Development and Clinical Validation of a Blood Test Based on 29-Gene Expression for Early Detection of Colorectal Cancer

Laura Ciarloni[1], Sahar Hosseinian Ehrensberger[1], Natsuko Imaizumi[1], Sylvain Monnier-Benoit[1], Cristina Nichita[2], Seung-Jae Myung[3], Joo Sung Kim[4], Si Young Song[5], Tae Il Kim[5], Boudewijn van der Weg[6], Rémy Meier[7], Jan Borovicka[8], Christoph Beglinger[9], Cédric Vallet[10], Philippe Maerten[11], Curzio Rüegg[1,12], and Gian Dorta[2]

Abstract

Purpose: A blood test for early detection of colorectal cancer is a valuable tool for testing asymptomatic individuals and reducing colorectal cancer–related mortality. The objective of this study was to develop and validate a novel blood test able to differentiate patients with colorectal cancer and adenomatous polyps (AP) from individuals with a negative colonoscopy.

Experimental Design: A case–control, multicenter clinical study was designed to collect blood samples from patients referred for colonoscopy or surgery. Predictive algorithms were developed on 75 controls, 61 large AP (LAP) ≥ 1 cm, and 45 colorectal cancer cases and independently validated on 74 controls, 42 LAP, and 52 colorectal cancer cases (23 stages I–II) as well as on 245 cases including other colorectal findings and diseases other than colorectal cancer. The test is based on a 29-gene panel

expressed in peripheral blood mononuclear cells alone or in combination with established plasma tumor markers.

Results: The 29-gene algorithm detected colorectal cancer and LAP with a sensitivity of 79.5% and 55.4%, respectively, with 90.0% specificity. Combination with the protein tumor markers carcinoembryonic antigen (CEA) and CYFRA21-2 resulted in a specificity increase (92.2%) with a sensitivity for colorectal cancer and LAP detection of 78.1% and 52.3%, respectively.

Conclusions: We report the validation of a novel blood test, Colox®, for the detection of colorectal cancer and LAP based on a 29-gene panel and the CEA and CYFRA21-1 plasma biomarkers. The performance and convenience of this routine blood test provide physicians a useful tool to test average-risk individuals unwilling to undergo upfront colonoscopy. *Clin Cancer Res; 1–8. ©2016 AACR.*

Introduction

Colorectal cancer is the second most common malignancy in women, the third one in men, and is the second leading cause of

[1]Novigenix SA, Epalinges, Switzerland. [2]Department of Gastroenterology and Hepatology, Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland. [3]Asan Medical Centre, Department of Gastroenterology, University of Ulsan College of Medicine, Seoul, South Korea. [4]Department of Internal Medicine, College of Medicine, Seoul National University, Seoul, South Korea. [5]Severance Hospital, Division of Gastroenterology, Department of Internal Medicine, Yonsei University College of Medicine, Seoul, South Korea. [6]Spital Thun, Gastroenterology, Thun, Switzerland. [7]Kantonsspital Liestal, Gastroenterology, Hepatology and Nutrition Department, University Hospital, Liestal, Switzerland. [8]Kantonsspital St. Gallen, Department of Gastroenterology and Hepatology, St. Gallen, Switzerland. [9]Universitätsspital Basel, Department of Biomedicine, Basel, Switzerland. [10]Ensemble Hospitalier de la Côte, Surgery Services, Morges, Switzerland. [11]Clinique Cecil, Gastroenterology, Lausanne, Switzerland. [12]Department of Medicine, Faculty of Science, University of Fribourg, Fribourg, Switzerland.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

L. Ciarloni, S.H. Ehrensberger, and N. Imaizumi contributed equally to this article.

Corresponding Authors: Curzio Rüegg, University of Fribourg, 1, Rue Albert Gockel, CH-1700 Fribourg, Switzerland. Phone: 41-(0)-26-300-8766; Fax: 41-(0)-26-300-9733; E-mail: curzio.ruegg@unifr.ch; or Gian Dorta, Department of Gastroenterology and Hepatology, Centre Hospitalier Universitaire Vaudois, Rue du Bugnon 56, CH-1011 Lausanne, Switzerland. E-mail: gian.dorta@chuv.ch

cancer mortality in Europe and North America (1). Each year, more than 1.4 million patients are diagnosed with colorectal cancer globally, and approximately half of these patients die from the disease. Colorectal cancer represents therefore a significant public health issue, especially in industrialized countries. When diagnosed early, colorectal cancer has a good therapeutic response with high survival rates, whereas advanced and metastatic stages of the disease are associated with poor prognosis (2). The development from premalignant lesions, such as adenomatous polyps (AP), and the long presymptomatic course of the disease highlight the importance of screening for colorectal cancer in asymptomatic average-risk individuals.

Recent studies have confirmed that regular screening by guaiac fecal occult blood tests (gFOBT) or flexible sigmoidoscopy significantly reduces colorectal cancer mortality (3–5). Moreover, systematic colorectal cancer screening is cost-effective as it lowers global healthcare costs in comparison with no screening (6–8). Average-risk individuals are recommended to be tested at regular intervals above the age of 50 by either gFBOT, fecal immunochemical test, colonoscopy, or sigmoidoscopy (3). Despite these recommendations, physicians today are facing a challenge in convincing average-risk asymptomatic individuals to be screened for colorectal cancer due to the invasive nature of colonoscopy and the reluctance of individuals to manipulate stool samples. There is therefore an unmet need for an effective and routine test for asymptomatic average-risk individuals. A blood-based test is highly attractive due to its minimal invasiveness, convenience, and high acceptance by individuals at average risk (9–11).

Translational Relevance

Early detection of colorectal cancer dramatically improves outcome of the disease. However, physicians face a challenge in convincing average-risk individuals to be screened for colorectal cancer due to the unpleasant nature of existing methods. Here, we report a novel blood test that effectively differentiates patients with colorectal cancer and large adenomatous polyps from individuals with negative colonoscopy. The test is based on a 29-gene panel expressed in peripheral blood mononuclear cells, alone or in combination with the established plasma tumor markers CEA and CYFRA21-1. The test's performance is consistent with its use as a colorectal cancer detection tool for average-risk individuals reluctant to undergo upfront colonoscopy. This study is being followed by a prospective study to monitor and further evaluate performance of the test in daily clinical practice. Finally, this test might become a useful option to increase compliance of average-risk individuals to colorectal cancer screening.

Myelomonocytic cells from the bone marrow are recruited at the tumor site to create an inflammatory-like microenvironment promoting tumor cell proliferation, survival, motility, angiogenesis, and immune escape (12–15). Based on preclinical and clinical observations, we hypothesized that mobilized cells circulating in the blood may be used as a biomarker for the presence of a tumor (16–18). We have recently reported a novel blood-based approach to detect for AP and colorectal cancer based on this hypothesis (19). In a follow-up study, we reported a 29-gene panel in peripheral blood mononuclear cells (PBMC) for the detection of patients with colorectal cancer or AP (20).

The objective of the present study was to use this 29-gene panel to develop and validate a blood test that discriminates patients with colorectal cancer and AP from individuals without lesions. The clinical sensitivity and specificity of the test were validated in a multicenter, case–control study. Finally, we investigated whether combining the newly developed 29-gene test with protein tumor markers already in routine clinical use for cancer management would benefit the test's accuracy.

Materials and Methods

Study design

The DGNP-COL-0310 was a multicenter, case–control clinical study in which, from June 2010 to February 2012, 1,579 participants were enrolled in seven Swiss and three South Korean hospitals (Supplementary Materials and Methods). In order to have a second independent sample set to prospectively validate the algorithm, an extension to the study was conducted from November 2012 to April 2013 in four Swiss centers, which enrolled 85 additional subjects (Fig. 1). The study protocol (DGNP-COL-0310) was approved by the competent ethics committees and Institutional Review Boards for research on human subjects (Supplementary Materials and Methods). All participants provided a written informed consent.

Study population

The study recruited subjects older than 50 years who were referred by general practitioners for the great majority (∼80%) for a diagnostic colonoscopy and for a minority (∼20%) for a screening colonoscopy. Most of the colorectal cancer cases (64%) were recruited among subjects who were scheduled for surgery.

Subjects with a personal history of polyps, colorectal cancer, other types of cancer, or with family history indicating high risk for colorectal cancer (≥ 2 first-degree relatives with colorectal cancer or ≥1 with colorectal cancer at the age< 50 years, familial polyposis, or Lynch syndrome) were not eligible for the study.

Following diagnosis, the enrolled participants were retrospectively allocated to clinically predefined groups as summarized in Fig. 2. Main study groups included controls, large AP (LAP), and colorectal cancer. Subjects in the control group were clear from any colorectal lesions, including hyperplastic polyps and small AP. The LAP group included patients with at least one adenoma ≥ 1 cm and no colorectal cancer lesion. Categorization of adenomas into the standard classification (i.e., advanced and nonadvanced adenoma) could not be performed due to the fact that the histologic grading of small adenomas was not collected in the case report form. Nevertheless, because adenomas larger than 1 cm constitute the vast majority of the advanced adenomas, the result relative to this group could be generalized with a good approximation for the advanced adenomas. The colorectal cancer group was diagnosed with adenocarcinoma or, rarely, mucinous adenocarcinoma, ranging from stages I to IV (AJCC system, 7th edition). Colorectal cancer cases with unknown stages were collected but not included in the per-protocol analysis. The remaining study groups included patients with small AP (<1 cm), hyperplastic polyps, inflammatory bowel diseases (IBD), other inflammatory diseases (e.g., psoriasis, rheumatoid arthritis, hepatitis), other gastrointestinal (GI) disorders (e.g., gall stones, diverticulitis), infectious viral diseases (e.g., HIV, Herpes Simplex), and malignancies other than colorectal cancer.

Enrolled subjects meeting all inclusion and exclusion criteria, including group-specific criteria, were designated as per-protocol subjects and used for per-protocol analysis. The study also enrolled subjects not meeting all the protocol criteria, but deemed to be valuable for the algorithm evaluation. For instance, some of them presented with comorbidities involving more than one medical condition/disease that defined the study group (e.g., LAP and inflammatory disease, colorectal cancer, and other GI disorder). These subjects were classified as "non per-protocol" and intended to be analyzed separately to test the algorithm response on a larger spectrum of medical conditions and comorbidities.

Clinical procedures

A colonoscopy was performed for all study patients, except for those with malignancies other than colorectal cancer, due to ethical reasons. Colonoscopy to the cecum was required in all cases except for colorectal cancer patients with a stenosis caused by a tumor mass. Resected LAP or colorectal cancer tissues were sent to a central pathology board composed of three independent pathologists blinded to all subject information, group allocation,
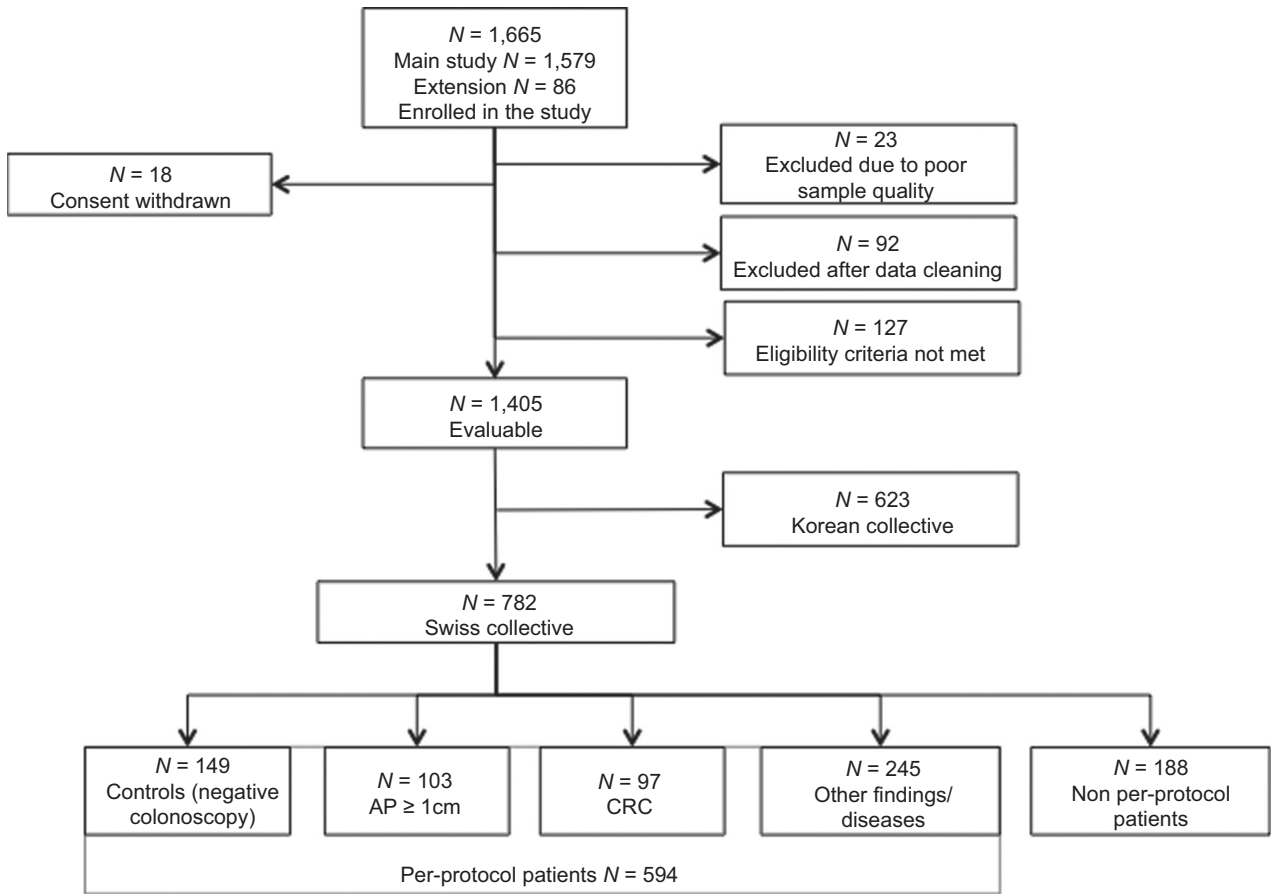
Figure 1.
Flowchart showing patient disposition.

and assessments made by other pathologists (Supplementary Materials and Methods).

For each participant, 16 mL of peripheral blood was collected into 4 × 4 mL Vacutainer CPT tubes (Becton Dickinson), before or immediately after colonoscopy. Most of the samples from subjects diagnosed with colorectal cancer requiring a surgery were collected up to 12 weeks after colonoscopy. In all cases, blood was collected prior to polyp resection and prior to any measures in preparation for surgery or any cancer-specific treatment.

Laboratory procedures

Sample preparation and analysis were performed in a blinded manner by laboratory personnel. PBMC separation was performed locally within 6 hours according to Vacutainer CPT tube manufacturer's instructions. PBMC pellets were resuspended and stored at −80°C. Plasma from each sample was collected and stored at −80°C.

Automated purification of total RNA from PBMC pellets was performed at a central laboratory. RNA integrity was analyzed and, on average, RNA samples showed an RNA integrity
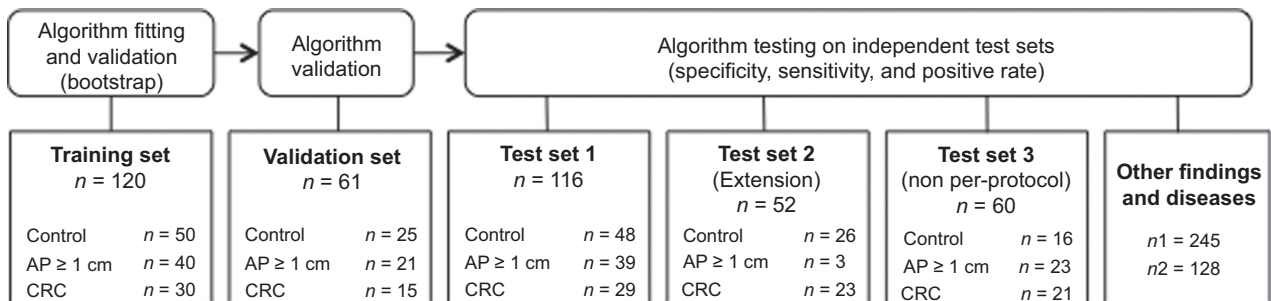

Figure 2.
Study design diagram. The training, validation, Test Sets 1 and 2 and part of subjects with other findings and diseases (n1) constituted the per-protocol cases. Test Set 3 and remaining patients with other findings and diseases (n2) were enrolled as non per-protocol cases.

number (RIN) of 8.6 ± 0.7. RNA samples with RIN < 5 were considered of poor quality and discarded. RNA samples were aliquoted and stored at −80°C.

For each sample, 200 ng of total RNA was reverse-transcribed into cDNA using the SuperScript VILO cDNA Synthesis Kit according to the manufacturer's instructions. Real-time PCR analysis was performed on the Lightcycler 480 instruments using a 29-gene panel previously reported (20). Briefly, PCR reactions were carried out on a 384-well plate, preloaded with RealTime ready Custom RT-qPCR assays using RealTime Ready DNA Probes Master Mix. Gene expression values (Cp) were normalized by the ΔCT method using the mean of the three housekeeping genes RPLP0, NACA, and TPT1.

Plasma concentration of CEA, CYFRA21-1, CA125, and CA19-9 tumor biomarkers was measured for each patient on the Architect immunoassay analyzer platform according to the manufacturer's instructions.

Details on the equipment used during the laboratory procedures and relative manufacturers can be found in Supplementary Materials and Methods.

Statistical design and modeling

The study aimed at enrolling 1,600 participants in order to have 1,400 evaluable subjects' samples. Of those, 900 were expected to be equally distributed between the three main study groups (control, LAP, and colorectal cancer) for the development and clinical validation of the test. The remaining subjects, presenting with other predefined diseases, were allocated to secondary groups for testing the algorithm specificity. In order to reach a significance level $\alpha = 0.05$ and a power $1 - \beta = 0.80$, the study required at least 200 controls, 190 colorectal cancer cases, and 240 LAP cases. Out of those, 120 controls, 70 colorectal cancer cases, and 95 LAP cases had to be collected at the Swiss sites. Samples from the three main groups were stratified by gender, age, site, colorectal cancer stage, and LAP size and randomly distributed into training, validation, and Test Set 1, in the proportion of 40%, 20%, and 40%, respectively (Fig. 2). Training and validation sets were used in the discovery phase to define the predictive algorithms, including 45 samples previously used for the identification of the 29-gene panel (20). In the testing phase, the algorithms were validated on Test Set 1, Test Set 2, collected in the extension study, and Test Set 3, which included "non per-protocol" control, LAP, and colorectal cancer subjects. Samples from subjects with other colorectal lesions or diseases were used to test the positive rate of the predictive algorithms.

Classifiers and multiple classifier systems were generated by penalized logistic regression (21–23) and fuzzy logic (24, 25) modeling techniques. First, an algorithm based on the 29-gene signature was defined following a multiple classifier combination strategy. We refer to it as a multigene multiclassifier (MGMC) algorithm. The MGMC algorithm was then combined with CEA and CYFRA21-1 proteins, using a decision tree classification approach (MGMC-P algorithm). Both algorithms release a binary result, which suggests the presence or absence of advanced colorectal neoplasia. The methodology used to develop the predictive algorithms of Colox is detailed in the Supplementary Materials and Methods.

Specificity was estimated as $1 -$ positivity among subjects belonging to the control group. For test characteristics, 95%

confidence intervals (CI) were computed with the use of an exact binomial test.

The *R* statistics environment was used for all statistical analyses.

Results

Study population

Out of 1,665 enrolled subjects, 1,405 were considered fully evaluable (84.4%). Of those, 782 were recruited in Switzerland and 623 in South Korea. This report focuses on the Swiss dataset only. Korean data will be presented separately.

In the Swiss collective, the per-protocol population was constituted of 594 participants (Fig. 1), including 149 control subjects, 103 LAP, and 97 colorectal cancer patients from stages I to IV, as well as 245 subjects diagnosed with other types of colorectal lesions (e.g., small adenomas, hyperplastic polyps), solid cancers, or other diseases such as IBD or diverticulitis, for testing the algorithm specificity. In addition, 188 cases classified as "non per-protocol" were enrolled for testing the algorithm behavior on a larger spectrum of medical conditions and comorbidities. Demographic and clinical characteristics of the subjects are reported in Table 1. Subjects in the control, LAP, and colorectal cancer groups had a mean age of 60.6, 67.3, and 69.4, respectively. The control group enrolled 54% women and 46% men, whereas the LAP and colorectal cancer group enrolled 64% and 62% of men, respectively.

29-gene expression analysis

All blood samples were profiled with a 29-gene panel (Supplementary Table S1) previously reported (20) and discovered through a combined univariate and multivariate approach. No differences in normalized gene expression were observed across collecting centers (data not shown) or across different age categories within the control or the LAP group (Supplementary Table S2). A few genes were found to be differentially expressed between individuals of different age classes in the colorectal cancer group, as well as between men and women within all

Table 1. Clinical and demographic characteristics of the study participants

Category	Patient (*N*)	Age (mean ± SD)	Male (%)
Control	149	60.6 ± 7.8	46
LAP	103	67.3 ± 8.2	64.1
1–2 cm	57	67.5 ± 8.4	66.7
>2 cm	46	67.1 ± 8.1	60.9
Colorectal cancer[a]	97	69.4 ± 9.7	61.9
Stage I	22	71.2 ± 9.2	59.1
Stage II	22	69.8 ± 8.6	63.6
Stage III	31	68.1 ± 12.3	54.8
Stage IV	22	69 ± 7.0	72.7
Other colorectal lesions and diseases	245	—	—
AP < 1 cm	67	65.2 ± 8.9	77.6
Hyperplastic polyps	61	60.4 ± 7.6	59
IBD[b]	14	58.7 ± 6.1	71.4
Other GI diseases[c]	19	62.7 ± 10.3	42.1
Inflammatory diseases[d]	11	65 ± 8.3	18.2
Hepatitis, Herpes, AIDS	10	61.8 ± 5.9	40
Other cancers	63	67.2 ± 8.0	71.4
Non per-protocol cases	188	65.1 ± 10.0	56.4

[a]All colorectal cancers were adenocarcinomas except one squamous carcinoma. Two were mucinous type.
[b]Eight Crohn's disease and 6 ulcerative colitis.
[c]Including 7 gallstones and 9 diverticulitis.
[d]Three autoimmune hepatitis, 5 psoriasis, and 3 rheumatoid arthritis.

Table 2. Specificity and sensitivity of the MGMC algorithm for colorectal cancer and LAP detection

	Test set 1		Test set 2		Test sets 1 + 2	Test sets 1 + 2 + 3	
	Total	Sensitivity (95% CI)	Total	Sensitivity (95% CI)	Sensitivity (95% CI)	Total	Sensitivity (95% CI)
	N	%	*N*	%	%	*N*	%
LAP	39	51.3 (35–68)	3	100 (29–100)	54.8 (39–70)	65	55.4 (43–68)
1–2 cm	22	54.5 (32–76)	1	100	56.5	—	—
>2 cm	17	47.1 (23–72)	2	100	52.6	—	—
Low grade	30	50.0 (31–69)	0	—	50.0 (31–69)	—	—
High grade	6	50.0 (12–88)	3	100 (29–100)	66.6 (30–93)	—	—
Tubular	9	44.4 (14–78)	0	—	44.4 (14–78)	—	—
Tubulovillous	21	47.6 (26–70)	3	100 (29–100)	54.2 (33–74)	—	—
Villous	5	60.0 (15–95)	0	—	60.0 (15–95)	—	—
Serrated	4	75.0 (19–99)	0	—	75.0 (19–99)	—	—
Colorectal cancer	29	75.9 (57–90)	23	73.9 (52–90)	75.0 (61–86)	73	79.5 (68–88)
Stages I and II	14	71.4 (42–92)	9	44.4 (14–78)	60.9 (39–80)	33	66.7 (48–82)
Stages III and IV	15	80.0 (52–96)	14	92.9 (66–100)	86.2 (68–96)	40	90.0 (76–97)
Stages I–III	22	81.8 (60–95)	19	68.4 (44–87)	75.6 (60–88)	58	79.3 (67–89)
Low grade	23	69.6 (47–87)	18	66.7 (45–89)	68.3 (63–92)	—	—
High grade	5	100 (48–100)	3	100 (29–100)	100 (63–100)	—	—
		Specificity (95% CI)		Specificity (95% CI)	Specificity (95% CI)		Specificity (95% CI)
	N	%	*N*	%	%	*N*	%
Controls (negative colonoscopy)	48	89.6 (77–97)	26	88.5 (70–98)	89.2 (80–95)	90	90.0 (82–95)

three groups. These differences in individual genes had no impact on the output of the 29-gene algorithms as shown by the results described further below. The most significant genes for colorectal cancer discrimination ($P < 0.0008$) were *S100A8*, *IL1B*, *CCR1*, and *PTGS2*, with 1.7- to 2.1-fold increase in gene expression (linear scale, Supplementary Table S1). Some of these genes (*S100A8*/*Calprotectin* and *PTGS2*/*COX2*) are well-known players in colorectal cancer (13, 26).

Clinical performance of the MGMC algorithm

The MGMC algorithm was applied to Test Set 1 showing a specificity of 89.6% and a sensitivity of 75.9% for colorectal cancer detection (Table 2). These performance characteristics were confirmed when the algorithm was tested on a second independent sample set (Test Set 2), showing the stability of the algorithm (Table 2). When the two test sets were combined, the algorithm showed a sensitivity of 75.0% for colorectal cancer and of 54.8% for LAP detection with specificity of 89.2%. The test performed well on early stage colorectal cancer with a sensitivity of 60.9% for colorectal cancer stages I–II, which increased to 75.6% when colorectal cancer stage III was included. No differences were observed in LAP detection in patients with only one versus more than one LAP (data not shown). Small AP (<1 cm) and hyperplastic polyps were also detected by the test but with lower sensitivity (29.9% and 29.5%, respectively; Table 3).

Table 3. Positive rate of the MGMC and MGMC-P algorithms for other colorectal lesions and diseases

		MGMC	MGMC-P
	Total	Positive rate	Positive rate
	N	%	%
AP < 1 cm	67	29.9	28.4
Hyperplastic polyps	61	29.5	27.9
IBD	14	42.9	28.6
Other GI diseases	19	36.8	31.6
Inflammatory diseases	11	18.2	18.2
Hepatitis, Herpes, AIDS	10	20.0	10.0

The MGMC algorithm was also tested on subjects with other diseases (Table 3 and Supplementary Table S3). Non-GI inflammatory and viral diseases showed a low rate of positivity, comparable with the controls, whereas IBD and other GI diseases had a higher positivity rate. The other cancers showed a very heterogeneous behavior ranging from a positive rate of 41.7% for pancreatic cancer to 80.0% for other GI cancers.

The algorithm was further tested on the Test Set 3, and the results are presented in Supplementary Table S4. They confirmed the findings of the per-protocol analyses. When these samples were added to the per-protocol samples, MGMC sensitivity for LAP and colorectal cancer was 55.4% and 79.5%, respectively, with a specificity of 90.0% (Table 2).

Combination of the MGMC algorithm with protein tumor markers

Among the four tumor markers analyzed, CEA and CYFRA21-2 showed the most significant *P* values (less than 0.001) for discrimination between the colorectal cancer and control groups, and the highest area under the curve for colorectal cancer and control classification (data not shown). In the training set, these tumor markers displayed a specificity of 86.0% and 94.0% and sensitivity for colorectal cancer detection of 63.3% and 43.3%, respectively. These results were confirmed in the validation set with a specificity of 88.0% and 96.0% and sensitivity for colorectal cancer of 60.0% and 26.7%, respectively. CEA and CYFRA21-1 sensitivity was directly correlated to colorectal cancer stage (data not shown).

In order to explore whether tumor-specific markers could complement and strengthen the MGMC algorithm, we added the variables CEA and CYFRA21-1. The resulting overall specificity (Test Sets 1+2) was 91.9%, higher than the MGMC algorithm alone (Table 4), although this difference was not statistically significant ($P > 0.05$). Sensitivity for colorectal cancer (75.0%) and LAP (52.4%) was similar to the one reported by MGMC (Table 4). When the Test Set 3 (Supplementary Table S4) was also taken into consideration, the test reached a sensitivity of 52.3% and 78.1%, respectively, with a specificity of 92.2% (Table 4).

Table 4. Specificity and sensitivity of the MGMC-P algorithm for colorectal cancer and LAP detection

	Test sets 1 + 2		Test sets 1 + 2 + 3	
	Total	Sensitivity (95% CI)	Total	Sensitivity (95% CI)
	N	%	*N*	%
LAP	42	52.4 (37–69)	65	52.3 (40–65)
1–2 cm	23	52.2 (31–73)	—	—
>2 cm	19	52.6 (31–79)	—	—
Colorectal cancer	52	75.0 (61–86)	73	78.1 (67–87)
Stages I and II	23	56.5 (35–77)	33	60.6 (42–75)
Stages III and IV	29	89.7 (73–98)	40	92.5 (80–98)
Stages I–III	41	70.7 (55–84)	58	74.1 (61–85)
		Specificity (95% CI)		**Specificity (95% CI)**
	N	%	*N*	%
Controls (negative colonoscopy)	74	91.9 (83–97)	90	92.2 (85–97)

When colorectal cancer cases were stratified according to the time of recruitment and blood draw, i.e., before surgery or at colonoscopy, a slightly higher detection rate was observed in the cases collected before surgery (79% vs 74%). However, this difference was not statistically significant ($p > 0.05$).

Due to the observed differences in age and sex distribution between the control and the colorectal cancer group and differences in the expression levels of some individual biomarkers, statistical testing was carried out to evaluate a possible confounding effect of age or sex on colorectal cancer discrimination by the algorithm based on the 29-gene panel. None of the considered factors showed an association with the test result (P value > 0.05; Supplementary Table S5). Moreover, age and sex were tested as covariates in multivariate logistic regression analysis. None of these two variables played a statistically significant role in differentiating the control from colorectal cancer group.

Discussion

Early detection of colorectal cancer dramatically improves the clinical outcome of the disease, and regular screening has proven to reduce mortality (2). The "gold standard" for AP and colorectal cancer detection is colonoscopy. It has the highest sensitivity for AP or colorectal cancer among all screening tests and has the advantage to allow direct visualization of the entire colon and removal of precancerous lesions. However, due to the unpleasant bowel preparation required by the method as well as its invasive nature, compliance to screening colonoscopy remains unsatisfactorily low. Other recommended options are stool-based tests, such as the guaiac or immunochemical FOBTs or the recently FDA-approved stool DNA test (Cologuard; refs. 27–31). Stool tests require stool manipulation at home, which is unappealing and inconvenient to some, thus contributing to the low screening compliance despite their promising performances.

An accurate blood test would be an attractive alternative for asymptomatic, average-risk individuals reluctant to undergo screening by a stool test or colonoscopy. This is supported by a Dutch study, based on a population survey and mathematical simulations which concluded that the target population preferred to be screened using a blood- rather than a stool-based test (26% vs. 17%; ref. 32). Another study reported that of 172 age-eligible individuals who were indicated for colorectal cancer screening by colonoscopy, 63% refused the procedure, whereas 97% of those who refused colonoscopy agreed to be screened by a less invasive test such as a blood test (83%) or a stool test (17%; ref. 10). It is therefore reasonable to expect that a blood test can help convincing average-risk individuals to be tested for colorectal cancer, and in the long term to contribute to reduction of colorectal cancer–related mortality.

A blood test based on detecting aberrantly methylated DNA in the *Septin9* gene (Epi proColon) has recently been reviewed by the FDA, but approval is still pending. In a large prospective screening study, the test showed sensitivity for colorectal cancer of 68% at specificity of 79%. The test was not able to detect advanced adenomas, showing sensitivity equivalent to the false-positive detection rate in the control group (22% vs. 21%; ref. 33).

Here, we report a validation study for a newly developed blood test, which can identify 78% of colorectal cancer and 52% of large adenomas. The latter result is of particular interest as large adenomas are one the most clinically relevant precursor lesions of colorectal cancer. Two different predictive algorithms were developed: one based on a 29-gene panel expressed in PBMC (MGMC algorithm) and one combining the MGMC algorithm with the plasma tumor markers CEA and CYFRA21-1 (MGMC-P algorithm). Although both algorithms had a comparable sensitivity for colorectal cancer, MGMC-P demonstrated a slightly higher specificity (92.2% vs. 90.0%). The clinical advantage of higher specificity is that the number of false-positive individuals sent to colonoscopy would be lower, thus avoiding unnecessary invasive procedures.

The test was also evaluated in patients presenting other colorectal lesions and a various spectrum of different diseases. The test detected around 30% of AP smaller than 1 cm. It would be of interest to evaluate how this detection rate correlates with the histology, i.e., nonadvanced versus advanced AP, the latter representing approximately 13% of the small AP (34).

The test detected also 28% to 30% of subjects with hyperplastic polyps which have traditionally not been considered as a relevant risk factor for colorectal cancer, unless present as hyperplastic polyposis syndrome (35). Whether the detection of these lesions is an added value for the test may be debatable. Under the scenario that considers the detection of hyperplastic polyps and small adenomas a false positive event, the test's specificity would be estimated at 83%, based on the study data and published prevalence data (36, 37). The full biologic and clinical relevance of isolated hyperplastic polyps to colorectal cancer

development, however, has not been thoroughly investigated (38). Emerging evidence indicates that hyperplastic polyps share similar morphologic and molecular features with sessile serrated adenomas, which are now recognized as precancerous lesions that can rapidly progress into invasive carcinoma (39). KRAS and BRAF mutations and aberrant CpG-island methylation are observed in a significant fraction of hyperplastic polyps (approximately 30%–60%) consistent with the notion that hyperplastic polyps are a subset of serrated adenoma (40–42). On the basis of these considerations, and in the context of a more personalized medicine, the detection of hyperplastic polyps with carcinogenic potential would be a desirable feature for such a test, although it may come with the burden of detecting also benign hyperplastic polyps. The clinical significance of a positive test in individuals with hyperplastic polyps will be explored in future studies.

The positive rate of the test in patients with IBD and other GI benign diseases was close to 30%. Because these patients are normally under close surveillance and/or are presenting symptoms, they are considered ineligible for testing by this kind of test.

In clinical practice, this test will not be applied to individuals diagnosed with other cancers. We nevertheless evaluated a limited number of patients with malignancies other than colorectal cancer and found different degrees of positivity. These exploratory observations did not yield relevant conclusions because the study was not designed to evaluate the test for the early detection of cancers other than colorectal cancer, and most patients were not investigated by colonoscopy.

In summary, a new blood test has been developed and clinically validated for detection of LAP and colorectal cancer based on a 29-gene expression panel and the CEA and CYFRA21-1 plasma biomarkers. The test's performance is consistent with its use as a colorectal cancer detection tool for average-risk individuals reluctant to undergo upfront colonoscopy. This convenient and routine blood test can help identify individuals who will benefit the most from a colonoscopy. A prospective study is currently on-going to monitor and further evaluate the performance of the test in daily clinical practice.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–86.
2. Brenner H, Kloor M, Pox CP. Colorectal cancer. Lancet 2014;383: 1490–502.
3. Garborg K, Holme O, Loberg M, Kalager M, Adami HO, Bretthauer M. Current status of screening for colorectal cancer. Ann Oncol 2013;24: 1963–72.
4. Shaukat A, Mongin SJ, Geisser MS, Lederle FA, Bond JH, Mandel JS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;369:1106–14.
5. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: A multicentre randomised controlled trial. Lancet 2010;375: 1624–33.
6. Cruzado J, Sanchez FI, Abellan JM, Perez-Riquelme F, Carballo F. Economic evaluation of colorectal cancer (CRC) screening. Best Pract Res Clin Gastroenterol 2013;27:867–80.
7. Tilson L, Sharp L, Usher C, Walsh C, S W, O'Ceilleachair A, et al. Cost of care for colorectal cancer in Ireland: A health care payer perspective. Eur J Health Econ 2012;13:511–24.
8. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. Epidemiol Rev 2011;33: 88–100.
9. Ganepola GA, Nizin J, Rutledge JR, Chang DH. Use of blood-based biomarkers for early diagnosis and surveillance of colorectal cancer. World J Gastrointest Oncol 2014;6:83–97.
10. Adler A, Geiger S, Keil A, Bias H, Schatz P, deVos T, et al. Improving compliance to colorectal cancer screening using blood and stool based tests in patients refusing screening colonoscopy in Germany. BMC Gastroenterol 2014;14:183.
11. Taber JM, Aspinwall LG, Heichman KA, Kinney AY. Preferences for blood-based colon cancer screening differ by race/ethnicity. Am J Health Behav 2014;38:351–61.
12. Lorusso G, Ruegg C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. Histochem Cell Biol 2008; 130:1091–103.
13. Peddareddigari VG, Wang D, Dubois RN. The tumor microenvironment in colorectal carcinogenesis. Cancer Microenviron 2010;3:149–66.
14. McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nat Cell Biol 2014;16:717–27.

15. Hanahan D, Coussens LM. Accessories to the crime: Functions of cells recruited to the tumor microenvironment. Cancer Cell 2012; 21:309–22.

16. Laurent J, Hull EF, Touvrey C, Kuonen F, Lan Q, Lorusso G, et al. Proangiogenic factor PlGF programs CD11b(+) myelomonocytes in breast cancer during differentiation of their hematopoietic progenitors. Cancer Res 2011;71:3781–91.

17. Guex N, Crespo I, Bron S, Ifticene-Treboux A, Faes-van't Hull E, Kharoubi S, et al. Angiogenic activity of breast cancer patients' monocytes reverted by combined use of systems modeling and experimental approaches. PLOS Biol 2015;11:e1004050.

18. Kuonen F, Laurent J, Secondini C, Lorusso G, Stehle JC, Rausch T, et al. Inhibition of the Kit ligand/c-Kit axis attenuates metastasis in a mouse model mimicking local breast cancer relapse after radiotherapy. Clin Cancer Res 2012;18:4365–74.

19. Nichita C, Ciarloni L, Monnier-Benoit S, Hosseinian S, Dorta G, Ruegg C. A novel gene expression signature in peripheral blood mononuclear cells for early detection of colorectal cancer. Aliment Pharmacol Ther 2014;39:507–17.

20. Ciarloni L, Hosseinian S, Monnier-Benoit S, Imaizumi N, Dorta G, Ruegg C. Discovery of a 29-gene panel in peripheral blood mononuclear cells for the detection of colorectal cancer and adenomas using high throughput real-time PCR. PLoS One 2015;10:e0123904.

21. Park MY, Hastie T. L1 regularization path algorithm for generalized linear models. J R Stat Soc B 2007;69:659–77.

22. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33:1–22.

23. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. Biom J 2010;52:70–84.

24. Vose MD. Simple genetic algorithm: Foundation and theory. Ann Arbor, MI: MIT Press; 1999.

25. Pena-Reyes CA.Evolutionary fuzzy modeling human diagnostic decisions. Ann N Y Acad Sci 2004;1020:190–211.

26. Duan L, Wu R, Ye L, Wang H, Yang X, Zhang Y, et al. S100A8 and S100A9 are associated with colorectal carcinoma progression and contribute to colorectal carcinoma cell survival and migration via Wnt/beta-catenin pathway. PLoS One 2013;8:e62092.

27. Guittet L, Bouvier V, Mariotte N, Vallee JP, Arsene D, Boutreux S, et al. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. Gut 2007;56:210–4.

28. Imperiale TF, Ransohoff DF, Itzkowitz SH. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014;371:187–8.

29. de Wijkerslooth TR, Stoop EM, Bossuyt PM, Meijer GA, van Ballegooijen M, van Roon AH, et al. Immunochemical fecal occult blood testing is equally sensitive for proximal and distal advanced neoplasia. Am J Gastroenterol 2012;107:1570–8.

30. Park DI, Ryu S, Kim YH, Lee SH, Lee CK, Eun CS, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;105:2017–25.

31. Rozen P, Liphshitz I, Barchana M. Follow-up of patients undergoing both semiquantitated immunochemical fecal occult blood and colonoscopy examinations. Eur J Cancer Prev 2012;21:247–53.

32. Benning TM, Dellaert BG, Dirksen CD, Severens JL. Preferences for potential innovations in non-invasive colorectal cancer screening: A labeled discrete choice experiment for a Dutch screening campaign. Acta Oncol 2014;53:898–908.

33. Potter NT, Hurban P, White MN, Whitlock KD, Lofton-Day CE, Tetzner R, et al. Validation of a real-time PCR-based qualitative assay for the detection of methylated SEPT9 DNA in human plasma. Clin Chem 2014;60:1183–91.

34. Tsai FC, Strum WB. Prevalence of advanced adenomas in small and diminutive colon polyps using direct measurement of size. Dig Dis Sci 2011;56:2384–8.

35. Orlowska J. Hyperplastic polyposis syndrome and the risk of colorectal cancer. Gut 2012;61:470–1.

36. Heitman SJ, Ronksley PE, Hilsden RJ, Manns BJ, Rostom A, Hemmelgarn BR Prevalence of adenomas and colorectal cancer in average risk individuals: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009;7:1272–8.

37. Hazewinkel Y, de Wijkerslooth TR, Stoop EM, Bossuyt PM, Biermann K, van de Vijver MJ, et al. Prevalence of serrated polyps and association with synchronous advanced neoplasia in screening colonoscopy. Endoscopy 2014;46:219–24.

38. Melato M, Rizzardi C, Nardon E, Stanta G. Scarce information about the risk of cancer in colorectal hyperplastic polpys and polyposis. Am J Gastroenterol 2001;96:2267–8.

39. Langner C. Serrated and non-serrated precursor lesions of colorectal cancer. Dig Dis 2015;33:28–37.

40. Kim KM, Lee EJ, Ha S, Kang SY, Jang KT, Park CK, et al. Molecular features of colorectal hyperplastic polyps and sessile serrated adenoma/polyps from Korea. Am J Surg Pathol 2011;35:1274–86.

41. O'Brien MJ. Hyperplastic and serrated polyps of the colorectum. Gastroenterol Clin North Am 2007;36:947–68, viii.

42. Janjua HG, Hogdall E, Linnemann D. Hyperplastic polyps of the colon and rectum - reclassification, BRAF and KRAS status in index polyps and subsequent colorectal carcinoma. APMIS 2015; 123:298–304.