Machine-learning-based Analysis Identifies miRNA Expression Profile for Diagnosis and Prediction of Colorectal Cancer: A Preliminary Study

DOROTA PAWELKA[1#], IZABELA LACZMANSKA[2#], PAWEL KARPINSKI[2,3],
STANISLAW SUPPLITT[2], WOJCIECH WITKIEWICZ[4], BARŁOMIEJ KNYCHALSKI[1],
JOANNA PELAK[5], PAULINA ZEBROWSKA[3] and LUKASZ LACZMANSKI[3]

[1]Department of Surgery Teaching, Wroclaw Medical University, Wroclaw, Poland;
[2]Department of Genetics, Wroclaw Medical University, Wroclaw, Poland;
[3]Laboratory of Genomics and Bioinformatics, Hirszfeld Institute of Immunology and Experimental Therapy,
Polish Academy of Sciences, Wroclaw, Poland;
[4]Research and Development Center of Lower Silesian Regional Specialist Hospital, Wroclaw, Poland;
[5]The Copper Health Center, Sklodowskiej-Curie, Lubin, Poland

Abstract. *Background: The stage of colorectal cancer (CRC) at the day of diagnosis has the greatest influence on survival rate. Thus, for CRC, which is mainly identified as advanced disease, non-invasive, molecular blood or stool tests could boost the diagnosis and lower mortality. Evaluation of miRNA expression levels in serum of patients diagnosed with CRC is a potential tool in early screening. Screening can be supported by machine learning (ML) as a tool for developing a cancer risk predictive model based on genetic data. Materials and Methods: miRNA was isolated from the serum of 8 patients diagnosed with CRC and 10 patients from a control group matched for age and sex. The expression of 179 miRNAs was determined using a serum/plasma panel (Exiqon). Determinations were conducted using real-time PCR technique on an Applied Biosystems QuantStudio3 device in 96-well plates. A predictive model was developed through the Azure Machine Learning platform. Results: A wide panel of 29 up-regulated miRNAs in CRC were identified and divided into two subgroups: 1) miRNAs with significantly higher serum level in cancer patients vs. controls (24 miRNAs) and 2) miRNAs detected only in cancer patients and not in controls (5 miRNAs). Re-analysis of published miRNA profiles of CRC tumours or CRC exosomes revealed that only 2 out of 29 miRNAs were up-regulated in all datasets including ours (miR-34a and miR-25-3p). Conclusion: Our research suggests the potential role of overexpressed miRNAs as diagnostic or prognostic biomarkers among CRC patients. Such clustering of miRNAs may be a potential direction for discovering new diagnostic panels of cancer (including CRC), especially using ML. The low correspondence between deregulation of miRNAs in serum and tumour tissue revealed in our study confirms previously published reports.*

[#]These Authors contributed equally to this study.

Correspondence to: Izabela Laczmanska, Department of Genetics, Wroclaw Medical University, Marcinkowskiego 1, 50-368 Wroclaw, Poland. Tel.: +48 0717841258, Fax: +48 0717840063, e-mail: izabela.laczmanska@umw.edu.pl

Colorectal cancer (CRC) is one of the most common cancers and the third leading cause of cancer-related deaths, especially in developed countries (1). Morbidity from CRC is strongly related to certain environmental factors, such as low-fibre diet, addictions (alcohol consumption and smoking) and lack of physical activity (2, 3). Because of the high frequency and impact of CRC on general public health, the studies on its genetic background include searching early genetic markers of this cancer at the chromosomal, molecular, and epigenetic levels (4, 5). Although some non-invasive techniques for CRC screening exist, colonoscopy still remains the gold standard (6). As some patients avoid or refuse this procedure, establishing a new genetic screening test from blood would be greatly beneficial for CRC prevention and early diagnosis.

The stage of CRC at the day of diagnosis plays the largest role in survival rates. Notably, when CRC is diagnosed at a localized stage, the 5-year survival rate ranges between 89%

Table I. *Clinicopathological characteristics of patients.*

No	Age	Sex	TNM	Grading	Tumour size (cm)	Localization	Histopathology	CEA (ng/ml)
1	54	M	T3N0M0	G1	3×4	Sigmoid	Adenocarcinoma	5.69
2	43	M	T2N0M0	G2	3	Sigmoid	Adenocarcinoma	1.58
3	77	M	T3N0M0	G2	4	Rectum	Adenocarcinoma	0.3
4	48	F	T3N0M0	G2	8	Ascending colon	Adenocarcinoma	0.5
5	58	F	T3N3M0	G2	5	Caecum	Adenocarcinoma	1.2
6	71	M	T4N3M0	G2	>10	Rectum	Adenocarcinoma	8.67
7	68	F	T3N2M0	G2	8	Caecum	Adenocarcinoma	99.37
8	56	M	T4N3M0	G3	>15	Rectum	Adenocarcinoma	33.19

CEA: Carcinoembryonic antigen; M: male; F: female; TNM: tumour, node, metastasis.

and 92%. Overall survival is significantly shorter when CRC is diagnosed at an advanced stage (7). Thus, for CRC, which is mainly diagnosed after noting specific symptoms of advanced disease (*e.g.* blood in stool and then by biopsy during colonoscopy), non-invasive molecular tests from blood or stool could influence the percentage of early stage recognition and lower the mortality (5, 8).

The role of microRNA (miRNA) in initiation, development and metastasis is widely described beside the known genetic background in various cancers including colorectal cancer (3, 4, 9-12). MiRNAs are small 19- to 25-nucleotide noncoding RNAs that by binding to target messenger RNAs (mRNAs) can regulate the expression of different genes. They can act either as oncomirs which are responsible for down-regulation of tumour suppressor genes or as suppressor miRNAs down-regulating oncogenes (3). It is estimated that about one-third of human genes are regulated in this way. Therefore for cancers, where this regulation is aberrant, miRNAs are good potential biomarkers suitable for patient screening and diagnosis even in early stages (3, 13-15).

The role of some miRNAs affecting genetic pathways involved in oncogenesis of CRC has been previously identified (3). It was revealed that the miR-135 family and miR-34 family were involved in the Wnt signalling pathway together with miR-145, miR-29b and miR-146a (16, 17). Also, miR-224 was reported to activate this pathway (18). Another signalling pathway crucial for CRC, the EGFR pathway, can be activated in different ways by the miRNAs let 7, miR-143, miR-145, miR-18a, miR-126, miR-21, miR-32, miR-92a and miR-181a (19, 20). Moreover, miR-520a and miR525a acting on *PIK3CA* can modulate activation of Akt-dependent signalling (21). Also *KRAS* expression can be inhibited by let-7, miR18a, miR30b, miR143 and miR-145 (3, 11). For the TGF-β signalling pathway miR-21 and miR-135b and also miR106a and miR301a were reported to play an important role while changes in expression of miR-25 and miR187 influenced this pathway *via* the *SMAD7* gene (22,

23). Additionally, for *TP53* deficiency the miR-34 family, miR-125b and miR-215 were reported to be dysregulated, promoting cancer cell growth and proliferation (3, 11).

Machine learning, a subdiscipline of artificial intelligence, has been previously described as a useful tool for analysis of large, complex data sets (24, 25). Computer algorithms involved in machine learning are expected to improve with experience. In the field of genomics, the method can be used, for example, to recognize the transcription start sites in genome sequence or to analyse the miRNA expression patterns (26, 27). Azure Machine Learning (Azure ML, Microsoft, Redmond, WA, USA; https://azure.microsoft.com/en-us/services/genomics/) is a cloud service that enables the machine learning process to be performed. This website provides a workspace that helps users build and test predictive models.

The aim of our study was to evaluate the level of expression of a miRNA panel in serum of patients diagnosed with CRC in comparison to serum of healthy patients and to find potential and applicative biomarkers.

Materials and Methods

Genetic analysis. miRNA was isolated from serum of 8 patients diagnosed with CRC (Table I) and 10 patients from a control group matched for age and sex. The study was accepted by the Ethical Committee of Wroclaw Medical University (approval number 644/2014).

Serum samples from study and control groups were used for total RNA isolation (using EXIQON miRCURY, Qiagen, Hilden, Germany). The control of extraction efficacy was checked with two internal standards – UniSp2 and UniSp4 – but the cDNA synthesis control was checked with UniSp6. The signal of listed controls (Cp) over 37 was an indicator of a weak yield during material preparation. To assess the level of possible haemolysis in the material and the impact of reaction inhibitors in each sample, the levels of miR-451 and miR-23a expression were determined. Samples with expression of tested isoforms – ΔCp (23a-451) above 7 were rejected from further analysis due to high haemolysis. Expression of 179 circulating miRNA isoforms in serum was determined using a kit provided by EXIQON (Serum/plasma
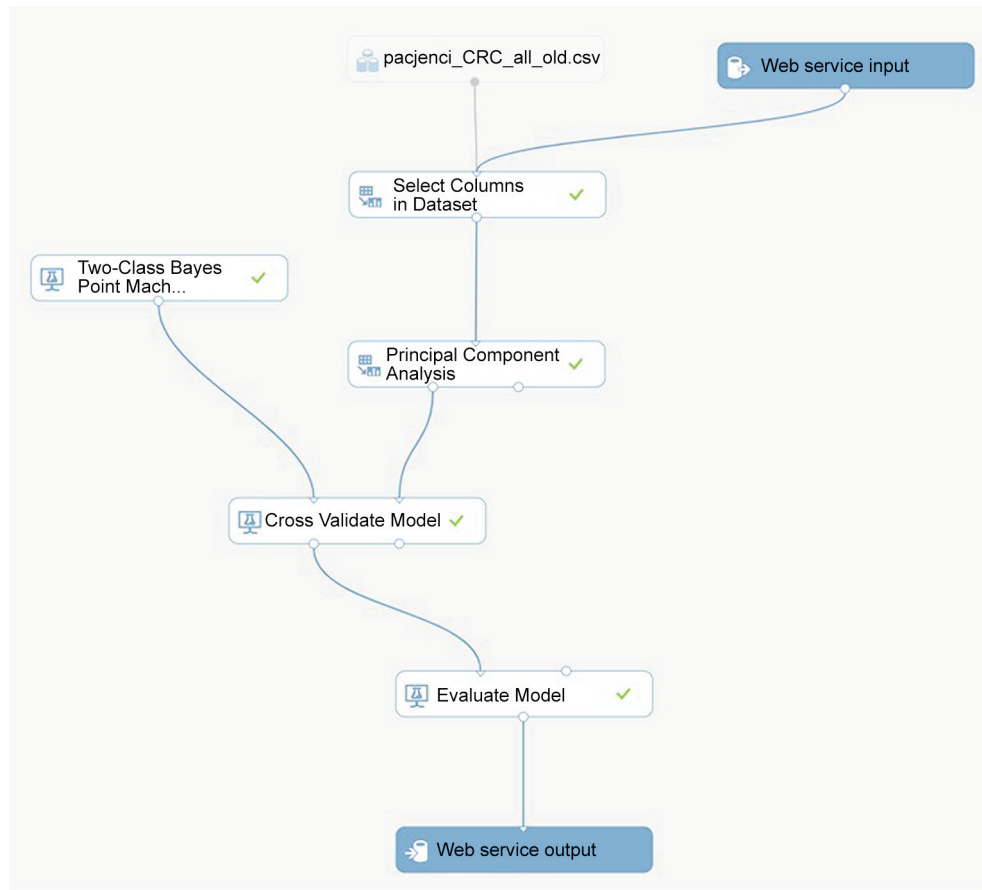
Figure 1. *Workflow of modelling using Azure ML service showing the steps of bioinformatics analysis using Two-Class Bayes Point Match machine-learning algorithm.*

miRNA miRCURY LNA, Qiagen). Determinations were conducted using real-time PCR technique on an Applied Biosystems QuantStudio3 device in 96-well plates (Thermo Fisher Scientific, Waltham, MA, USA). Expression profile in cancer samples compared to expression levels of miRNAs in healthy individuals' serum were calculated by:

$$\Delta Cp = Cp_{cancer} - Cp_{control}$$

From the obtained log ratio score (LR) a base 2 logarithm was calculated. LR values over 1 (overexpression) and lower than -1 (suppression) were taken into consideration in further analysis.

Machine learning. We built a supervised classification learning-machine model using Azure Machine Learning (Microsoft, Redmond, WA, USA) (Figure 1). First, to reduce data dimension we used Principal Component Analysis to select 20 miRNAs, for which we observed a statistically significant difference in expression levels between the control and cancer groups. Subsequently, normalisation of the columns of dataset to zero mean was performed.

We constructed predictive models using various two-class algorithms, such as Decision Forest, Decision Jungle, Bayes Point Machine, Support Vector Machine, and Neural Network. After accuracy and precision analysis we decided to train our set using the Bayes Point Machine method for tumour prediction based on the expression of selected serum miRNAs. Bayes Point Machine was run with 30 iterations, bias inclusion and allowance of unknown values. Small sample size limited our ability to split data into training and test sets, therefore, we used the cross validation algorithm (10 folds were used) which is dedicated to evaluate machine learning models on a limited data sample. The data set was randomly divided into 10 subsets of equal size. One subset was selected for each iteration and was used to test the trained model on 9 subsets (Figure 1). The experiment has been published to link: https://gallery.cortanaintelligence.com/Experiment/New-CRC-miRNA

miRNA data re-analysis. Data on previously published miRNA expression profiles of CRC and healthy adjacent tissues were obtained from Gene Expression Omnibus (GEO), accession numbers: GSE35834 and GSE41655 (28). GSE35834 included 23 normal tissues and 31 tumours, whereas GSE41655 included 15 normal tissues and 33 tumours. In addition we re-analysed GSE40247 including miRNA profiles obtained from serum exosomes from 11 healthy volunteers and 68 CRC patients (we excluded individuals with stage I of disease) (29). To assess which

Table II. *Two groups of up-regulated miRNAs in analysed CRC samples (p<0.05).*

Group	miRNA	Mann-Whitney test	*p*-Value
miRNAs overexpressed	19b-3p	2.632	0.0040
	22-3p	1.953	0.0484
	144-3p	2.633	0.0041
	19a-3p	2.631	0.0042
	103a-3p	2.463	0.0081
	151a-5p	2.123	0.0283
	let-7i-5p	2.347	0.0121
	146a-5p	2.293	0.0162
	30e-5p	2.632	0.0041
	16-5p	2.551	0.0061
	25-3p	2.633	0.0041
	451a	2.293	0.0162
	let-7g-5p	2.123	0.0283
	125b-5p	2.293	0.0162
	185-5p	2.347	0.0121
	106a-5p	2.632	0.0040
	29c-3p	2.051	0.0333
	324-3p	2.143	0.0242
	374a-5p	2.144	0.0242
	199a-5p	2.347	0.0121
	210-3p	2.348	0.0121
	130a-3p	2.195	0.0238
	15a-5p	2.347	0.0121
	let-7d-5p	2.051	0.0333
miRNAs expressed only in cancer patients	34a		
	7-1-3p		
	543		
	335-3p		
	136-3p		

miRNAs were differentially expressed we first limited data to intersection with 29 miRNAs that were up-regulated in our initial study (see Results). Subsequently we employed non parametric ANOVA type analysis using the npmv 2.4.0 R package to assess the significance of differences. A plot was generated using the nmf 0.23.0 R package.

Results

Genetic analysis. Congruent with previously published studies, our findings revealed the up-regulation of dozens of miRNAs in CRC-diagnosed patients' serum, in comparison to healthy cases. We identified a wide panel of 29 up-regulated miRNAs in CRC divided into two subgroups: 24 of statistically significantly higher serum level in cancer patients and 5 detected repeatedly only in cancer patients compared to the control group (full list in Table II).

miRNA data re-analysis. Results of previously published miRNA data re-analysis are visualized in Figure 2. Out of 29 up-regulated miRNAs (see Table II) 23 miRNAs were present in all three re-analysed datasets. Significant up-regulation in CRC was detected in all three datasets for miR-34a and miR-25-3p. We also detected significant down-regulation of several miRNAs in colorectal tumours including miR-324-3p, miR-125b-5p, miR-199a-5p, miR-29c-3p and miR-30e-5p. Interestingly, three of these miRNAs (miR-125b-5p, miR-199a-5p and miR-29c-3p) were up-regulated in serum exosomes of CRC patients (see Figure 2, GSE40247). The remaining 16 miRNAs (out of 23) displayed no significant changes in gene expression between normal and CRC status.

Machine learning. We developed a predictive model with 19 cases through the Azure ML platform (Figure 1). First, we tested a model using various algorithms, such as Two-class Decision Forest, Two-class Decision Jungle, Two-class Bayes Point Machine, Two-class Support Vector Machine, and Two-class Neural Network. Two-class Bayes Point Machine showed the best accuracy and precision, which were respectively 0.947 and 0.917. The AUC of the ROC curve was 0.955 (Figure 3). The threshold has been set to the optimum cut-off value of 0.4.

Discussion

Recently, the dysregulated expression of various miRNAs was highlighted in terms of CRC diagnostics (30). It has been shown that miRNAs contribute to the initiation and progression of numerous oncogenic molecular events (10). The miRNA's involvement in CRC development may be bidirectional: tumour-promoting or tumour-suppressing (31). It means that miRNAs can induce cell proliferation and inhibit apoptosis – leading to tumour initiation – or act as suppressors through reducing tumour susceptibility to metastasis and invasion. Transcripts up-regulated in cancers, called oncogenic miRNAs (oncomirs), usually inhibit the expression of tumour-suppressor genes and stimulate carcinogenesis (31). Thus, miRNA may be utilized as clinical cancer biomarkers (CRC included), since they exhibit high tissue specificity, altered expression in tumour cells and relative stability (32). MiRNAs which are expressed significantly more highly or only in CRC patients, can be used in the early detection of this cancer in non-invasive screening tests (33, 34). Moreover, miRNA-oriented diagnostics offers flexibility in choice of clinical material, because it can be obtained from tumour tissue, the patient's blood serum or stool. Examples of widely described CRC oncomirs are miR-21, 92a, 96, 135a/b, 155, 224, 214, 31, 210, 182/503, 200c, 301a (10, 35). Well-identified tumour-suppressive CRC miRNAs include let-7, miR-194, 143/145, 34a, 126, 27b, 7, 18a-3p, 26b, 101, 144, 320a, 330, 455, 149 (3, 10). Nevertheless, the function of multiple miRNAs can be ambiguous, and some transcripts are described as both oncomirs and tumour suppressors, such as mir155 (36, 37).
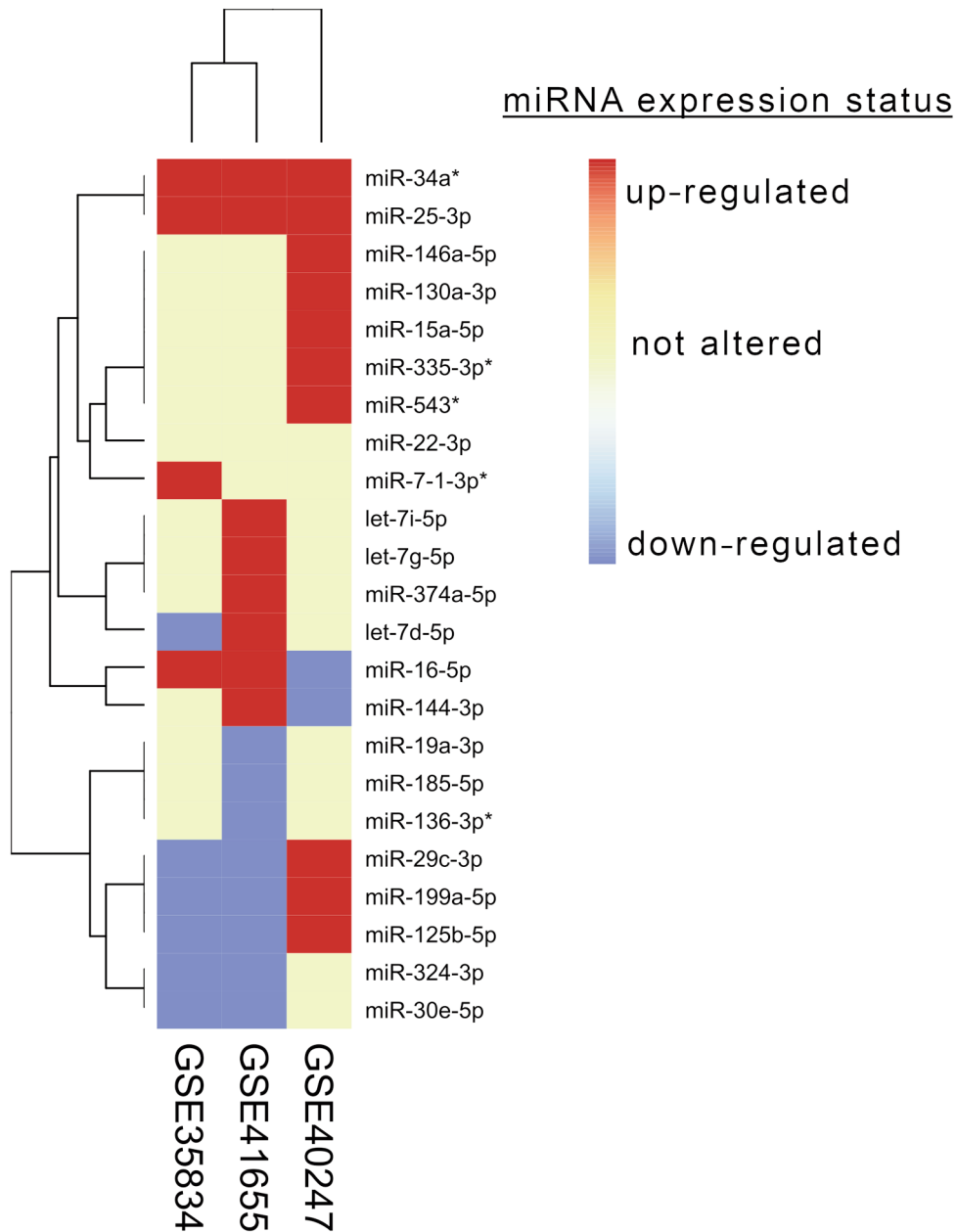
Figure 2. *Heatmap illustrating the results of re-analysis of three miRNA datasets using nonparametric ANOVA. Stars indicate miRNAs that were detected exclusively in the serum of CRC patients in our study. Note that GSE40247 includes miRNA profiles obtained from serum exosomes.*

Most identified miRNAs mediate in various patho-physiological conditions, such as cancer (38), chronic inflammatory diseases (39, 40), osteoporosis and diabetes (41), through interfering in different molecular and metabolic pathways: targeting tumour-suppressor genes and activating oncogenic transcription factors (42), affecting cell cycle arrest, pluripotency (43), glucose metabolism (44) and angiogenesis (45). Some miRNAs have already been associated with CRC

risk, patient prognosis or treatment outcomes (11). Interestingly, many sources indicate that above-mentioned CRC miRNAs present contrasting functions.

Our research also suggests the potential role of overexpressed miRNAs as a diagnostic or prognostic panel among CRC patients. We revealed a group of 29 miRNAs which present significantly higher expression in CRC patients than in the healthy population. Compared to other

A

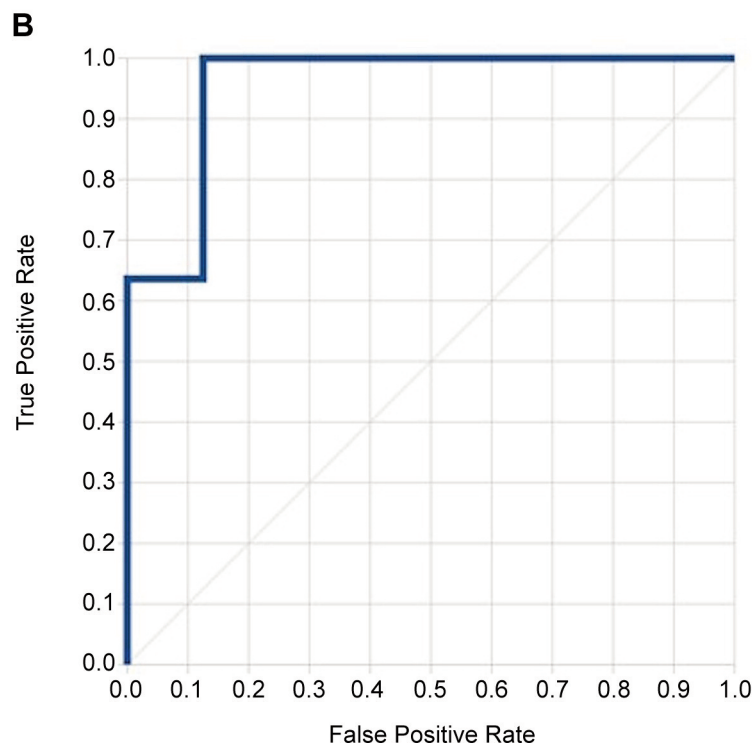|						
True Positive	False Negative	Accuracy	Precision	Threshold		AUC
11	0	0.947	0.917	0.4		0.955
False Positive	True Negative	Recall	F1 Score			
1	7	1.000	0.957			
Positive Label	Negative Label					
Control	Cancer					

Figure 3. *Results of the validation of Two-class Bayes Point Machine. (A) Accuracy, Precision, Recall, F1 Score and Threshold. (B) ROC curve.*

studies, our project developed a wider panel of 29 overexpressed miRNAs, found regularly in every analysed cancerous sample, with high precision and a high accuracy rate – up to 0.947. In the study of Kanaan *et al.* a panel of 8 miRNAs (miR-532-3p, miR-331, miR-195, miR-17, miR-142-3p, miR-15b, miR-532, and miR-652) was detected with a lower accuracy rate of 0.868 (46). In the project of Tan *et al.* a panel of 3 miRNAs (miR-144-3p, miR-425-5p, and miR-1260b) was identified with sensitivity and specificity of 93.8% and 91.3%, respectively, while Zhang *et al.* detected a seven-miRNA signature for CRC diagnosis (miR-103a-3p, miR-127-3p, miR-151a-5p, miR-17-5p, miR-181a-5p, miR-

18a-5p and miR-18b-5p) with an accuracy up to 0.895 (47, 48). Thus, our study developed an expanded and specific potential model for CRC diagnostics. Moreover, all the above-mentioned projects are based mostly on overexpression or dysregulated expression of various miRNAs, while our panel combines overexpressed CRC miRNAs with miRs detectable only in cancer patients.

Despite the overexpression in analysed samples, it has been reported that some of the defined miRNAs may act as CRC oncomirs (such as miR-19a-3p) and tumour suppressors as well [such as miR-34a (49) or let-7 (50)]. Similarly to other research related to CRC miRNA panels, we show the overexpression of

the following transcripts: miR-144-3p (47), 19a-3p (51), miR-103a-3p, miR-151a-5p (48). Compared to other studies, our project revealed a wider panel of 29 overexpressed miRNAs, found regularly in every analysed sample.

We also proposed a predictive model using machine learning, which could be a potentially useful and readily available tool for diagnosing a colorectal cancer risk group using serum miRNA tests. Using machine learning, it is possible to analyse many biological variables that may affect the process of cancer formation (52). On the basis of the entered data, the tool creates an *in silico* model of the probability of cancer occurrence according to a given biological profile, in our case a specific serum miRNA panel. To assess the suitability of this model, it is necessary to test a large group of people.

Finally, by re-analysing published miRNA profiles of CRC tumours or CRC exosomes we revealed that only miR-34a and miR-25-3p were up-regulated in all datasets including ours. Such low correspondence between miRNA deregulation serum and tumour tissue has been previously reported by others. For example, Zhu *et al*. found that only 10 out of ~100 deregulated miRNAs are shared between serum and corresponding breast tumours (53). Recently, Gmerek *et al*. (2019) found no overlap in miRNAs regulated in CRC tissue and serum of CRC patients. This suggests that serum deregulated miRNAs may not be directly secreted from colorectal tumour cells but rather corresponds to other cancer-related conditions, *i.e*. systemic inflammation and oxidative stress (54).

The small sample size in the discovery dataset limited our ability to split data into training and test sets, which is the limitation that this study hopes to partially mitigate through the application of cross-validation. Another limitation is that the public miRNA data sets included in this analysis were profiled from different platforms which may lead to poor extrapolation of the findings.

Conclusion

Our preliminary results show that numerous miRNAs of various functions discovered to date can be overexpressed in CRC. As described, the identified panel consists of multiple miRNAs presenting different spectra of molecular involvement, which together appear to be a predictive panel with high diagnostic efficiency for CRC. To confirm our results the study should be continued on a larger group of samples. However, even for the present group the analysis clearly indicated repeatability of co-incidence of detected miRNAs. The possible detectability of above-mentioned miRNAs in CRC may be related to interference or functional synergism of individual miRNAs in the mediation of carcinogenesis. Such clustering of miRNAs may be promising for discovering new diagnostic panels of cancer, including CRC.

Conflicts of Interest

The Authors declare no conflicts of interest.

Authors' Contributions

IL, LL, PK and PZ carried out the molecular genetic studies. DP, IL, LL and SS drafted the manuscript. DP, PK, JP and BK created a database including the clinical data of patients. DP, IL and LL conceived of the study, participated in its design and coordination. LL and PK prepared bioinformatic analyses. WW provided financial support. All authors read and approved the final manuscript.

References

1   Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin *68(6)*: 394-424, 2018. PMID: 30207593. DOI: 10.3322/caac.21492

2   Hsing AW, McLaughlin JK, Chow WH, Schuman LM, Co Chien HT, Gridley G, Bjelke E, Wacholder S and Blot WJ: Risk factors for colorectal cancer in a prospective study among U.S. white men. Int J Cancer *77(4)*: 549-553, 1998. PMID: 9679757. DOI: 10.1002/(sici)1097-0215(19980812)77:4<549::aid-ijc13>3.0.co;2-1

3   Shirafkan N, Mansoori B, Mohammadi A, Shomali N, Ghasbi M and Baradaran B: MicroRNAs as novel biomarkers for colorectal cancer: New outlooks. Biomed Pharmacother *97*: 1319-1330, 2018. PMID: 29156521. DOI: 10.1016/j.biopha.2017.11.046

4   Yang S, Sun Z, Zhou Q, Wang W, Wang G, Song J, Li Z, Zhang Z, Chang Y, Xia K, Liu J and Yuan W: MicroRNAs, long noncoding RNAs, and circular RNAs: potential tumor biomarkers and targets for colorectal cancer. Cancer Manag Res *10*: 2249-2257, 2018. PMID: 30100756. DOI: 10.2147/CMAR.S166308

5   Brenner H and Chen C: The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention. Br J Cancer *119(7)*: 785-792, 2018. PMID: 30287914. DOI: 10.1038/s41416-018-0264-x

6   Niederreiter M, Niederreiter L, Schmiderer A, Tilg H and Djanani A: Colorectal cancer screening and prevention-pros and cons. memo - Magazine of European Medical Oncology *12(3)*: 239-243, 2020. DOI: 10.1007/s12254-019-00520-z

7   Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA and Jemal A: Colorectal cancer statistics, 2020. CA Cancer J Clin *70(3)*: 145-164, 2020. PMID: 32133645. DOI: 10.3322/caac.21601

8   Dollinger MM, Behl S and Fleig WE: Early detection of colorectal cancer: a multi-center pre-clinical case cohort study for validation of a combined DNA stool test. Clin Lab *64(10)*: 1719-1730, 2018. PMID: 30336540. DOI: 10.7754/Clin.Lab.2018.180521

9   Guraya S: Prognostic significance of circulating microRNA-21 expression in esophageal, pancreatic and colorectal cancers; a systematic review and meta-analysis. Int J Surg *60*: 41-47, 2018. PMID: 30336280. DOI: 10.1016/j.ijsu.2018.10.030

10 Ding L, Lan Z, Xiong X, Ao H, Feng Y, Gu H, Yu M and Cui Q: The dual role of microRNAs in colorectal cancer progression. Int J Mol Sci *19(9)*: 2791, 2018. PMID: 30227605. DOI: 10.3390/ijms19092791

11 To KK, Tong CW, Wu M and Cho WC: MicroRNAs in the prognosis and therapy of colorectal cancer: From bench to bedside. World J Gastroenterol *24(27)*: 2949-2973, 2018. PMID: 30038463. DOI: 10.3748/wjg.v24.i27.2949

12 Hibner G, Kimsa-Furdzik M and Francuz T: Relevance of microRNAs as potential diagnostic and prognostic markers in colorectal cancer. Int J Mol Sci *19(10)*: 2944, 2018. PMID: 30262723. DOI: 10.3390/ijms19102944

13 Filipów S and Łaczmański Ł: Blood circulating miRNAs as cancer biomarkers for diagnosis and surgical treatment response. Front Genet *10*: 169, 2019. PMID: 30915102. DOI: 10.3389/fgene.2019.00169

14 Friedman RC, Farh KK, Burge CB and Bartel DP: Most mammalian mRNAs are conserved targets of microRNAs. Genome Res *19(1)*: 92-105, 2009. PMID: 18955434. DOI: 10.1101/gr.082701.108

15 Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS and Sung JJ: Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut *58(10)*: 1375-1381, 2009. PMID: 19201770. DOI: 10.1136/gut.2008.167817

16 Zhu L, Gao J, Huang K, Luo Y, Zhang B and Xu W: miR-34a screened by miRNA profiling negatively regulates Wnt/β-catenin signaling pathway in Aflatoxin B1 induced hepatotoxicity. Sci Rep *5*: 16732, 2015. PMID: 26567713. DOI: 10.1038/srep16732

17 Li L, Wang A, Cai M, Tong M, Chen F and Huang L: Identification of stool miR-135b-5p as a non-invasive diaognostic biomarker in later tumor stage of colorectal cancer. Life Sci *260*: 118417, 2020. PMID: 32931801. DOI: 10.1016/j.lfs.2020.118417

18 Li T, Lai Q, Wang S, Cai J, Xiao Z, Deng D, He L, Jiao H, Ye Y, Liang L, Ding Y and Liao W: MicroRNA-224 sustains Wnt/β-catenin signaling and promotes aggressive phenotype of colorectal cancer. J Exp Clin Cancer Res *35*: 21, 2016. PMID: 26822534. DOI: 10.1186/s13046-016-0287-1

19 Pereira DM and Rodrigues CMP: miRNAs as modulators of EGFR therapy in colorectal cancer. Adv Exp Med Biol *1110*: 133-147, 2018. PMID: 30623370. DOI: 10.1007/978-3-030-02771-1_9

20 Mlcochova J, Faltejskova P, Nemecek R, Svoboda M and Slaby O: MicroRNAs targeting EGFR signalling pathway in colorectal cancer. J Cancer Res Clin Oncol *139(10)*: 1615-1624, 2013. PMID: 23817698. DOI: 10.1007/s00432-013-1470-9

21 Ding L, Jiang Z, Chen Q, Qin R, Fang Y and Li H: A functional variant at miR-520a binding site in PIK3CA alters susceptibility to colorectal cancer in a Chinese Han population. Biomed Res Int *2015*: 373252, 2015. PMID: 25834816. DOI: 10.1155/2015/373252

22 Peng W, Sha H, Sun X, Zou R, Zhu Y, Zhou G and Feng J: Role and mechanism of miR-187 in human cancer. Am J Transl Res *12(9)*: 4873-4884, 2020. PMID: 33042395.

23 Li Q, Zou C, Zou C, Han Z, Xiao H, Wei H, Wang W, Zhang L, Zhang X, Tang Q, Zhang C, Tao J, Wang X and Gao X: MicroRNA-25 functions as a potential tumor suppressor in colon cancer by targeting Smad7. Cancer Lett *335(1)*: 168-174, 2013. PMID: 23435373. DOI: 10.1016/j.canlet.2013.02.029

24 Eraslan G, Avsec Ž, Gagneur J and Theis FJ: Deep learning: new computational modelling techniques for genomics. Nat Rev Genet *20(7)*: 389-403, 2019. PMID: 30971806. DOI: 10.1038/s41576-019-0122-6

25 Libbrecht MW and Noble WS: Machine learning applications in genetics and genomics. Nat Rev Genet *16(6)*: 321-332, 2015. PMID: 25948244. DOI: 10.1038/nrg3920

26 Ohler U, Liao GC, Niemann H and Rubin GM: Computational analysis of core promoters in the Drosophila genome. Genome Biol *3(12)*: RESEARCH0087, 2002. PMID: 12537576. DOI: 10.1186/gb-2002-3-12-research0087

27 Parveen A, Mustafa SH, Yadav P and Kumar A: Applications of machine learning in miRNA discovery and target prediction. Curr Genomics *20(8)*: 537-544, 2019. PMID: 32581642. DOI: 10.2174/1389202921666200106111813

28 Pizzini S, Bisognin A, Mandruzzato S, Biasiolo M, Facciolli A, Perilli L, Rossi E, Esposito G, Rugge M, Pilati P, Mocellin S, Nitti D, Bortoluzzi S and Zanovello P: Impact of microRNAs on regulatory networks and pathways in human colorectal carcinogenesis and development of metastasis. BMC Genomics *14*: 589, 2013. PMID: 23987127. DOI: 10.1186/1471-2164-14-589

29 Ogata-Kawata H, Izumiya M, Kurioka D, Honma Y, Yamada Y, Furuta K, Gunji T, Ohta H, Okamoto H, Sonoda H, Watanabe M, Nakagama H, Yokota J, Kohno T and Tsuchiya N: Circulating exosomal microRNAs as biomarkers of colon cancer. PLoS One *9(4)*: e92921, 2014. PMID: 24705249. DOI: 10.1371/journal.pone.0092921

30 Vafaee F, Diakos C, Kirschner MB, Reid G, Michael MZ, Horvath LG, Alinejad-Rokny H, Cheng ZJ, Kuncic Z and Clarke S: A data-driven, knowledge-based approach to biomarker discovery: application to circulating microRNA markers of colorectal cancer prognosis. NPJ Syst Biol Appl 4: 20, 2018. PMID: 29872543. DOI: 10.1038/s41540-018-0056-1

31 Schetter AJ, Okayama H and Harris CC: The role of microRNAs in colorectal cancer. Cancer J *18(3)*: 244-252, 2012. PMID: 22647361. DOI: 10.1097/PPO.0b013e318258b78f

32 Dong Y, Wu WK, Wu CW, Sung JJ, Yu J and Ng SS: MicroRNA dysregulation in colorectal cancer: a clinical perspective. Br J Cancer *104(6)*: 893-898, 2011. PMID: 21364594. DOI: 10.1038/bjc.2011.57

33 Desmond BJ, Dennett ER and Danielson KM: Circulating extracellular vesicle microRNA as diagnostic biomarkers in early colorectal cancer-a review. Cancers (Basel) *12(1)*: 52, 2019. PMID: 31878015. DOI: 10.3390/cancers12010052

34 Wang S, Xiang J, Li Z, Lu S, Hu J, Gao X, Yu L, Wang L, Wang J, Wu Y, Chen Z and Zhu H: A plasma microRNA panel for early detection of colorectal cancer. Int J Cancer *136(1)*: 152-161, 2015. PMID: 23456911. DOI: 10.1002/ijc.28136

35 Strubberg AM and Madison BB: MicroRNAs in the etiology of colorectal cancer: pathways and clinical implications. Dis Model Mech *10(3)*: 197-214, 2017. PMID: 28250048. DOI: 10.1242/dmm.027441

36 Liu J, Chen Z, Xiang J and Gu X: MicroRNA-155 acts as a tumor suppressor in colorectal cancer by targeting CTHRC1 *in vitro*. Oncol Lett *15(4)*: 5561-5568, 2018. PMID: 29556299. DOI: 10.3892/ol.2018.8069

37 Li Y, Duo Y, Bi J, Zeng X, Mei L, Bao S, He L, Shan A, Zhang Y and Yu X: Targeted delivery of anti-miR-155 by functionalized mesoporous silica nanoparticles for colorectal cancer therapy. Int J Nanomedicine *13*: 1241-1256, 2018. PMID: 29535520. DOI: 10.2147/IJN.S158290

38 Sun N, Zhang L, Zhang C and Yuan Y: miR-144-3p inhibits cell proliferation of colorectal cancer cells by targeting BCL6 *via* inhibition of Wnt/β-catenin signaling. Cell Mol Biol Lett *25*: 19, 2020. PMID: 32206063. DOI: 10.1186/s11658-020-00210-3

39 Majd M, Hosseini A, Ghaedi K, Kiani-Esfahani A, Tanhaei S, Shiralian-Esfahani H, Rahnamaee SY, Mowla SJ and Nasr-

Esfahani MH: MiR-9-5p and miR-106a-5p dysregulated in CD4[+] T-cells of multiple sclerosis patients and targeted essential factors of T helper17/regulatory T-cells differentiation. Iran J Basic Med Sci *21(3)*: 277-283, 2018. PMID: 29511494. DOI: 10.22038/ijbms.2018.25382.6275

40 Huang L, Ma Q, Li Y, Li B and Zhang L: Inhibition of microRNA-210 suppresses pro-inflammatory response and reduces acute brain injury of ischemic stroke in mice. Exp Neurol *300*: 41-50, 2018. PMID: 29111308. DOI: 10.1016/j.expneurol.2017.10.024

41 Guarino E, Delli Poggi C, Grieco GE, Cenci V, Ceccarelli E, Crisci I, Sebastiani G and Dotta F: Circulating microRNAs as biomarkers of gestational diabetes mellitus: updates and perspectives. Int J Endocrinol *2018*: 6380463, 2018. PMID: 29849620. DOI: 10.1155/2018/6380463

42 Nazarov PV, Reinsbach SE, Muller A, Nicot N, Philippidou D, Vallar L and Kreis S: Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. Nucleic Acids Res *41(5)*: 2817-2831, 2013. PMID: 23335783. DOI: 10.1093/nar/gks1471

43 Champeris Tsaniras S, Delinasios GJ, Petropoulos M, Panagopoulos A, Anagnostopoulos AK, Villiou M, Vlachakis D, Bravou V, Stathopoulos GT and Taraviras S: DNA replication inhibitor geminin and retinoic acid signaling participate in complex interactions associated with pluripotency. Cancer Genomics Proteomics *16(6)*: 593-601, 2019. PMID: 31659113. DOI: 10.21873/cgp.20162

44 Esteves JV, Yonamine CY, Pinto-Junior DC, Gerlinger-Romero F, Enguita FJ and Machado UF: Diabetes modulates microRNAs 29b-3p, 29c-3p, 199a-5p and 532-3p expression in muscle: Possible role in GLUT4 and HK2 repression. Front Endocrinol (Lausanne) *9*: 536, 2018. PMID: 30258406. DOI: 10.3389/fendo.2018.00536

45 Hu Y, Xu R, He Y, Zhao Z, Mao X, Lin L and Hu J: Downregulation of microRNA 106a 5p alleviates ox LDL mediated endothelial cell injury by targeting STAT3. Mol Med Rep *22(2)*: 783-791, 2020. PMID: 32626987. DOI: 10.3892/mmr.2020.11147

46 Kanaan Z, Roberts H, Eichenberger MR, Billeter A, Ocheretner G, Pan J, Rai SN, Jorden J, Williford A and Galandiuk S: A plasma microRNA panel for detection of colorectal adenomas: a step toward more precise screening for colorectal cancer. Ann Surg *258(3)*: 400-408, 2013. PMID: 24022433. DOI: 10.1097/SLA.0b013e3182a15bcc

47 Tan Y, Lin JJ, Yang X, Gou DM, Fu L, Li FR and Yu XF: A panel of three plasma microRNAs for colorectal cancer diagnosis. Cancer Epidemiol *60*: 67-76, 2019. PMID: 30925282. DOI: 10.1016/j.canep.2019.01.015

48 Zhang H, Zhu M, Shan X, Zhou X, Wang T, Zhang J, Tao J, Cheng W, Chen G, Li J, Liu P, Wang Q and Zhu W: A panel of seven-miRNA signature in plasma as potential biomarker for colorectal cancer diagnosis. Gene *687*: 246-254, 2019. PMID: 30458288. DOI: 10.1016/j.gene.2018.11.055

49 Jiang H, Ge F, Hu B, Wu L, Yang H and Wang H: rs35301225 polymorphism in miR-34a promotes development of human colon cancer by deregulation of 3'UTR in E2F1 in Chinese population. Cancer Cell Int *17*: 39, 2017. PMID: 28293146. DOI: 10.1186/s12935-017-0402-1

50 Saridaki Z, Weidhaas JB, Lenz HJ, Laurent-Puig P, Jacobs B, De Schutter J, De Roock W, Salzman DW, Zhang W, Yang D, Pilati C, Bouché O, Piessevaux H and Tejpar S: A let-7 microRNA-binding site polymorphism in KRAS predicts improved outcome in patients with metastatic colorectal cancer treated with salvage cetuximab/panitumumab monotherapy. Clin Cancer Res *20(17)*: 4499-4510, 2014. PMID: 25183481. DOI: 10.1158/1078-0432.CCR-14-0348

51 Zhu M, Huang Z, Zhu D, Zhou X, Shan X, Qi LW, Wu L, Cheng W, Zhu J, Zhang L, Zhang H, Chen Y, Zhu W, Wang T and Liu P: A panel of microRNA signature in serum for colorectal cancer diagnosis. Oncotarget *8(10)*: 17081-17091, 2017. PMID: 28177881. DOI: 10.18632/oncotarget.15059

52 Shellman MH and Shellman YG: Human against machine? Machine learning identifies microRNA ratios as biomarkers for melanoma. J Invest Dermatol *140(1)*: 18-20, 2020. PMID: 31864430. DOI: 10.1016/j.jid.2019.07.688

53 Zhu J, Zheng Z, Wang J, Sun J, Wang P, Cheng X, Fu L, Zhang L, Wang Z and Li Z: Different miRNA expression profiles between human breast cancer tumors and serum. Front Genet *5*: 149, 2014. PMID: 24904649. DOI: 10.3389/fgene.2014.00149

54 Gmerek L, Martyniak K, Horbacka K, Krokowicz P, Scierski W, Golusinski P, Golusinski W, Schneider A and Masternak MM: MicroRNA regulation in colorectal cancer tissue and serum. PLoS One *14(8)*: e0222013, 2019. PMID: 31469874. DOI: 10.1371/journal.pone.0222013