

## Supplementary file 1

# The diagnostic and prognostic value of *Clorf174* in colorectal cancer

Elham Nazari<sup>1,2\*#</sup>, Ghazaleh Khalili-Tanha<sup>2,3</sup>, Ghazaleh Pournali<sup>2</sup>, Fatemeh Khojasteh-Leylakoohi<sup>3</sup>, Hanieh Azari<sup>3</sup>, Mohammad Dashtiahangar<sup>4</sup>, Hamid Fiuji<sup>5#</sup>, Zahra Yousefli<sup>2,3</sup>, Alireza Asadnia<sup>2,3</sup>, Mina Maftooh<sup>2,6</sup>, Hamed Akbarzade<sup>2</sup>, Mohammadreza Nassiri<sup>7</sup>, Seyed Mahdi Hassanian<sup>2</sup>, Gordon A Ferns<sup>8</sup>, Godefridus J Peters<sup>5,9</sup>, Elisa Giovannetti<sup>5,10</sup>, Jyotsna Batra<sup>11,12</sup>, Majid Khazaei<sup>2</sup>, Amir Avan<sup>2,12\*</sup>

<sup>1</sup>Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup>Metabolic Syndrome Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>3</sup>Medical Genetics Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>4</sup>School of Medicine, Gonabad University of Medical Sciences, Gonabad, Iran

<sup>5</sup>Department of Medical Oncology, Cancer Center Amsterdam, Amsterdam U.M.C., VU. University Medical Center (VUMC), Amsterdam, The Netherlands

<sup>6</sup>College of Medicine, University of Warith Al-Anbiyaa, Karbala, Iraq

<sup>7</sup>Recombinant Proteins Research Group, The Research Institute of Biotechnology, Ferdowsi University of Mashhad, Mashhad, Iran

<sup>8</sup>Brighton & Sussex Medical School, Division of Medical Education, Falmer, Brighton, Sussex BN1 9PH, UK

<sup>9</sup>Professor In Biochemistry, Medical University of Gdansk, Gdansk, Poland

<sup>10</sup>Cancer Pharmacology Lab, AIRC Start up Unit, Fondazione Pisana per La Scienza, Pisa, Italy

<sup>11</sup>Centre for Genomics and Personalised Health, Queensland University of Technology, Brisbane 4059, Australia

<sup>12</sup>Faculty of Health, School of Biomedical Sciences, Queensland University of Technology, Brisbane 4059, Australia

#Equally contributed as first author.

**Table S1.** Characteristics of population

Characteristics	Sub groups	Number (%)
Sex	Male	185 (46.5)
	Female	213 (53.3)
History of colon polyp	yes	72 (89.9)
	no	326(18.1)
History of neoadjuvant treatment	yes	2 (0.5)
	no	396 (95.5)
Cancer status	Tumor free	299(75.1)
	With tumor	99 (24.9)
Stages	1 and 2	207 (52)
	3	117 (29.4)
	4	55 (13.8)
	missing	19 (4.8)
	Metastasis (M)	No
	Yes	54(13.6)
Lymph node involvement (N)	No	225 (56.5)
	Yes	101 (25.4)
	No data	72(18.1)

Depth of tumor invasion (T)	No data	2(.5)
	T1	12(3.0)
	T2	56(14.1)
	T3	278(69.8)
	T4	14(3.5)
	T5	26(6.5)
	T6	10(2.5)
Synchronous colon cancer present	Yes	139(3.3)
	No	385(96.7)
Veous invasion	Yes	79(19.8)
	No	319(82)
Status	Alive	349(87.7)
	Death	49(12.3)
patient.number_of_first_degree_relatives_with_cancer_d iagnosis	No	353(88.7)
	Yes	39(9.8)
	No data	69(1.5)
	Mean	$\pm$ SD
Month	6.25	0.107
patient.number_of_lymphnodes_positive_by_he	1.99	0.225
patient.lymph_node_examined_count	19.47	0.717
Height	122.43	3.83
Age	64.85	0.663
Day	15.26	0.452
<b>Total Number=398</b>		

**Table S2.** Clinicopathological characteristics of CRC patients

Clinicopathological Variables	No. of patient (%) /mean $\pm$ SD
Patients	47
Mean age (Years, mean $\pm$ SD)	55.1 $\pm$ 16.8
Sex	
Male	26 (55.3)
Female	21 (44.7)
Depth of tumor invasion (T)	
T1	2 (4)
T2	11 (26)
T3	27 (63)

T4	3 (7)
Lymph node involvement (N)	
Yes	36 (80)
No	9 (20)
Metastasis (M)	
Yes	6 (20.7)
No	23 (79.3)
Family history	
Yes	7 (24.1)
No	22 (75.9)

**Table S3. Confusion Matrix**

	true cancer	true Normal
pred. cancer	52	3
pred. Normal	3	30
class recall	94.55%	90.91%

**Table S4. Results of the diagnostic tests performed between the Adenocarcinoma colorectal patients and the controls.**

Stages	Markers		Cutoff	AUC (95%CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Stage 1 & 2	Popdc3		0.679	0.668	0.691	0.606	0.917	0.238
	TPO		0.873	0.786	0.763	0.848	0.969	.364
	TLX2		0.783	0.875	0.758	0.939	0.987	0.383
	Combination75	CA14-POPDC3-TLX2-TPO	0.849	0.876	0.841	0.879	0.978	0.468

	<b>Combination76</b>	CA14-SNAP91-TLX2-TPO	0.862	0.892	0.841	0.939	0.989	0.484
	<b>Combination77</b>	CATSPER1-GRIA4-POPDC3-SNAP91	0.797	0.897	0.879	0.879	0.978	0.537
	<b>Combination78</b>	CATSPER1-GRIA4-POPDC3-TLX2	0.840	0.912	0.841	0.909	0.983	0.476
	<b>Combination79</b>	CATSPER1-GRIA4-POPDC3-TPO	0.837	0.880	0.836	0.879	0.977	0.460
<b>Stage 3</b>	<b>LOC100128239</b>		0.790	0.800	0.778	0.788	0.929	0.500
	<b>Combination68</b>	FAM151A-LILRA4-LOC100128239-MESTI	0.740	0.894	0.863	0.879	0.962	0.644
	<b>Combination69</b>	FAM151A-LILRA4-LOC100128239-SH2D6	0.874	0.908	0.726	0.970	0.988	0.500
	<b>Combination70</b>	FAM151A-LILRA4-LRAT-MESTIT1	0.679	0.907	0.906	0.818	0.946	0.711
	<b>Combination71</b>	FAM151A-LILRA4-LRAT-SH2D6	0.697	0.918	0.923	0.909	0.769	0.973
	<b>Combination72</b>	FAM151A-LILRA4-MESTIT1-SH2D6	0.672	0.903	0.915	0.818	0.730	0.947
<b>Stage4</b>	<b>DIRC1</b>		0.529	0.858	0.727	0.879	0.909	0.659
	<b>FAM118A</b>		0.625	0.672	0.618	0.697	0.773	0.523
	<b>AKAP4</b>		0.683	0.700	0.400	0.909	0.880	0.476
	<b>Combination159</b>	AKAP4-C1orf174-DIRC1-PROZ-SKIL	0.449	0.924	0.927	0.788	0.879	0.867
	<b>Combination160</b>	AKAP4-C1orf174-DIRC1-PROZ-SLC29A4	0.567	0.947	0.909	0.939	0.962	0.861
	<b>Combination161</b>	AKAP4-C1orf174-DIRC1-PROZ-SV2C	0.602	0.924	0.855	0.879	0.922	0.784

	<b>Combination16 2</b>	AKAP4- C1orf174- DIRC1-SKIL- SLC29A4	0.660	0.956	0.909	0.939	0.96 2	0.86 1
	<b>Combination16 3</b>	AKAP4- C1orf174- DIRC1-SKIL- SV2C	0.551	0.939	0.873	0.879	0.92 3	0.80 6
<b>Prognostic biomarker s</b>	<b>HPRT1</b>		0.473	0.792	0.927	0.606	0.79 7	0.83 3
	<b>RPIA</b>		0.541	0.832	0.836	0.879	0.92 0	0.76 3
	<b>ATP8B1</b>		0.566	0.910	0.909	0.879	<b>0.92 6</b>	0.85 3
	<b>Combination29 1</b>	ATP8B1- C1orf174- CASS4- KCNIP2	0.574	0.917	0.945	0.879	0.92 9	0.90 6
	<b>Combination29 2</b>	ATP8B1- C1orf174- CASS4- KCNK13	0.533	0.950	0.945	0.879	0.92 9	0.90 6
	<b>Combination29 3</b>	ATP8B1- C1orf174- CASS4-RPIA	0.468	0.922	0.927	0.879	0.92 7	0.87 9
	<b>Combination29 4</b>	ATP8B1- C1orf174- CASS4- SRRM5	0.570	0.932	0.909	0.879	0.92 6	0.85 3
	<b>Combination29 5</b>	ATP8B1- C1orf174- CASS4- ZNF805	0.498	0.938	0.909	0.879	0.92 6	0.85 3

Table S5. GLM model analysis for diagnosis testing.

Stages	Markers and combination	Coefficients		Degrees of Freedom	AIC	Residual Deviance
Stage 1&2	Popdc3	CA14-POPDC3- TLX2-TPO	(Intercept) log(POPDC3 + 1) 0.4692 1.2141	239	186.3	182.3

	TPO		(Intercept) log(TPO + 1) -0.2648 2.0347	239	164.7	160.7
	TLX2		(Intercept) log(TLX2 + 1) -2.093 3.307	239	146.1	142.1
	Combination75	CA14-POPDC3-TLX2-TPO	(Intercept) log(CA14 + 1) log(POPDC3 + 1) log(TLX2 + 1) log(TPO + 1) -3.0882 0.6041      0.3574 2.4817      0.9510	239	144.8	134.8
	Combination76	CA14-SNAP91-TLX2-TPO	(Intercept) log(CA14 + 1) log(SNAP91 + 1) log(TLX2 + 1) log(TPO + 1) -2.9629 0.5356      1.2567 2.1968      0.3087	239	139.1	129.1
	Combination77	CATSPER1-GRIA4-POPDC3-SNAP91	(Intercept) log(CATSPER1 + 1) log(GRIA4 + 1) log(POPDC3 + 1) log(SNAP91 + 1) 2.1137 -2.7311 1.8172 0.6307 1.3686	239	129.5	119.5
	Combination78	CATSPER1-GRIA4-POPDC3-TLX2	(Intercept) log(CATSPER1 + 1) log(GRIA4 + 1) log(POPDC3 + 1) log(TLX2 + 1) 0.1263 -2.0790 2.1089 0.2182 2.3513	239	123.7	113.7
	Combination79	CATSPER1-GRIA4-POPDC3-TPO	(Intercept) log(CATSPER1 + 1) log(GRIA4 + 1) log(POPDC3 + 1) log(TPO + 1)	239	135.3	125.3

			2.0075 -2.4791 2.0728 0.5912 0.9887			
Stage3	DIRC1		(Intercept) log(DIRC1 + 1) 17.79 - 10.03	87	82.44	78.44
	FAM118A		(Intercept) log(FAM118A + 1) -11.972 7.418	87	111.4	107.4
	LOC100128239		(Intercept) log(LOC100128239 + 1) -1.269 2.535	149	133	129
	Combination68	FAM151A- LILRA4- LOC100128239- MESTI	(Intercept) log(FAM151A + 1) log(LILRA4 + 1) log(LOC100128239 + 1) -9.067 1.512 2.660 2.261 log(MESTIT1 + 1) 2.297	149	101.7	91.7
	Combination69	FAM151A- LILRA4- LOC100128239- SH2D6	(Intercept) log(FAM151A + 1) log(LILRA4 + 1) log(LOC100128239 + 1) -7.006 1.027 2.093 2.735 log(SH2D6 + 1) 1.617	149	99.99	89.99
	Combination70	FAM151A- LILRA4-LRAT- MESTIT1	(Intercept) log(FAM151A + 1) log(LILRA4 + 1) log(LRAT + 1) log(MESTIT1 + 1) -8.586 1.516 2.208 2.234 2.107	149	98.98	88.98

	Combination71	FAM151A- LILRA4-LRAT- SH2D6	(Intercept) log(FAM151A + 1) log(LILRA4 + 1) log(LRAT + 1) log(SH2D6 + 1) -7.192 1.034      1.679 2.810      1.755	149	94.9	84.9
	Combination72	FAM151A- LILRA4- MESTIT1- SH2D6	(Intercept) log(FAM151A + 1) log(LILRA4 + 1) log(MESTIT1 + 1) log(SH2D6 + 1) -10.3772 0.8292      2.4469 3.7681      2.0591	149	101.3	91.34
Stage4	AKAP4		(Intercept) log(AKAP4 + 1) 6.554      - 3.552	87	109.1	105.1
	Combination159	AKAP4- C1orf174- DIRC1-PROZ- SKIL	(Intercept) log(AKAP4 + 1) log(C1orf174 + 1) log(DIRC1 + 1) log(PROZ + 1) -20.569 -0.977 13.435      - 10.236      - 0.813 log(SKIL + 1) 10.731	87	69.48	57.48
	Combination160	AKAP4- C1orf174- DIRC1-PROZ- SLC29A4	(Intercept) log(AKAP4 + 1) log(C1orf174 + 1) log(DIRC1 + 1) log(PROZ + 1) 17.0511 -1.8067 8.0107      - 12.5232      - 0.7747 log(SLC29A4 + 1) -4.1949	87	58.45	46.45
	Combination161	AKAP4- C1orf174- DIRC1-PROZ- SV2C	(Intercept) log(AKAP4 + 1) log(C1orf174 + 1) log(DIRC1 + 1) log(PROZ + 1) 7.6504 -0.6973 5.6637      -	87	70.49	58.49



			11.5992 - 0.9944 log(SV2C + 1) 3.8315			
	Combination162	AKAP4- C1orf174- DIRC1-SKIL- SLC29A4	(Intercept) log(AKAP4 + 1) log(C1orf174 + 1) log(DIRC1 + 1) log(SKIL + 1) -1.558 -1.515 10.155 - 12.375 7.357 log(SLC29A4 + 1) -3.634	87	55.6	43.6
	Combination163	AKAP4- C1orf174- DIRC1-SKIL- SV2C	(Intercept) log(AKAP4 + 1) log(C1orf174 + 1) log(DIRC1 + 1) log(SKIL + 1) -16.4929 -0.6158 9.4753 - 11.4548 9.9866 log(SV2C + 1) 3.0440	87	64.62	52.62
Prognostic biomarkers	HPRT1		(Intercept) log(HPRT1 + 1) -18.209 2.825	87	98.94	94.94
	RPIA		(Intercept) log(RPIA + 1) -19.089 2.854	87	95.6	91.6
	ATP8B1		(Intercept) log(ATP8B1 + 1) 31.781 - 3.639	87	65.15	61.15
	Combination291	ATP8B1- C1orf174- CASS4-KCNIP2	(Intercept) log(ATP8B1 + 1) log(C1orf174 + 1) log(CASS4 + 1) log(KCNIP2 + 1) 29.2408 -3.5468 0.5576 0.0557 - 0.6540	87	67.13	57.13
	Combination292	ATP8B1- C1orf174-	(Intercept) log(ATP8B1 + 1)	87	57.9	47.9

		CASS4- KCNK13	log(C1orf174 + 1) log(CASS4 + 1) log(KCNK13 + 1) 39.9124 -5.0667 0.4295 - 2.7568 3.6532			
	Combination293	ATP8B1- C1orf174- CASS4-RPIA	Intercept) log(ATP8B1 + 1) log(C1orf174 + 1) log(CASS4 + 1) log(RPIA + 1) 8.7949 -3.0909 1.0238 0.2464 1.6263	87	66.85	56.85
	Combination294	ATP8B1- C1orf174- CASS4-SRRM5	(Intercept) log(ATP8B1 + 1) log(C1orf174 + 1) log(CASS4 + 1) log(SRRM5 + 1) 12.5016 -3.0863 1.7559 - 0.1757 1.5242	87	65.53	55.53
	Combination295	ATP8B1- C1orf174- CASS4-ZNF805	(Intercept) log(ATP8B1 + 1) log(C1orf174 + 1) log(CASS4 + 1) log(ZNF805 + 1) -9.8351 -3.3892 3.3102 0.1482 3.6330	87	62.95	52.95

Table S6. Clinicopathological characteristics of CRC patients

<b>Clinicopathological Variables</b>	<b>No. of patient (%) /mean ± SD</b>
Patients	47
Mean age (Years, mean ± SD)	55.1 ± 16.8
Sex	
Male	26 (55.3)
Female	21 (44.7)

Depth of tumor invasion (T)	
T1	2 (4)
T2	11 (26)
T3	27 (63)
T4	3 (7)
Lymph node involvement (N)	
Yes	36 (80)
No	9 (20)
Metastasis (M)	
Yes	6 (20.7)
No	23 (79.3)
Family history	
Yes	7 (24.1)
No	22 (75.9)

---