An optimized predictor panel for colorectal cancer diagnosis based on the combination of tumor-associated antigens obtained from protein and phage microarrays

Rodrigo Barderas¹⁺, Ingrid Babel^{1,+,¶}, Ramón Díaz-Uriarte², Víctor Moreno^{3,4}, Adolfo Suárez⁵, Felix Bonilla⁶, Roi Villar-Vázquez¹, Gabriel Capellá³ and J. Ignacio Casal^{1*}

¹ Functional Proteomics Laboratory. Centro de Investigaciones Biológicas (CIB-CSIC). 28040 Madrid. Spain.

² Department of Biochemistry, Universidad Autónoma de Madrid (UAM) - Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM), Madrid, Spain.

³ IDIBELL- Institut Català d'Oncologia. 08907 Barcelona, Spain.

⁴ Facultad de Medicina, Universidad de Barcelona. 08907 Barcelona, Spain

⁵ Hospital de Cabueñes. 33394 Gijón. Spain

⁶ Hospital Puerta de Hierro Majadahonda. 28222 Madrid. Spain

⁺, these authors contributed equally to the manuscript

Keywords: autoantibodies, colorectal cancer, diagnosis, protein microarrays, tumor-associated antigens

Running title: Autoantibodies for colorectal cancer diagnosis

¶. Current Address. Molecular Diagnostics, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, M5-C800, Seattle, WA 98109, USA

***Corresponding author:**

J. Ignacio Casal
Functional Proteomics Laboratory
Centro de Investigaciones Biológicas (CIB-CSIC)
Ramiro de Maeztu, 9
28040 Madrid. Spain
Phone: +34 91 8373112
Fax: +34 91 5360432
e-mail: icasal@cib.csic.es

SUMMARY

Humoral response in cancer patients appears early in cancer progression and can be used for diagnosis, including early detection. By using human recombinant protein and T7 phages microarrays displaying colorectal cancer (CRC)-specific peptides, we previously selected 6 phages and 6 human recombinant proteins as tumor-associated antigens (TAAs) with high diagnostic value. After completing validation in biological samples, TAAs were classified according to their correlation, redundancy in reactivity patterns and multiplex diagnostic capabilities. For predictor model optimization, TAAs were reanalyzed with a new set of samples. A combination of three phages displaying peptides homologous to GRN, NHSL1 and SREBF2 and four proteins PIM1, MAPKAPK3, FGFR4 and ACVR2B, achieved an area under the curve (AUC) of 94%, with a sensitivity of 89.1% and specificity of 90.0%, to correctly predict the presence of cancer. For early colorectal cancer stages, the AUC was 90%, with a sensitivity of 88.2% and specificity of 82.6%. In summary, we have defined an optimized predictor panel, combining TAAs from different sources, with highly improved accuracy and diagnostic value for colorectal cancer.

1 Introduction

Colorectal cancer (CRC) is the major cause of cancer-associated mortality in developed countries [1]. An early detection of the disease is critical for higher patient survival rates. The most common CRC protein biomarker, CEA, is mostly adequate for advanced stages and for monitoring recurrence of the disease [2, 3]. Final outcome of patients improves substantially with early detection of CRC, with 5-year survival rates of 91.1%, %, 69.8% and 11.4% for localized, regional and distant stages, respectively (Surveillance Epidemiology and End Results (2011). National Cancer Institute. USA).

In cancer, self-proteins altered before or during tumor formation and progression, can elicit an immune response when they are released from the tumoral cells to the blood of the patients [4-7]. These proteins can suffer point mutations, misfolding, truncation, over-expression, aberrant splicing or degradation as a consequence of mutations in the DNA of tumoral cells [8-14]. Patients with cancer spontaneously develop a humoral response against these altered tumor-associated autoantigens (TAAs), producing autoantibodies [15-17]. The immune response occurs early during tumor development. Measuring levels of autoantibodies in blood would allow for a non-invasive cancer diagnosis at earlier stages, providing an effective method for cancer screening and preclinical diagnosis. This approach has already been successfully used in other cancers [4, 15, 18-24].

The development of noninvasive screening tests would represent a major advance in the fight against cancer, as pre-clinical or early diagnosis are considered the best prevention to reduce cancer mortality. Screening tests based on autoantibodies present in serum or plasma of cancer patients are a promising alternative to fulfill this goal. Because antibodies are very stable molecules and their differences between control and case samples are easier to detect, we have used the humoral response in cancer patients' to identify autoantibodies and their respective TAAs as biomarkers of CRC.

High-density protein microarrays overcome the limitation inherent to other methods, such as SERPA [25], and increase the chances of finding new autoantibodies against low-abundant proteins. Only combinations of multiple TAAs will cope with the individual variability and low prevalence of specific autoantibodies in cancer patients to render a diagnostic assay with the sufficient sensitivity and specificity [26]. By using two different microarray approaches, human recombinant protein microarrays (ProtoArray) and T7-phage microarrays, we previously identified several TAAs with high CRC diagnostic potential [18, 27]. However, to improve the diagnostic accuracy with a better and more robust predictor, we analyzed different combinations of the identified TAAs with a new set of samples. The new set of samples included a new panel of sera from CRC patients in A-D stages and healthy reference controls.

In this study, we used 11 previously reported TAAs for optimization of the predictor panel, since MST1/STK4 was coincident among both studies. We expressed and purified the six full-length recombinant proteins obtained from ProtoArray™, as well as SULF1 protein, and four phage-peptides derived from T7-phage microarrays. They were validated and tested by ELISA with an independent cohort of samples, different from the sera used in the discovery phase with the microarrays. The best TAA combination was composed of three phages and 4 proteins. This panel showed a high discriminatory power between CRC samples and controls (AUC = 94%) for all stages as well as for early CRC detection (AUC = 90%).

2 Materials and methods

2.1 Clinical information and serum collection.

The Institutional Ethical Review Boards of the Centro de Investigaciones Biológicas (CIB) and the Spanish National Research Council (CSIC) approved this study on biomarker discovery in colorectal cancer. Written informed consent was obtained from all patients. Serum samples for validation analysis were obtained from patients in the Bellvitge University Hospital & Institut Català d'Oncologia (Barcelona), Puerta de Hierro Hospital (Madrid) and Cabueñes Hospital (Gijón). Sample collection was approved by the Ethical Review Boards of these institutions.

For optimization studies, we used a new panel comprising 50 CRC samples (representative of the different Dukes stages A-D) and 46 healthy reference control sera (Table 1). Serum samples from control subjects were selected to match the median age and the same gender proportion that the CRC cohort. The median age was 70.8 years (range 23-90). All sera were processed as previously described [18, 27]. Samples were handled anonymously according to ethical and legal guidelines at the CSIC.

2.2 Proteins and phages.

Proteins were either internally produced or purchased from various sources. The 6xHis fusion proteins PIM1, MAPKAPK3, MST1/STK4, FGFR4, ACVR2B, SRC and SULF1 were expressed in *E. coli* and purified by affinity chromatography on a HiTrap Chelating column (GE Healthcare) followed by gel filtration on a Superdex 200 column (GE Healthcare), as previously described [18, 27].

For amplification of T7 phages displaying CRC-specific peptides, BL21 *E. coli* cells were grown until $OD_{600\text{ nm}} \geq 0.4$. Then, T7 phages displaying peptides for SREBF2, GRN, GTF2i and NHSL1 were used to infect bacteria for 3-4 h until lysis was observed. Bacterial lysates were centrifuged and phage-containing supernatants were stored at -80°C until use.

2.3 Dot blot and western blot analysis.

Dot blot analysis to determine the expression of FGFR4, SRC and MST1/STK4 in different healthy and cancer tissues was performed using DiscoverLight Human Tissue Arrays (Pierce). For western blot analysis, protein extracts from CRC paired tissues were prepared as described [28, 29]. Briefly, 50 µg of protein extracts were run in 10% SDS-PAGE and transferred to nitrocellulose membranes (Hybond-C extra). After blocking, membranes were incubated overnight at 4°C with optimized dilutions of specific mono- or polyclonal antibodies (FGFR4, Santa Cruz Biotechnology; SRC, R&D Systems; MST1, Cell Signaling). Immunodetection on the membranes was achieved by using peroxidase-conjugated secondary antibodies (Sigma). The ECL signal was developed with SuperSignal West Femto substrate (Thermo). The abundance of the proteins in dot blot and western blot assays was determined by densitometry using Quantity One 1D Analysis Software v4.6 (Bio-Rad Laboratories).

2.4 ELISA tests

ELISA experiments with full-length recombinant proteins were performed as described before [18]. Briefly, microtiter plates (Maxisorp, Nunc) were coated overnight with 0.3 µg of the purified proteins in 50 µl of PBS. For multiplex analysis, we coated the same well with 0.3 µg of each protein, PIM1, MAPKAPK3 and FGFR4, in 50 µl of PBS. After washing three times with PBS, plates were blocked with 3% skimmed milk in PBS (MPBS) for 2 h at room temperature. Following additional washing, serum samples (dilution 1:50 in 3% MPBS) were incubated for 2 h at room temperature. After washing, anti-human IgG peroxidase-labelled (DAKO) at a 1:3000 dilution in 3% MPBS was added for 2 h at room temperature. Then, the signal was developed with 3,3',5,5'-tetramethylbenzidine (TMB) substrate (Sigma). The reaction was stopped with 1M H₂SO₄ and absorption was measured at 450 nm. To test T7 phages displaying CRC peptides,

we used previously published protocols [27]. CEA concentration in serum was determined using a specific immunoassay test kit (MP Biomedicals).

2.5 Statistical analysis

For the analysis of ELISA datasets, a one-tailed Student's *t* test was performed, assuming unequal variances to assess whether the means of normal group and tumoral group were statistically different from each other. Each individual marker and the combinations of biomarkers were evaluated by receiver operating characteristics (ROC) curve analysis, and the corresponding area under the curve (AUC) was calculated using JMP[®] 7 (SAS) and the JMP Stepwise Regression option to eliminate unnecessary terms in the model.

Correlation between markers was performed to determine the overlap in the sera reactivity in order to discern whether the information provided by each TAA was, or not, partially redundant with other TAAs. ELISA results were plotted for all pairs of the 11 TAAs and Pearson's correlation calculated for all combinations using the R statistical computing system [30]. A low Pearson's correlation would mean that the reactivity with each TAA does not show a linear association with the reactivity to any other TAA. Whereas a high Pearson's correlation would indicate that one marker is partially redundant given the other. Therefore, it is more likely that we find a better combined classifier if the correlation is small between two markers than if the two markers are highly correlated. Note, however, that in some cases (see below the example of GTF2i and SREBF2 phages), markers with high correlation can, however, do a good patient discrimination when used together if, for instance, the ratio between the two markers differs between types of patients.

Logistic regression with variable selection using Akaike Information Criterion (AIC) was bootstrapped as previously described with minor modifications [18, 27]. Bootstrapping was applied to the complete procedure (i.e. including the variable selection step) to determine the effect of the combination of the biomarkers together with the structural variables sex and age

of the patients in order to determine their importance in the model since these variables usually affect most of the biological process. Within each bootstrap replication (and with the complete data set) variable selection was carried out using AIC. Logistic regression was fitted in R [30] using the function "glm" (from R) and the function "lrm" from the "rms" package for R by F. Harrell (version 3.3-3. <http://CRAN.R-project.org/package=rms> and [31]). Model validation via bootstrap was carried out with the validate function in the "rms" package [31, 32]. Variable selection included CEA in the model in order to clarify if the model would benefit of the measurement of the marker currently used in clinic.

3 Results and discussion

3.1 Identification of seven proteins and four phage-peptides as CRC biomarkers.

Using two different protein microarrays approaches [18, 27], we observed 43 proteins with significantly different autoantibody levels in CRC sera ($p \leq 0.04$) and 43 phages differentially recognized by CRC patients. A total of 11 TAAs were selected based on the statistical significance and diagnostic prevalence. From this set, seven full length recombinant proteins were expressed in *E. coli* and purified to homogeneity (Fig. 1) and four phages displaying CRC-specific peptides were amplified and purified. MST1/STK4, SULF1, ACVR2B, SRC, MAPKAPK3, FGFR4 and PIM1 recombinant proteins and NHSL1, SREBF2, GRN and GTF2i-like phage lysates were tested by ELISA to assess their ability to discriminate between CRC sera and healthy control sera. We used an independent cohort of 96 samples (50 colorectal cancer, including 17 samples from early colorectal cancer stages (A-B), and 46 control samples (Fig. 2).

For individual TAAs, autoantibody titers in CRC group were significantly higher than in control group by ELISA, except for ACVR2B, as expected [18] (Supplementary Table 1), p values were statistically significant, except for GRN phage. Using ROC curves analysis, we observed that all purified proteins and phages showed individual diagnostic potential for discriminating between CRC and reference control samples, with AUCs $\geq 52\%$ (Table 2). Interestingly, full-length recombinant proteins showed higher AUC values in comparison to phages displaying CRC-specific peptide TAAs, with AUCs up to 73.2% for MAPKAPK3. The specificity and sensitivity of the different markers was quite variable (Table 2). We also tested the ability of the TAAs for their prediction ability according to the stage of the patients for early diagnosis purposes. All the TAAs showed individual ability for discriminating early CRC samples from controls, with AUCs going from 53.8% for NHSL1 phage up to 79.2% for MST1/STK4 (Supplementary Table 2).

3.2 Expression of FGFR4, SRC and MST1/STK4 in colorectal cancer

Protein lysate arrays containing tissue extracts from colon, rectum and other cancer tissues along with matched controls were used to evaluate expression of FGFR4, SRC and MST1/STK4 in different pairs of cancer and normal tissues (Fig. 3A). Antibodies to FGFR4, SRC and MST1/STK4 exhibited a higher expression in colon cancer relative to matched healthy colon. MST1/STK4 also presented more expression in rectal cancer in comparison to normal rectum. SRC was not detected in rectal tissues and FGFR4 was present in normal rectum but not in rectal cancer. The three proteins exhibited quite different patterns of expression in other normal and cancer tissues (i.e. cervix) (Fig. 3A).

For further confirmation, FGFR4 and SRC were analyzed by immunoblotting with protein samples from colorectal cancer and reference cells (Fig. 3B) and paired normal/tumoral colorectal extracts (Fig. 3C). MST1/STK4 had been previously characterized [27]. Interestingly, we observed an overexpression of FGFR4 in two metastatic cell lines (KM12SM and SW48) and a lower and variable pattern in the rest of metastatic and non-metastatic CRC cell lines (Fig. 3B). SRC expression was quite variable among different CRC cell lines with basal level of expression in reference control cell lines. No significant expression for FGFR4 or SRC was observed in reference cells, except embryonic HEK293 (Fig. 3B). FGFR4 expression was cancer-specific in early cancer stages, with slight overexpression in some tumoral tissues, preferably at B and C Dukes' stages. SRC expression was barely different between tumor and normal adjacent tissue (Fig. 3C). Remarkably, two clear SRC bands were observed. The upper band corresponded to the SRC expected migration. The origin of the lower band is unclear, but it was SRC specific.

3.3 Classification of Tumor-Associated Antigens according to their reactivity with cancer patients

Then, we studied the correlation between the identified TAAs for redundancy and discrimination ability for CRC diagnosis. This information can be used to develop a multiplex

test, since those with high correlation values can be used together. We plotted the ELISA results for all pairs of TAAs. Many TAAs showed low correlation between them (≤ 0.34), indicating that their reactivity was not redundant and they discriminated better CRC from controls if tested separately (Supplementary Table 3). However, we found different scenarios for several pairs of TAAs. MAPKAPK3 and ACVR2B showed a small linear correlation (0.34) and, thus, gave together a good discriminatory ability (Fig. 4A), since the amount of redundant information each one of the markers provides about the other marker is small. In contrast, PIM1 and MAPKAPK3, with a correlation of 0.709, FGFR4 and PIM1 with 0.592 and FGFR4 and MAPKAPK3 with 0.635, showed a very similar predictive ability in recognizing the same patient's population, suggesting the using each of these three markers individually and separately, in the same model, would not improve the discrimination ability of a panel predictor for CRC diagnosis. Therefore, PIM1, MAPKAPK3 and FGFR4 can be tested together to enhance the diagnostic ability of the panel. Multiplexing the assay reduces the amount of sera needed and simplifies the test. Indeed, when multiplexing FGFR4, PIM1 and MAPKAPK3, the AUC, sensitivity and specificity increased to 79.7%, 84.8% and 79.6%, respectively (Table 2). For early diagnosis purposes, we observed again that multiplexing these antigens produced an increase in the AUC, sensitivity and specificity in comparison to the proteins separately (Supplementary Table 2).

Regarding the correlation between the other markers, we noticed a high correlation (0.82) for two phages: GTF2i and SREBF2. Phages GTF2i and SREBF2 appeared in most of the multivariate models fitted (see section 3.5, for description of bootstrapping). The AUC of each of these phages on their own is small, but GTF2i and SREBF2 together can do a good discrimination (Fig. 4B). For a fixed absorbance level for SREBF2 (e.g., orange dot-slash rectangle), the probability that a patient be considered disease-free increased with the levels of GTF2i. In contrast, if we fix the level of GTF2i (e.g., green slash rectangle) the probability that a patient is considered diseased increases as we increase the value of SREBF2. By using

those two variables (SREBF2 and GTF2i) the two classes can be separated, as the “normals” tend to be those that have a high GTF2i value relative to the amount of SREBF2. Although the two markers showed a high correlation (0.82), they can discriminate between tumor and normal patients because the ratio GTF2i/SREBF2 was different between tumors and controls (that ratio is higher for controls than tumors). Interestingly, although the ratio GTF2i/SREBF2 was highly discriminatory, there was not enough evidence of significant interaction (p value: 0.1448) to specify an interaction between them in the logistic regressions. In summary, correlation between independent variables is a property of the relationship of these independent variables, but is not inherently related to the joint relationship of the independent and the dependent variables (the disease status in this case) and their discriminatory capability. It should be analyzed case by case.

3.4 Building an optimized predictor panel for colorectal cancer

Next, we investigated which combination of proteins and phages shows the best diagnostic ability for CRC. ELISA data for each TAA and the combination of FGFR4, PIM1 and MAPKAPK3, which were used together, were fitted to a logistic curve and performed logistic regressions. Different models were produced using different combinations of phages and proteins.

The combination of three phages displaying peptides homologues to GTF2i, NHSL1 and SREBF2 and four proteins PIM1, ACVR2B, MAPKAPK3 and FGFR4, with the data of the last three proteins altogether, showed the best ability to discriminate among CRC serum samples and reference sera (Fig. 5A). ROC curves obtained after combining the results for the seven TAAs showed an AUC of 94%, with a sensitivity and specificity of 89.1% and 90%, respectively.

Interestingly, using the 7 TAAs optimized predictor with CRC serum samples corresponding to early A-B stages, we found an AUC of 90%, with a sensitivity and specificity of 88.2% and 82.6%, respectively (Fig. 5B). With these same samples, CEA gave an AUC of 74%, with a sensitivity and specificity of 82.4% and 63.0%, respectively (Fig. 5C).

3.5 Bootstrapping analysis confirmed the predictor panel

We performed bootstrapping (using 1000 bootstrap samples) to obtain a bias-corrected AUC and to validate this TAA combination for CRC screening. The initial model included linear terms for all phages and proteins, together with the gender and age of the patients. Since the gender and age of the patients usually affect multiple biological processes, they were included in the analysis to know if they affect the panel of biomarkers. In addition, we also combined the model with and without CEA quantification in all samples to determine if CEA levels in serum could improve the model. With CEA in this model, the value of the bias-corrected AUC was 92.5%. The model without CEA, gave a bias-corrected AUC of 90.5%. Both models got better AUC than CEA alone, which gave a bias-corrected AUC of 82.4%.

After variable selection, using Akaike Information Criterion, the final model with CEA retained the same three phages displaying homologue peptides to GTF2i, NHSL1 and SREBF2 and four full-length recombinant proteins PIM1, ACVR2B, MAPKAPK3 and FGFR4 (Table 3). To avoid an overestimation of the predictive capacity of the model, we obtained bias-corrected estimates of the AUC, by bootstrapping the complete process of variable selection. The bias-corrected AUC was 91.6%. Bootstrapping also provided information on the stability of the selection procedure: among the bootstrapped models, most contained between 7 and 11 variables. Interestingly, two of the variables appeared in all the models: GTF2i and SREBF2 phages. ACVR2B appeared in 96% of the models, PIM1, MAPKAPK3 and FGFR4 in 92%, CEA in 86% and, finally, NHSL1 phage in 75% of the models. Regarding the final model without CEA after variable selection, we observed that the model contained the same variables, with a bias-corrected AUC of 90.5%, slightly lower than the model including CEA (**Supplementary Table 4**).

Then, we tested the value of the prediction for early diagnosis purposes. We started from the best model composed of three phages and four proteins plus sex and age, as structural variables. The bias-corrected AUC using bootstrapping was 90% for stages A-B. In

306 this case, we did not include CEA in the analysis, since CEA is not recommended for early
307 diagnosis [2, 3]. If we apply the same test with the CEA values, the bias-corrected AUC was
308 74.2% for A-B. As expected our panel performed better than CEA for stages A and B.

4 Conclusions

In this report, we have optimized a TAA combination to improve CRC diagnosis, starting from two different panels of TAAs previously reported [18, 27]. The final combination was composed of three phages and four recombinant proteins. Interestingly, SRC that was excluded from this model did not show a clear overexpression in cancer respect to normal samples, suggesting a potential relevance for overexpression in TAA selection. This final panel gave higher specificity and sensitivity (AUC: 94%; bias-corrected AUC: 92.5%) to diagnose CRC than any previously reported serum biomarkers, especially for early CRC stages (AUC: 90%; bias-corrected AUC: 90%). These values are very promising for a non-invasive diagnosis of early stage colon tumors.

After observing a high correlation between the three markers: PIM1, MAPKAPK3 and FGFR4, they were combined in a multiplex assay, instead of testing one antigen at a time. However, ACVR2B and three phages were tested separately as they react with different patients. Further experiments might increase the specificity and sensitivity of the diagnostic tests [20] by testing i) the use of full-length recombinant proteins instead of the corresponding phages, ii) the most appropriate expression system (i.e. bacteria or insect cells) and iii) the use of different tags in the fusion protein (GST or 6xHis) that could affect the binding of the autoantibodies to its target.

In summary, we have described an improved CRC panel with a superior diagnostic capacity for CRC. Furthermore, since most of the identified TAAs are kinases, studies to determine if they could be new targets for therapeutic intervention in CRC are in progress.

330 ACKNOWLEDGEMENTS

331 RB was recipient of a JAE-DOC Contract of the CSIC. This research was supported by grants
332 from the Spanish Ministry of Education and Science BIO2009-08818 (JIC) and BIO2009-12458
333 (R.D-U.). Grant to stable cancer research groups of the Asociación Española Contra el Cáncer
334 and grants from the Program of Biosciences “Colomics2” (Comunidad de Madrid), Instituto de
335 Salud Carlos III (FIS 05/1006 and 08/1635), the CIBERESP G55, the “Acción transversal del
336 cancer” and the Proteored platform.

5 References

- [1] Edwards BK, Ward E, Kohler BA, Ehemann C, Zauber AG, Anderson RN, et al. Annual report to the nation on the status of cancer, 1975-2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer*. 2010;116:544-73.
- [2] Duffy MJ, van Dalen A, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, et al. Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. *Eur J Cancer*. 2007;43:1348-60.
- [3] Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol*. 2006;24:5313-27.
- [4] Orenes-Pinero E, Barderas R, Rico D, Casal JI, Gonzalez-Pisano D, Navajo J, et al. Serum and tissue profiling in bladder cancer combining protein and tissue arrays. *J Proteome Res*. 2010;9:164-73.
- [5] Chapman C, Murray A, Chakrabarti J, Thorpe A, Woolston C, Sahin U, et al. Autoantibodies in breast cancer: their use as an aid to early diagnosis. *Ann Oncol*. 2007;18:868-73.
- [6] Chapman CJ, Murray A, McElveen JE, Sahin U, Luxemburger U, Tureci O, et al. Autoantibodies in lung cancer: possibilities for early detection and subsequent cure. *Thorax*. 2008;63:228-33.
- [7] Hudson ME, Pozdnyakova I, Haines K, Mor G, Snyder M. Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. *Proceedings of the National Academy of Sciences of the United States of America*. 2007;104:17494-9.
- [8] Soussi T. p53 Antibodies in the sera of patients with various types of cancer: a review. *Cancer Res*. 2000;60:1777-88.

- [9] Chen G, Gharib TG, Huang CC, Thomas DG, Shedden KA, Taylor JM, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. Clin Cancer Res. 2002;8:2298-305.
- [10] Chen Y, Lin P, Qiu S, Peng XX, Looi K, Farquhar MG, et al. Autoantibodies to Ca²⁺ binding protein Calnuc is a potential marker in colon cancer detection. Int J Oncol. 2007;30:1137-44.
- [11] Chen YT, Scanlan MJ, Sahin U, Tureci O, Gure AO, Tsang S, et al. A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. Proceedings of the National Academy of Sciences of the United States of America. 1997;94:1914-8.
- [12] Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. Nature. 2000;404:770-4.
- [13] Ulanet DB, Torbenson M, Dang CV, Casciola-Rosen L, Rosen A. Unique conformation of cancer autoantigen B23 in hepatoma: a mechanism for specificity in the autoimmune response. Proceedings of the National Academy of Sciences of the United States of America. 2003;100:12361-6.
- [14] von Mensdorff-Pouilly S, Petrakou E, Kenemans P, van Uffelen K, Verstraeten AA, Snijdwint FG, et al. Reactivity of natural and induced human antibodies to MUC1 mucin with MUC1 peptides and n-acetylgalactosamine (GalNAc) peptides. International journal of cancer. 2000;86:702-12.
- [15] Casal JI, Barderas R. Identification of cancer autoantigens in serum: toward diagnostic/prognostic testing? Mol Diagn Ther. 2010;14:149-54.
- [16] Anderson KS, LaBaer J. The sentinel within: exploiting the immune system for cancer biomarkers. Journal of proteome research. 2005;4:1123-33.
- [17] Adam JK, Odhav B, Bhoola KD. Immune responses in cancer. Pharmacology & therapeutics. 2003;99:113-32.

386 [18] Babel I, Barderas R, Diaz-Uriarte R, Martinez-Torrecuadrada JL, Sanchez-Carbayo M, Casal
387 JI. Identification of tumor-associated autoantigens for the diagnosis of colorectal cancer in
388 serum using high density protein microarrays. *Mol Cell Proteomics*. 2009;8:2382-95.

389 [19] Chatterjee M, Ionan A, Draghici S, Tainsky MA. Epitomics: global profiling of immune
390 response to disease using protein microarrays. *Omics*. 2006;10:499-506.

391 [20] Sreekumar A, Laxman B, Rhodes DR, Bhagavathula S, Harwood J, Giacherio D, et al.
392 Humoral immune response to alpha-methylacyl-CoA racemase and prostate cancer. *J Natl*
393 *Cancer Inst*. 2004;96:834-43.

394 [21] Wang X, Yu J, Sreekumar A, Varambally S, Shen R, Giacherio D, et al. Autoantibody
395 signatures in prostate cancer. *N Engl J Med*. 2005;353:1224-35.

396 [22] Scanlan MJ, Chen YT, Williamson B, Gure AO, Stockert E, Gordan JD, et al. Characterization
397 of human colon cancer antigens recognized by autologous antibodies. *International journal of*
398 *cancer*. 1998;76:652-8.

399 [23] Scanlan MJ, Welt S, Gordon CM, Chen YT, Gure AO, Stockert E, et al. Cancer-related
400 serological recognition of human colon cancer: identification of potential diagnostic and
401 immunotherapeutic targets. *Cancer Res*. 2002;62:4041-7.

402 [24] Gnjjatic S, Ritter E, Buchler MW, Giese NA, Brors B, Frei C, et al. Seromic profiling of
403 ovarian and pancreatic cancer. *Proceedings of the National Academy of Sciences of the United*
404 *States of America*. 2010;107:5088-93.

405 [25] Klade CS, Voss T, Krystek E, Ahorn H, Zatloukal K, Pummer K, et al. Identification of tumor
406 antigens in renal cell carcinoma by serological proteome analysis. *Proteomics*. 2001;1:890-8.

407 [26] Anderson KS, Ramachandran N, Wong J, Raphael JV, Hainsworth E, Demirkan G, et al.
408 Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in
409 breast cancer. *J Proteome Res*. 2008;7:1490-9.

410 [27] Babel I, Barderas R, Diaz-Uriarte R, Moreno V, Suarez A, Fernandez-Acenero MJ, et al.
 411 Identification of MST1/STK4 and SULF1 proteins as autoantibody targets for the diagnosis of
 412 colorectal cancer by using phage microarrays. *Mol Cell Proteomics*. 2011;10:M110 001784.

413 [28] Alfonso P, Nunez A, Madoz-Gurpide J, Lombardia L, Sanchez L, Casal JI. Proteomic
 414 expression analysis of colorectal cancer by two-dimensional differential gel electrophoresis.
 415 *Proteomics*. 2005;5:2602-11.

416 [29] Barderas R, Desmet J, Timmerman P, Meloen R, Casal JI. Affinity maturation of antibodies
 417 assisted by in silico modeling. *Proceedings of the National Academy of Sciences of the United*
 418 *States of America*. 2008;105:9029-34.

419 [30] R Core Development Team. R: a Language and Environment for Statistical Computing. The
 420 R Foundation for Statistical Computing, Vienna. 2009.

421 [31] Harrell FE. *Regression Modeling Strategies*. Springer. 2001.

422 [32] Venables WN, Ripley BD. *Modern applied statistics with S*. Springer. 2001.

423

424 Table 1. Clinical information of the colorectal cancer patients tested during the study.

		Serum	Total	Age (years)	Sex	
					Female	Male
Screening with protein microarrays	ProtoArrays	Control	8	65.3 ± 13.6	2	6
		CRC (Duke's stage D)	12	64.5 ± 12.1	3	8
	T7-phage	Control	8	63.5 ± 15.6	6	9
		CRC (all stages)	15	66.3 ± 9.3	4	11
		Duke's stage A	3	75.7 ± 5.3	1	2
		Duke's stage B	3	62.3 ± 6.5	1	2
		Duke's stage C	5	67.8 ± 10.8	1	4
		Duke's stage D	4	60.2 ± 3.3	1	3
	Validation	Control	46	60.9 ± 11.4	18	28
		CRC (all stages)	50	70.8 ± 15.7	17	33
		Duke's stage A	6	60.0 ± 11.0	3	3
		Duke's stage B	11	73.0 ± 10.1	3	8
		Duke's stage C	16	76.0 ± 16.6	8	8
		Duke's stage D	17	68.7 ± 12.1	3	14

425 * The age of the different patients is represented as mean ± standard deviation.

426 **Table 2. TAA capacity to discriminate between CRC and control sera.**

Recombinant protein	Phage	Specificity (%)	Sensitivity (%)	AUC (%) ^b
FGFR4	-	61.9	71.1	69.2
MAPKAPK3	-	74.0	72.7	73.2
PIM1	-	83.3	48.1	65.2
ACVR2B	-	76.2	59.6	66.6
MST1/STK4	-	71.4	63.5	69.1
SRC	-	61.9	67.3	70.9
SULF1	-	78.6	62.0	67.1
PIM1-MAPKAPK3-FGFR4 ^a	-	79.6	84.8	79.7
-	GRN	57.1	55.0	52.3
-	GTF2i	52.2	60.1	57.3
-	NHSL1	50.2	56.1	59.4
-	SREBF2	69.6	54.4	61.0

427 ^a, These proteins were combined in a multiplexing assay.

428 ^b, AUC (%). Area under the curve in percentage.

Table 3. Predictor panel for colorectal cancer diagnosis of the final model after variable selection with Akaike Information Criterion

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-6.593	3.844	-1.715	0.086
Age	0.056	0.033	1.688	0.091
NHSL1-phage	6.152	2.917	2.109	0.035*
GTF2i-phage ^a	-10.597	3.313	-3.198	0.001**
SREBF2-phage	13.960	4.698	2.972	0.003**
PIM1, MAPKAPK3 and FGFR4	11.924	3.402	3.505	<.001***
ACVR2B ^b	-9.578	3.324	-2.881	0.004**
Sex	1.506	0.842	1.789	0.074
CEA ^c	1.633	0.665	2.455	0.014*

The variable selection statistics show the contribution and significance of the phages and human full-length recombinant proteins in the model. The intercept is the value of the logit, or $\log(p/(1-p))$, where p is the probability of being CRC patient, when the value of all the other terms in the model is 0. The intercept, therefore, represents the differences in the probability of being CRC or control patient, even if there are no effects of the independent variables. Estimate, estimated coefficient (slope). A negative coefficient would indicate that the probability of being CRC patient decreased when the immune response of the marker increases; Std. Error, the standard error of the specified variable; z value or Wald statistic, which is exactly equal to the estimated coefficient divided by its standard error; Pr(>|z|), p-value from the Wald test for that specific coefficient comparing the z-value against a standard normal; * to ***, grade of significance.

^a. GTF2i-phage coefficient is negative although its reactivity is higher in CRC than control due to the high correlation in the immune response between GTF2i-phage and SREBF2-phage (see legend to Fig.4B).

^b, ACVR2B coefficient is negative because the reactivity with this marker is higher in controls than in CRC patients.

^c CEA was included in the model to determine if CEA levels in serum could improve the model. Alternatively, we also performed variable selection without CEA in the model (see Supplementary Table 4).

Legend to Figures

Figure 1. Full length-recombinant TAAs were expressed in *E. coli* and purified to homogeneity. Recombinant proteins (1 µg) were separated by 10% SDS-PAGE and stained with Coomassie Blue to verify their purity. Identity of the purified proteins was confirmed by MALDI-TOF mass spectrometry.

Figure 2. Workflow followed for the identification and validation of autoantibodies and their respective target proteins as colorectal cancer diagnostic biomarkers using different sera collection and statistical analyses.

Figure 3. Expression of SRC, MST1/STK4 and FGFR4 in different cancer samples. (A) DiscoverLight Human Tissue Arrays spotted with protein lysates were probed with SRC, FGFR4 or MST1 specific antibodies. The tissues spotted in the microarrays are indicated at the bottom of the figure. (B) Western blot analysis of colorectal cancer cell lines in comparison to reference control cell lines. 50 µg of cell extracts were separated by 10% SDS-PAGE and transferred to nitrocellulose membranes and probed with specific antibodies to FGFR4 and SRC. The abundance of each marker is represented as bar graph after quantification of WB by densitometry. Tubulin was used as loading control. (C) Western blot analysis of protein extracts from paired normal (N) and tumoral (T) tissues from six CRC patients (Dukes stages A, 211 and 299; B, 699 and 700; and C, 704 and 713). Tissue protein extracts were resolved and transferred onto nitrocellulose membranes and incubated with the specific antibodies. The abundance of each marker is represented as bar graph after quantification of WB by densitometry. **Black triangles** in SRC indicate the band that was quantified by densitometry.

Figure 4. Correlation study of redundancy between colorectal cancer TAAs. (A) Relationship between MAPKAPK3, ACVR2B and status (CRC versus control). The linear correlation between

MAPKAPK3 and ACVR2B was 0.34. Black curve represents the line from a logistic regression model that includes only MAPKAPK3 and ACVR2B, with ACVR2B modeled as a restricted cubic spline with three knots to use a non-linear, low complexity model. More knots were not used to prevent overfitting. (B) Relationship between GTF2i-phage, SREBF2-phage and status (CRC versus control). The discrimination of CRC and control samples using SREBF2-phage and GTF2i-phage was good. Although these markers show a high correlation (0.82), they can discriminate CRC from control samples because of the ratio GTF2i-phage/SREBF2-phage is larger for controls than CRC samples. The dotted black line separates CRC samples from controls with a logistic regression model that included only GTF2i-phage and SREBF2-phage. As an illustration, within the orange dot slash rectangle, the probability that a patient is control increased with the levels of autoantibodies to GTF2i-phage for a fixed absorbance level of SREBF2-phage. Within the green slash rectangle, the probability that a patient is CRC positive increased as the reactivity with SREBF2-phage increased for a fixed absorbance level of GTF2i-phage.

Figure 5. Sensitivity and specificity of the optimized predictor panel. ROC curves were based on multiplex analyses of the three phages and four proteins using a total of 96 samples (50 samples from CRC patients and 46 healthy reference controls. A) CRC samples versus reference controls. B) CRC samples corresponding to Duke's stages A and B versus reference controls. C) CEA value in CRC samples corresponding to Duke's stages A and B versus reference controls. The corrected AUC after bootstrapping is also indicated in the figure.

Figure 1

Figure 1

Figure 2

Figure 2

Figure 3

Figure 3

Figure 4

Figure 4

Figure 5

Figure 5