastro.2008.07.068 PMID:<u>18835392</u>

- Thorlacius-Ussing G, Schnack Nielsen B, Andersen V, Holmstrøm K, Pedersen AE. Expression and localization of miR-21 and miR-126 in mucosal tissue from patients with inflammatory bowel disease. Inflamm Bowel Dis. 2017; 23:739–52. <u>https://doi.org/10.1097/MIB.000000000001086</u> PMID:28426456
- Rebane A, Akdis CA. MicroRNAs: essential players in the regulation of inflammation. J Allergy Clin Immunol. 2013; 132:15–26. <u>https://doi.org/10.1016/j.jaci.2013.04.011</u> PMID:<u>23726263</u>
- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004; 116:281–97. <u>https://doi.org/10.1016/s0092-8674(04)00045-5</u> PMID:<u>14744438</u>
- McKenna LB, Schug J, Vourekas A, McKenna JB, Bramswig NC, Friedman JR, Kaestner KH. MicroRNAs control intestinal epithelial differentiation, architecture, and barrier function. Gastroenterology. 2010; 139:1654–64. <u>https://doi.org/10.1053/j.gastro.2010.07.040</u> PMID:20659473
- Boros É, Nagy I. The role of MicroRNAs upon epithelialto-mesenchymal transition in inflammatory bowel disease. Cells. 2019; 8:1461. https://doi.org/10.3390/cells8111461 PMID:31752264
- Oikonomopoulos A, Polytarchou C, Joshi S, Hommes DW, Iliopoulos D. Identification of circulating MicroRNA signatures in Crohn's disease using the nanostring nCounter technology. Inflamm Bowel Dis. 2016; 22:2063–69. <u>https://doi.org/10.1097/MIB.00000000000883</u> PMID:<u>27542126</u>
- 22. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. Elife. 2015; 4:e05005. https://doi.org/10.7554/eLife.05005 PMID:26267216
- 23. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. PLoS One. 2018; 13:e0206239. <u>https://doi.org/10.1371/journal.pone.0206239</u> PMID:30335862
- 24. Wong N, Wang X. miRDB: an online resource for microRNA target prediction and functional annotations. Nucleic Acids Res. 2015; 43:D146–52. https://doi.org/10.1093/nar/gku1104 PMID:25378301
- 25. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: proteinprotein association networks with increased coverage, supporting functional discovery in genome-wide

experimental datasets. Nucleic Acids Res. 2019; 47:D607–13.

https://doi.org/10.1093/nar/gky1131 PMID:30476243

- 26. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol. 2014 (Suppl 4); 8:S11. https://doi.org/10.1186/1752-0509-8-S4-S11 PMID:25521941
- 27. Bandettini WP, Kellman P, Mancini C, Booker OJ, Vasu S, Leung SW, Wilson JR, Shanbhag SM, Chen MY, Arai AE. MultiContrast delayed enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. J Cardiovasc Magn Reson. 2012; 14:83. https://doi.org/10.1186/1532-429X-14-83 PMID:23199362
- Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 2019; 47:D590–95. https://doi.org/10.1093/nar/gky962 PMID:30321428
- Tian Y, Xu J, Li Y, Zhao R, Du S, Lv C, Wu W, Liu R, Sheng X, Song Y, Bi X, Li G, Li M, et al. MicroRNA-31 reduces inflammatory signaling and promotes regeneration in colon epithelium, and delivery of mimics in microspheres reduces colitis in mice. Gastroenterology. 2019; 156:2281–96.e6. https://doi.org/10.1053/j.gastro.2019.02.023 PMID:30779922
- Selcuklu SD, Donoghue MT, Spillane C. miR-21 as a key regulator of oncogenic processes. Biochem Soc Trans. 2009; 37:918–25. https://doi.org/10.1042/BST0370918 PMID:19614619
- 31. Khalili H, Chan SS, Lochhead P, Ananthakrishnan AN, Hart AR, Chan AT. The role of diet in the aetiopathogenesis of inflammatory bowel disease. Nat Rev Gastroenterol Hepatol. 2018; 15:525–35. <u>https://doi.org/10.1038/s41575-018-0022-9</u> PMID:<u>29789682</u>
- Ghafouri-Fard S, Eghtedarian R, Taheri M. The crucial role of non-coding RNAs in the pathophysiology of inflammatory bowel disease. Biomed Pharmacother. 2020; 129:110507. <u>https://doi.org/10.1016/j.biopha.2020.110507</u> PMID:32768980
- 33. GBD 2017 Inflammatory Bowel Disease Collaborators. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 2020; 5:17–30.

https://doi.org/10.1016/S2468-1253(19)30333-4 PMID:<u>31648971</u>

- 34. Kim D, Nguyen QT, Lee J, Lee SH, Janocha A, Kim S, Le HT, Dvorina N, Weiss K, Cameron MJ, Asosingh K, Erzurum SC, Baldwin WM 3rd, et al. Anti-inflammatory roles of glucocorticoids are mediated by Foxp3⁺ regulatory T cells via a miR-342-dependent mechanism. Immunity. 2020; 53:581–96.e5. <u>https://doi.org/10.1016/j.immuni.2020.07.002</u> PMID:<u>32707034</u>
- Zhang T, Hu J, Wang X, Zhao X, Li Z, Niu J, Steer CJ, Zheng G, Song G. MicroRNA-378 promotes hepatic inflammation and fibrosis via modulation of the NF-κB-TNFα pathway. J Hepatol. 2019; 70:87–96. <u>https://doi.org/10.1016/j.jhep.2018.08.026</u> PMID:<u>30218679</u>
- Seeley JJ, Baker RG, Mohamed G, Bruns T, Hayden MS, Deshmukh SD, Freedberg DE, Ghosh S. Induction of innate immune memory via microRNA targeting of chromatin remodelling factors. Nature. 2018; 559:114–19. https://doi.org/10.1038/s41586-018-0253-5

PMID:29950719

- Fisher K, Lin J. MicroRNA in inflammatory bowel disease: translational research and clinical implication. World J Gastroenterol. 2015; 21:12274–82. <u>https://doi.org/10.3748/wjg.v21.i43.12274</u> PMID:<u>26604636</u>
- Kalla R, Adams AT, Ventham NT, Kennedy NA, White R, Clarke C, Ivens A, Bergemalm D, Vatn S, Lopez-Jimena B, Ricanek P, Vatn MH, Söderholm J, et al, and IBD Character Consortium. Whole blood profiling of T-cell derived miRNA allows the development of prognostic models in inflammatory bowel disease. J Crohns Colitis. 2020; 14:1724–33. https://doi.org/10.1093/ecco-jcc/jjaa134

PMID:<u>32598439</u>

- 39. Fang B, Guo P, Yang M, Ma Y, Yan X, Jia Z, Gao W, Ahmad S, Xu C, Liu C, Zhang M. A novel fluorescent enhancing platform based on DNA-scaffolded silver nanoclusters for potential inflammatory bowel disease-associated microRNA detection. Talanta. 2020; 218:121122. <u>https://doi.org/10.1016/j.talanta.2020.121122</u> PMID:32797879
- Liu S, da Cunha AP, Rezende RM, Cialic R, Wei Z, Bry L, Comstock LE, Gandhi R, Weiner HL. The host shapes the gut microbiota via fecal MicroRNA. Cell Host Microbe. 2016; 19:32–43. <u>https://doi.org/10.1016/j.chom.2015.12.005</u> PMID:26764595
- 41. Iborra M, Bernuzzi F, Correale C, Vetrano S, Fiorino G,

Beltrán B, Marabita F, Locati M, Spinelli A, Nos P, Invernizzi P, Danese S. Identification of serum and tissue micro-RNA expression profiles in different stages of inflammatory bowel disease. Clin Exp Immunol. 2013; 173:250–58.

https://doi.org/10.1111/cei.12104 PMID:23607522

- Mohammadi A, Kelly OB, Filice M, Kabakchiev B, Smith MI, Silverberg MS. Differential expression of microRNAs in peripheral blood mononuclear cells identifies autophagy and TGF-beta-related signatures aberrantly expressed in inflammatory bowel disease. J Crohns Colitis. 2018; 12:568–81. <u>https://doi.org/10.1093/ecco-jcc/jjy010</u> PMID:29420705
- Chen P, Li Y, Li L, Yu Q, Chao K, Zhou G, Qiu Y, Feng R, Huang S, He Y, Chen B, Chen M, Zeng Z, Zhang S. Circulating microRNA146b-5p is superior to C-reactive protein as a novel biomarker for monitoring inflammatory bowel disease. Aliment Pharmacol Ther. 2019; 49:733–43.

https://doi.org/10.1111/apt.15159 PMID:30734320

- 44. Soubières AA, Poullis A. Emerging role of novel biomarkers in the diagnosis of inflammatory bowel disease. World J Gastrointest Pharmacol Ther. 2016; 7:41–50. https://doi.org/10.4292/wjgpt.v7.i1.41
 PMID:26855811
- 45. Ge Y, Sun M, Wu W, Ma C, Zhang C, He C, Li J, Cong Y, Zhang D, Liu Z. MicroRNA-125a suppresses intestinal mucosal inflammation through targeting ETS-1 in patients with inflammatory bowel diseases. J Autoimmun. 2019; 101:109–20. <u>https://doi.org/10.1016/j.jaut.2019.04.014</u> PMID:31014918
- Link A, Goel A. MicroRNA in gastrointestinal cancer: a step closer to reality. Adv Clin Chem. 2013; 62:221–68. <u>https://doi.org/10.1016/b978-0-12-800096-0.00006-8</u> PMID:<u>24772669</u>
- 47. Verdier J, Breunig IR, Ohse MC, Roubrocks S, Kleinfeld S, Roy S, Streetz K, Trautwein C, Roderburg C, Sellge G. Faecal micro-RNAs in inflammatory bowel diseases. J Crohns Colitis. 2020; 14:110–17. https://doi.org/10.1093/ecco-jcc/jjz120 PMID:31209454
- Duran-Sanchon S, Moreno L, Augé JM, Serra-Burriel M, Cuatrecasas M, Moreira L, Martín A, Serradesanferm A, Pozo À, Costa R, Lacy A, Pellisé M, Lozano JJ, et al. Identification and validation of MicroRNA profiles in fecal samples for detection of colorectal cancer. Gastroenterology. 2020; 158:947–57.e4. <u>https://doi.org/10.1053/j.gastro.2019.10.005</u>

- 49. Ostenfeld MS, Jensen SG, Jeppesen DK, Christensen LL, Thorsen SB, Stenvang J, Hvam ML, Thomsen A, Mouritzen P, Rasmussen MH, Nielsen HJ, Ørntoft TF, Andersen CL. miRNA profiling of circulating EpCAM(+) extracellular vesicles: promising biomarkers of colorectal cancer. J Extracell Vesicles. 2016; 5:31488. https://doi.org/10.3402/jev.v5.31488 PMID:27576678
- 50. Wu H, Wei M, Jiang X, Tan J, Xu W, Fan X, Zhang R, Ding C, Zhao F, Shao X, Zhang Z, Shi R, Zhang W, Wu G. IncRNA PVT1 promotes tumorigenesis of colorectal cancer by stabilizing miR-16-5p and interacting with the VEGFA/VEGFR1/AKT axis. Mol Ther Nucleic Acids. 2020; 20:438–50. https://doi.org/10.1016/j.omtn.2020.03.006
- Choi CR, Bakir IA, Hart AL, Graham TA. Clonal evolution of colorectal cancer in IBD. Nat Rev Gastroenterol Hepatol. 2017; 14:218–29. <u>https://doi.org/10.1038/nrgastro.2017.1</u> PMID:28174420

PMID:32276209

- 52. Krichevsky AM, Gabriely G. miR-21: a small multifaceted RNA. J Cell Mol Med. 2009; 13:39–53. <u>https://doi.org/10.1111/j.1582-4934.2008.00556.x</u> PMID:<u>19175699</u>
- 53. Matos LL, Menderico GM Jr, Theodoro TR, Pasini FS, Ishikawa MM, Ribeiro AAB, de Mello ES, Pinhal MA, Moyses RA, Kulcsar MAV, Dedivitis RA, Cernea CR, Kowalski LP. Cancer-associated fibroblast regulation by microRNAs promotes invasion of oral squamous cell carcinoma. Oral Oncol. 2020; 110:104909. <u>https://doi.org/10.1016/j.oraloncology.2020.104909</u> PMID:32702628
- 54. Jin XH, Lu S, Wang AF. Expression and clinical significance of miR-4516 and miR-21-5p in serum of patients with colorectal cancer. BMC Cancer. 2020; 20:241.

https://doi.org/10.1186/s12885-020-06715-6 PMID:<u>32293319</u>

55. Arhi C, Askari A, Nachiappan S, Bottle A, Arebi N, Athanasiou A, Ziprin P, Aylin P, Faiz O. Stage at diagnosis and survival of colorectal cancer with or without underlying inflammatory bowel disease: a population-based study. J Crohns Colitis. 2020. [Epub ahead of print]. https://doi.org/10.1093/ecco-jcc/jjaa196

```
PMID:32991688
```

56. Song R, Li Y, Hao W, Wang B, Yang L, Xu F. Identification and analysis of key genes associated with ulcerative colitis based on DNA microarray data. Medicine (Baltimore). 2018; 97:e10658. <u>https://doi.org/10.1097/MD.000000000010658</u> PMID:<u>29794741</u>

- 57. Bilal MY, Houtman JCD. Transmission of T Cell Receptor-Mediated Signaling via the GRB2 Family of Adaptor Proteins. In: Soboloff J, Kappes DJ, editors. Signaling Mechanisms Regulating T Cell Diversity and Function. Boca Raton (FL): CRC Press/Taylor & Francis; 2018. Chapter 9. PMID:30427631
- Wang Z, Vaughan TY, Zhu W, Chen Y, Fu G, Medrzycki M, Nishio H, Bunting ST, Hankey-Giblin PA, Nusrat A, Parkos CA, Wang D, Wen R, Bunting KD. Gab2 and Gab3 redundantly suppress colitis by modulating macrophage and CD8⁺ T-cell activation. Front Immunol. 2019; 10:486.

https://doi.org/10.3389/fimmu.2019.00486 PMID:<u>30936879</u>

- 59. Wang W, Xiao X, Chen X, Huo Y, Xi WJ, Lin ZF, Zhang D, Li YF, Yang F, Wen WH, Yang AG, Wang T. Tumorsuppressive miR-145 co-repressed by TCF4-β-catenin and PRC2 complexes forms double-negative regulation loops with its negative regulators in colorectal cancer. Int J Cancer. 2018; 142:308–21. https://doi.org/10.1002/ijc.31056 PMID:28921929
- Yan P, Wang Y, Meng X, Yang H, Liu Z, Qian J, Zhou W, Li J. Whole exome sequencing of ulcerative colitisassociated colorectal cancer based on novel somatic mutations identified in Chinese patients. Inflamm Bowel Dis. 2019; 25:1293–301. <u>https://doi.org/10.1093/ibd/izz020</u> PMID:30794281
- Klupp F, Giese C, Halama N, Franz C, Lasitschka F, Warth A, Schmidt T, Kloor M, Ulrich A, Schneider M. E3 ubiquitin ligase Smurf2: a prognostic factor in microsatellite stable colorectal cancer. Cancer Manag Res. 2019; 11:1795–803. <u>https://doi.org/10.2147/CMAR.S178111</u>
 - PMID:30863185
- Chen JJ, Zhang W. High expression of WWP1 predicts poor prognosis and associates with tumor progression in human colorectal cancer. Am J Cancer Res. 2018; 8:256–65.
 PMID:29511596
- Nadeem MS, Kumar V, Al-Abbasi FA, Kamal MA, Anwar F. Risk of colorectal cancer in inflammatory bowel diseases. Semin Cancer Biol. 2020; 64:51–60. <u>https://doi.org/10.1016/j.semcancer.2019.05.001</u> PMID:31112753
- Rutter MD, Saunders BP, Wilkinson KH, Rumbles S, Schofield G, Kamm MA, Williams CB, Price AB, Talbot IC, Forbes A. Thirty-year analysis of a colonoscopic surveillance program for neoplasia in ulcerative colitis. Gastroenterology. 2006; 130:1030–38. <u>https://doi.org/10.1053/j.gastro.2005.12.035</u> PMID:<u>16618396</u>

- 65. Aardoom MA, Joosse ME, de Vries AC, Levine A, de Ridder L. Malignancy and mortality in pediatric-onset inflammatory bowel disease: a systematic review. Inflamm Bowel Dis. 2018; 24:732–41. <u>https://doi.org/10.1093/ibd/izx104</u> PMID:<u>29522170</u>
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–504.

https://doi.org/10.1101/gr.1239303 PMID:14597658

67. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, et al, and The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000; 25:25–29.

https://doi.org/10.1038/75556 PMID:10802651