Electronic Supplementary Material (ESI) for Molecular BioSystems.

Affinity proteomics led identification of vimentin as a potential biomarker in Colon cancers: Insights from serological screening and computational modelling-Supplementary data

Supplementary Data

1. Materials and Methods

1.1 LDH assay:

Cell viability was evaluated by measuring the viability of various cell lines in different media preparation via total cytoplasmic lactate dehydrogenase, using the commercially available LDH assay kit (Gbioscience) in accordance with the manufacturer's instructions and as described by Weyermann [23]. OD at 450 nm was recorded and taken as a measure for the quantity of cells.

1.2 SDS-PAGE:

SDS-PAGE was carried out using the Laemmli buffer system under reducing conditions. The resolving gel was polymerized in a mini gel apparatus (Hoefer) for 30-45 min. The thickness of the gel was 1.5mm. The stacking gel prepared afresh was layered on top of the resolving gel and allowed to polymerize for 15-20 min. In a final set of experiments, 30-50 µg of samples (total protein extract or secretory fraction) were mixed with Laemmli sample buffer and placed in a heating block at 100°C for 5 min before loading into wells. Electrophoresis was carried out at a constant current of 30mA.

1.3 Western Blot:

Western blot was carried out by the method described by [24]. The sample to be analyzed separated in a 10% SDS-PAGE under reducing conditions and transferred onto a nitrocellulose (NC) membrane (Millipore, USA) at a constant current of 300mA for 2 h using a Bio-Rad transfer apparatus (BioRad, USA). The membrane was blocked (3% BSA, 0.5% Tween20, TBS; 0.1 M Tris-HCl, 0.5 M NaCl) for 3 h at room temperature and subsequently probed with primary antibody (diluted 1:300 in blocking buffer) and washed thrice in TBST followed by IR labeled secondary (1:30000 dilution in blocking buffer) 700 IR-anti-rabbit antibody (Licor Odyssey). Detection was carried out using Licor Infra-Red imager according to the manufacturer's protocol.

1.4 Selection of immunocompetent patient sera:

For a selection of the most immune-competent sera from colorectal patients, ELISA was performed according to the manual provided by Licor Odyssey. Briefly, normal and patient sera serving as a source of TAA, were diluted in coating buffer (Carbonate –Bicarbonate Buffer 50mM-pH 9.5) to a final concentration of 10µg/100µl and kept for coating overnight at 4 ºC. Free binding sites were blocked with blocking solution (3% BSA in PBS). Anti-secretome antibody (anti-Sc) (1:200) was incubated with the sera-coated wells at room temperature for 2 hr, then washed three times with PBST (PBS + 0.05% Tween 20) followed by incubation with secondary antibody IRdye 800 anti- rabbit IgG (licor odyssey) at a dilution of 1:20000 as recommended by manufacturer for 30 minutes at 37ºC in a dark. After five washings with PBST, plates were scanned and analyzed using an OdysseyH IR scanner using OdysseyH imaging software 3.0. Scan settings were set to medium image quality, 169 mm resolution, intensity 5.0 for the 700-channel, and 800-channel with an offset of 4.0 mm. For signal quantification, antibody signals were analyzed as the average 800-channel integrated intensities from duplicate

wells normalized to the 700-channel signal integrated intensity to correct for well-to-well variations. Results are expressed as percent relative response of patient sera (means ± standard errors of the mean) compared to control sera.

1.5 Vimentin expression in patient sera.

Peripheral venous blood samples were collected; serum was isolated and stored at -80 ^{O}C until used. The serum Vimentin concentrations were determined by indirect ELISA. Briefly the serum samples were diluted at 1:100 in carbonate and bicarbonate buffer pH 9.6 and 100 µL of diluted sera were added to each well. The plate was incubated at 4 ^{O}C overnight for serum antigen coating. The wells were washed three times with PBST. Serum antigen coated wells were blocked using 2% BSA. 100 µL blocking solution (PBST-BSA) was added to each well, incubated at 37 ^{O}C for 1 hr and washed three times with PBST. The plate was then added with 100 µL of 1:2000 dilution of anti-vimentin antibody diluted in 1%BSA- PBS. After 1 hr incubation at 30 ^{O}C, the wells were washed three times with PBST and 100 µL of secondary antibody (IRDye anti-rabbit IgG diluted to 1:30000) was added to each well. After 30 mins of incubation at 30 ^{O}C, well were washed thrice with PBST and absorbance was measured at a wavelength of 700nm in the Odyssey Infra Red imager.

2. Supplementary Figures

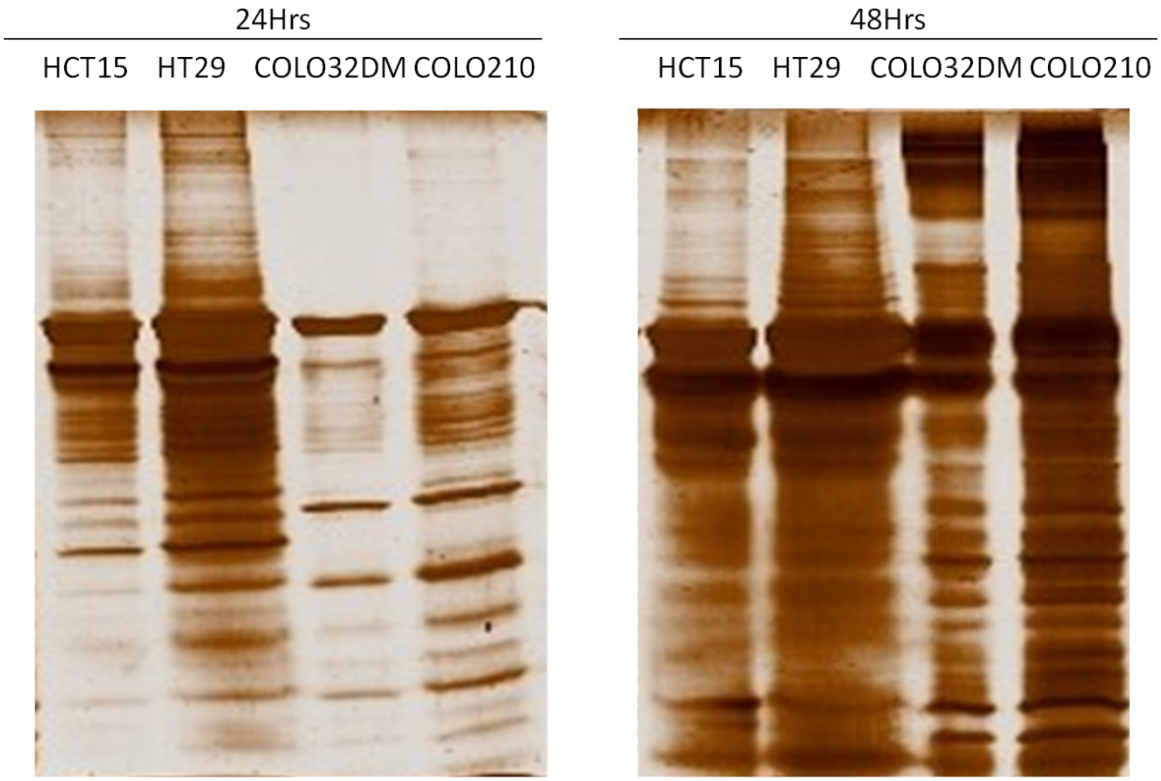


Fig S1: Secretome profiling of cell lines: Secretory fractions collected at 24hrs and 48hrs from human colon adenocarcinoma representative cell lines like HCT15, HT29, COLO320DM, COLO210 were concentrated and resolved on 10% SDS-PAGE and silver stained.

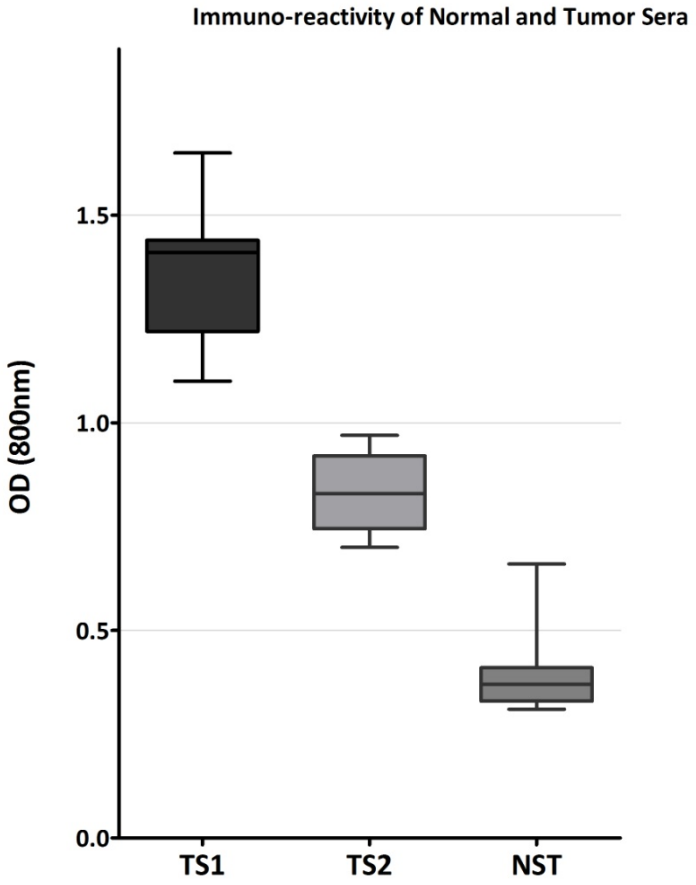
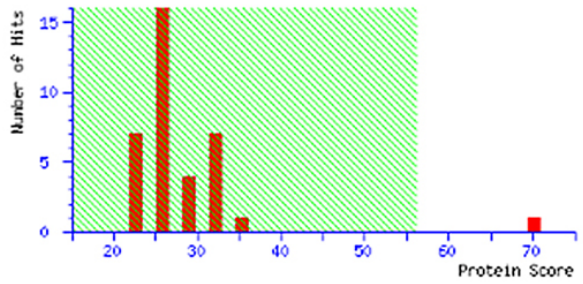


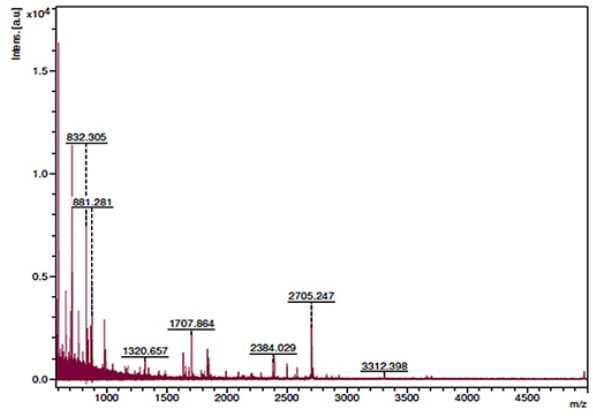
Fig S2: Graph represents the immune- reactivity of sera samples obtained by probing with affinity reagent (anti-Sc antibody.). TS1 comprises of sera samples that showed the highest immunoreactivity, TS2 group had moderately immunoreactive sera, all of them being tumor sera while as NST group comprised of those sera samples that showed the lowest immuno-reactivity. NST group had mixed cohort of tumor sera and normal sera. For signal quantification, antibody signals were analyzed as the average 800-channel integrated intensities from duplicate wells normalized to the 700-channel signal integrated intensity to correct for well-to-well variations Data represented as Mean±SD of atleast three independent experiments.

Database : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Taxonomy : Homo sapiens (human) (20271 sequences)
Timestamp : 25 Feb 2014 at 19:10:35 GMT
Top Score : 70 for PP16B_HUMAN, Protein phosphatase 1 regulatory inhibitor subunit 16B

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 56 are significant (p<0.05).





Protein sequence coverage: 11%

Matched peptides shown in ***bold red***.

```
   1 MASHVDLLTE LQLLEKVPTL ERLRAAQKRR AQQLKKWAQY EQDLQHRKRK
  51 HERKRSIGGR RKKVSFEASV ALLEASLRND AEEVRYFLKN KVSPDLCNED
 101 GLTALHQCCI DNFEEIVKLL LSHGANVNAR DNELWTPLHA AATCGHINLV
 151 KILVQYGADL LAVNSDGNMP YDLCEDEPTL DVIETCMAYQ GITQEKINEM
 201 RVAPEQQMIA DIHCMIAAGQ DLDWIDAQGA TLLHIAGANG YLRAAELLLD
 251 HGVRVDVKDW DGWEPLHAAA FWGQMQMAEL LVSHGASLSA RTSMDEMPID
 301 LCEEEEFKVL LLELKHKHDV IMKSQLRHKS SLSRRTSSAG SRGKVVRRAS
 351 LSDRTNLYRK EYEGEAILWQ RSAAEDQRTS TYNGDIRETR TDQENKDPNP
 401 RLEKPVLLSE FPTKIPRGEL DMPVENGLRA PVSAYQYALA NGDVWKVHEV
 451 PDYSMAYGNP GVADATPPWS SYKEQSPQTL LELKRQRAAA KLLSHPFLST
 501 HLGSSMARTG ESSSEGKAPL IGGRTSPYSS NGTSVYYTVT SGDPPLLKFK
 551 APIEEMEEKV HGCCRIS
```

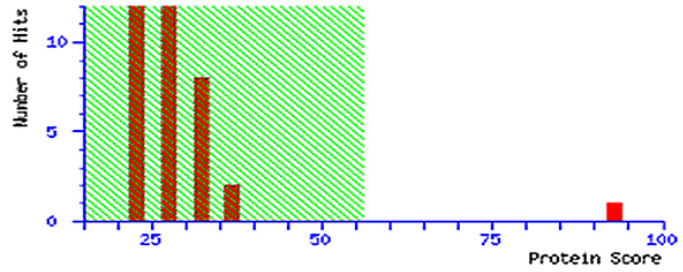
Unformatted sequence string: 567 residues (for pasting into other applications)

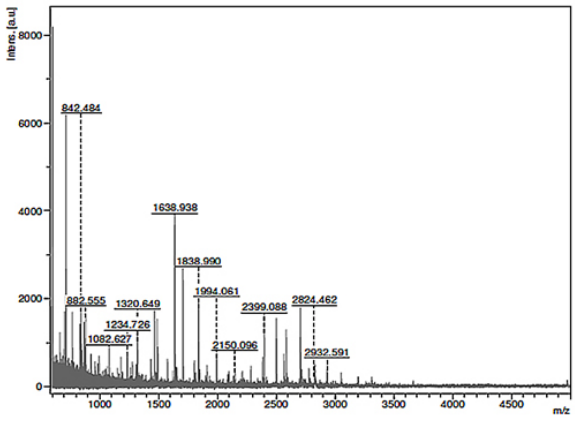
Fig S3: Identification of PP16B as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS. Graph represents MS spectrum of tryptic digests of PP16B.

Database : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Taxonomy : Homo sapiens (human) (20271 sequences)
Timestamp : 25 Feb 2014 at 21:19:08 GMT
Top Score : 93 for VIME_HUMAN, Vimentin OS=Homo sapiens GN=VIM PE=1 SV=4

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 56 are significant (p<0.05).





Protein sequence coverage: 28%

Matched peptides shown in *bold red*.

```
  1 MSTRSVSSSS YRRMFGGPGT ASRPSSSRSY VTTSTRTYSL GSALRPSTSR
 51 SLYASSPGGV YATRSSAVRL RSSVPGVRLL QDSVDFSLAD AINTEFKNTR
101 TNEKVELQEL NDRFANYIDK VRFLEQQNKI LLAELEQLKG QGKSRLGDLY
151 EEEMRELRRQ VDQLTNDKAR VEVERDNLAE DIMRLREKLQ EEMLQREEAE
201 NTLQSFRQDV DNASLARLDL ERKVESLQEE IAFLKKLHEE EIQELQAQIQ
251 EQHVQIDVDV SKPDLTAALR DVRQQYESVA AKNLQEAEEW YKSKFADLSE
301 AANRNNDALR QAKQESTEYR RQVQSLTCEV DALKGTNESL ERQMREMEEN
351 FAVEAANYQD TIGRLQDEIQ NMKEEMARHL REYQDLLNVK MALDIEIATY
401 RKLLEGEESR ISLPLPNFSS LNLRETNLDS LPLVDTHSKR TLLIKTVETR
451 DGQVINETSQ HHDDLE
```

Unformatted sequence string: 466 residues (for pasting into other applications).

Fig S4: Identification of Vimentin as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity-2DE-MS. Graph represents theMS spectrum of tryptic digests of Vimentin.

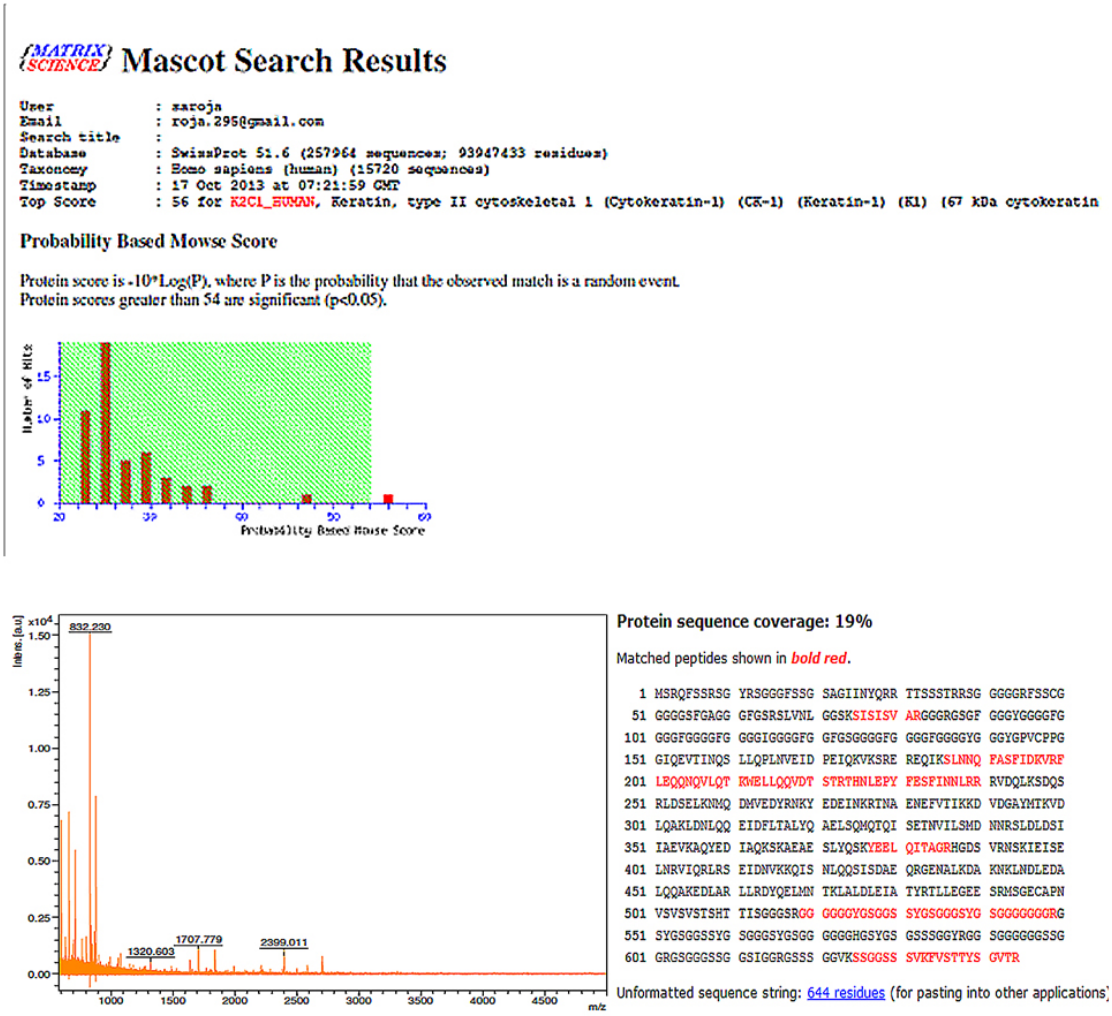


Fig S5: Identification of KRT1 as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS. Graph represents the **MS** spectrum of tryptic digests of KRT1

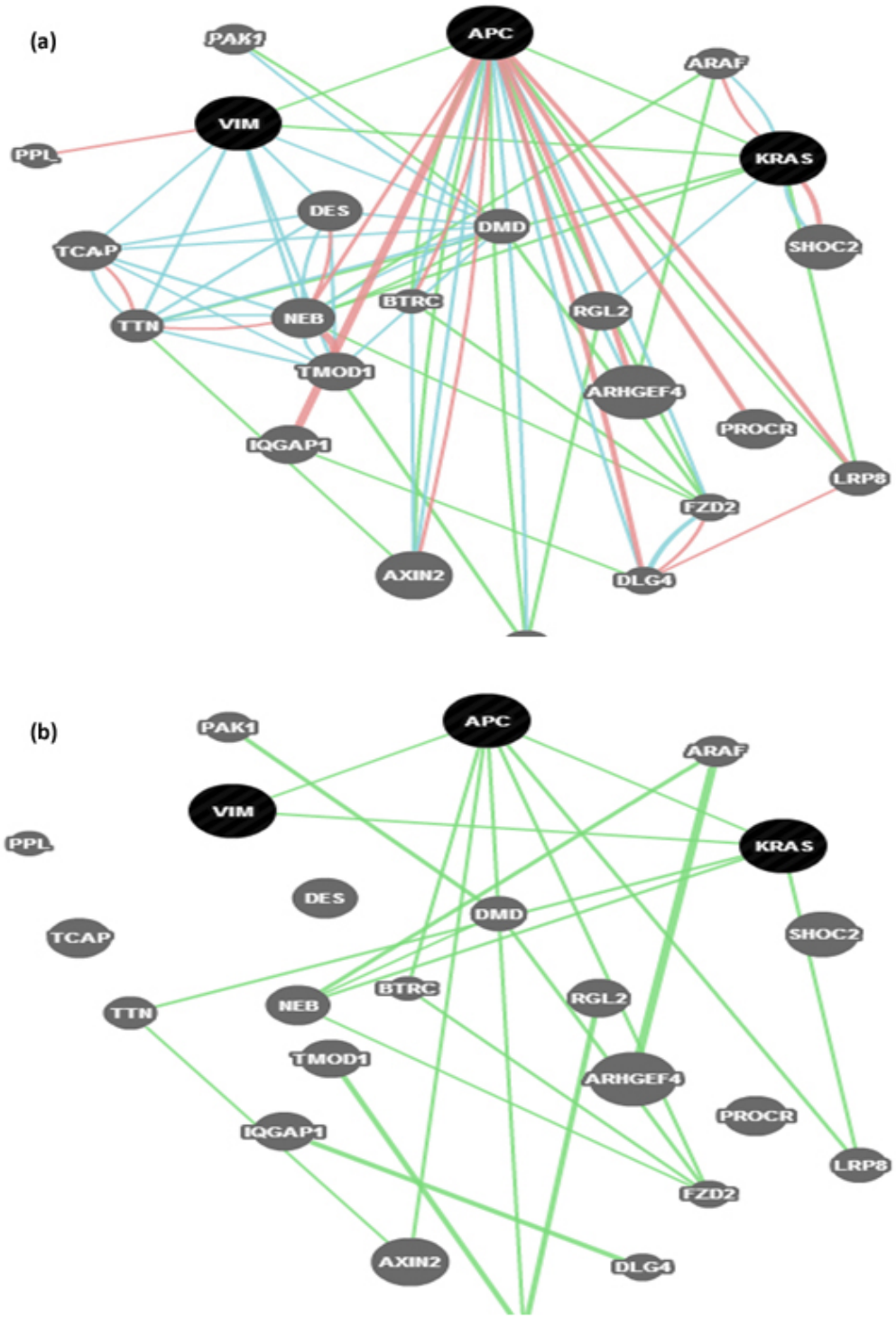


Fig S6.1 Protein subnetworks of Vimentin(Query protein) with KRAS and APC (Hubs): Protein-Protein Physical/Functional (a) and Genetic interaction (b) sub networks generated by Cytoscape demonstrate the secondary physical interaction of Vimentin with APC via NEB, DMD and GSK3B while as with KRAS via ARAF. Direct genetic interaction also evident between the query protein, Vimentin and Hubs genetrated via KRAS and APC. Black filled circles are the query proteins and Hubs. Only significant sub networks are shown in the figure

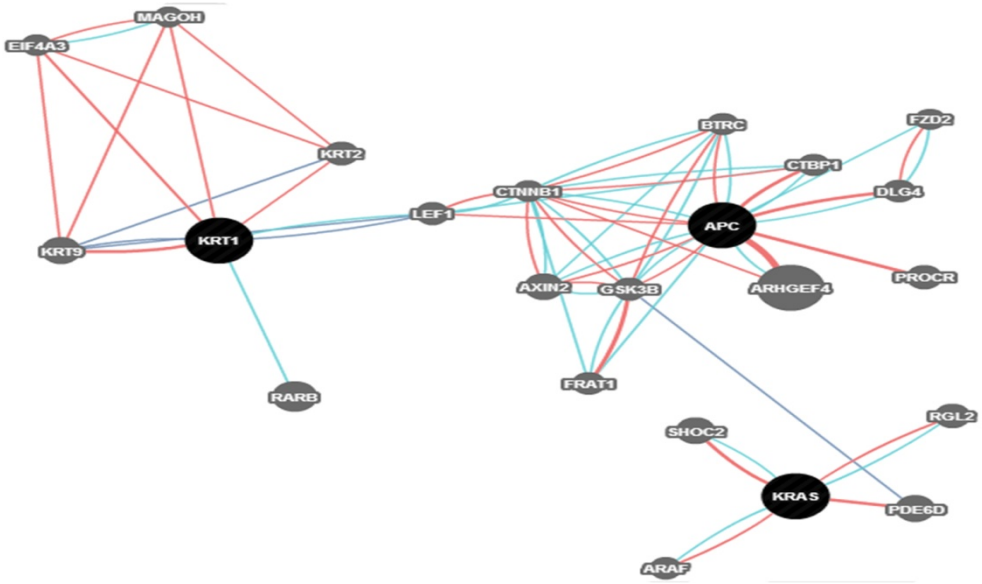


Fig S6.2 Protein subnetworks of KRT1 (Query protein) with KRAS and APC (Hubs): Protein-Protein Physical/Functional interaction between the sub networks generated by Cytoscape demonstrate that KRT1 shows a distant dependency via LEF1 on APC and K-ras mediated pathways.No genetic interaction was found in this network. Black filled circles are the query proteins and Hubs. Only significant sub networks are shown in the figure.

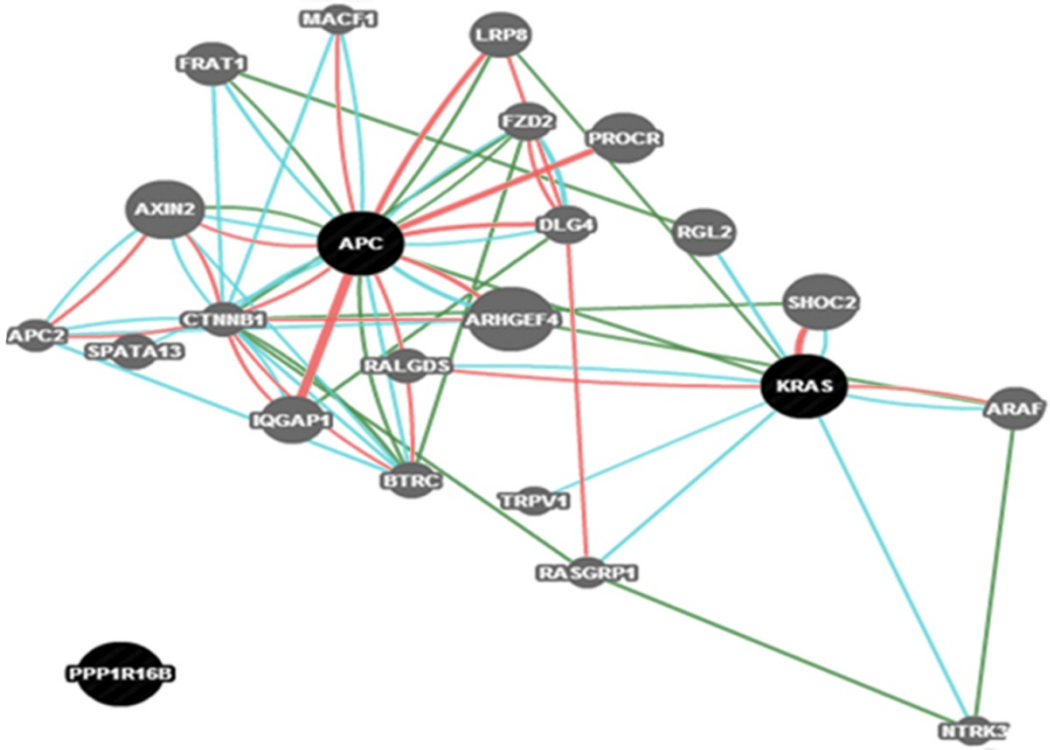


Fig S6.3 Protein subnetworks of PP16B denoted as PPP1R16B (Query protein) with KRAS and APC (Hubs): Protein-Protein Physical/Functional interaction between the sub networks demonstrate that PPP1R16B lies outside the interactive network generated by Cytoscape and therefore does not show any dependency either through the physical or genetic interaction. Black filled circles are the query proteins and Hubs. Only significant sub networks are shown in the figure.

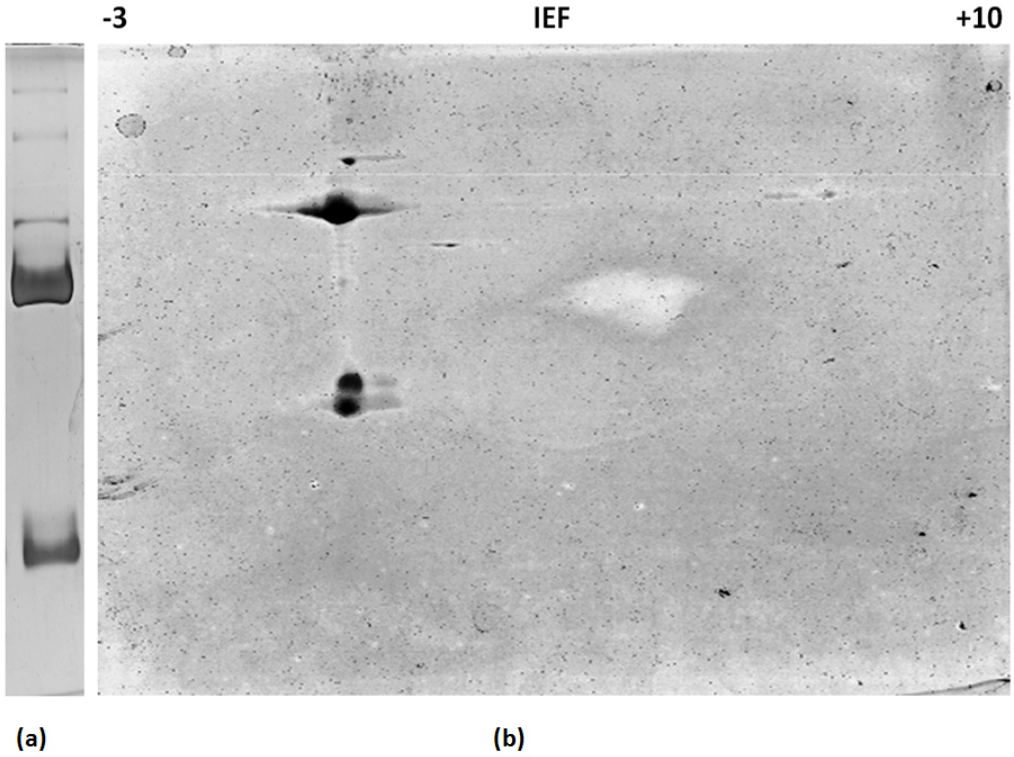


Fig S 7: a) Representative 10% SDS Gel of purified polyclonal anti-Sc antibody (affinity reagent). b) 2-Dimensional Electrophoresis of Anti-Sc antibody (affinity reagent)

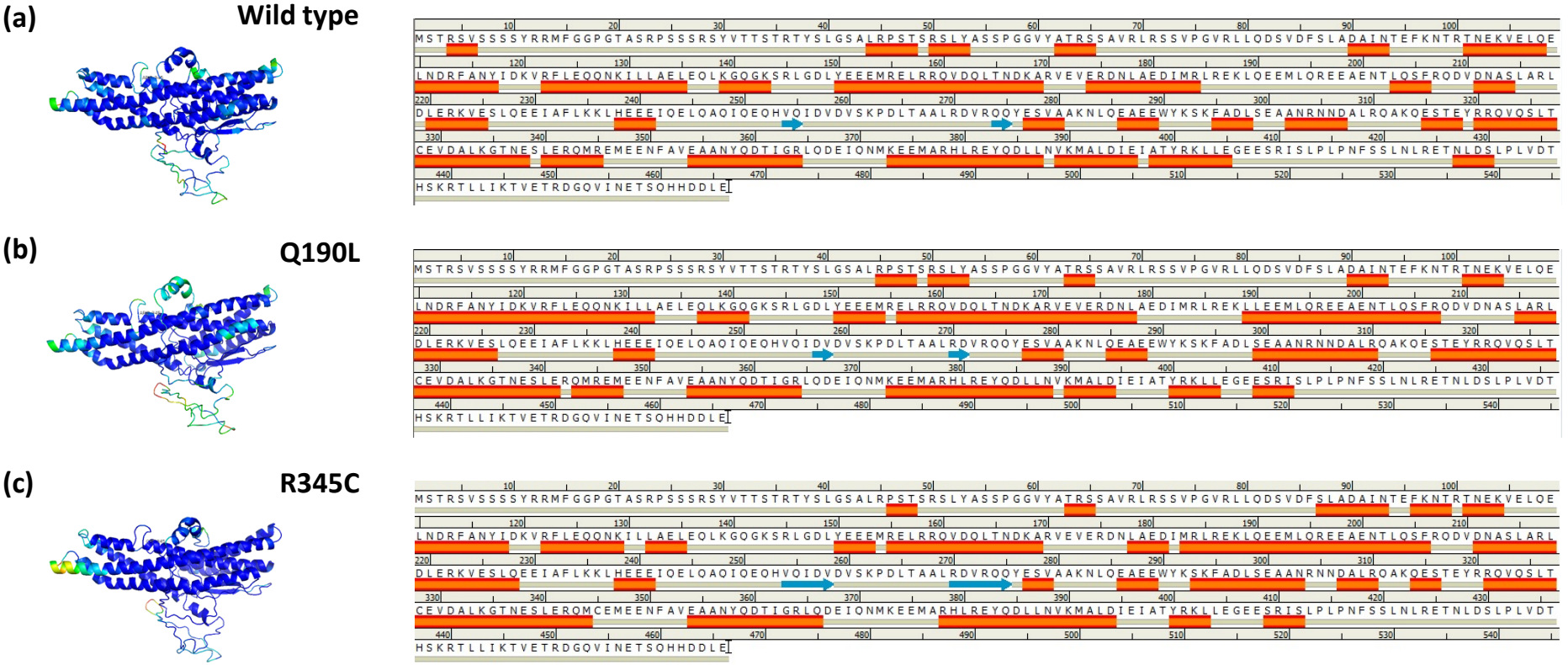


Fig S8: The average bfactor values implicited on the tertiary structure of (a) Wildtype, (b) Q190L and (c) R345C vimentin structure with their corresponding average secondary structure architecture, arrows represent the major variations.

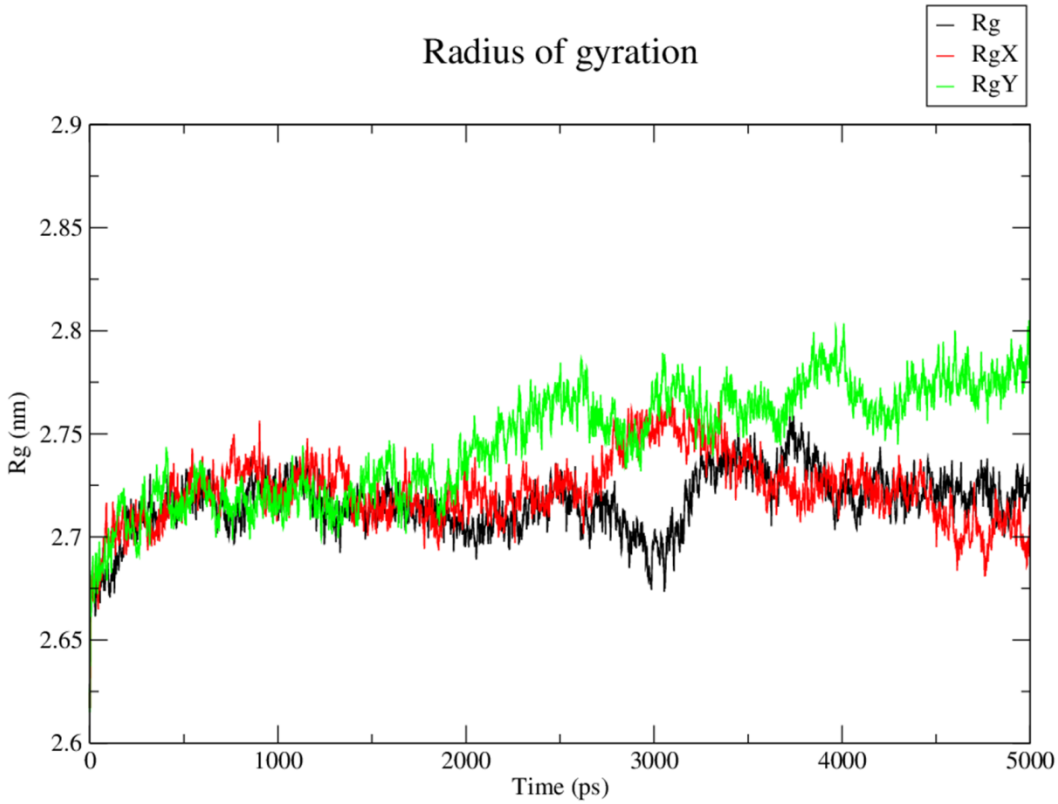


Fig S9: Variation in Rg of three structures at 300k (Black: WT, Red: Q190L, Green: R345C).

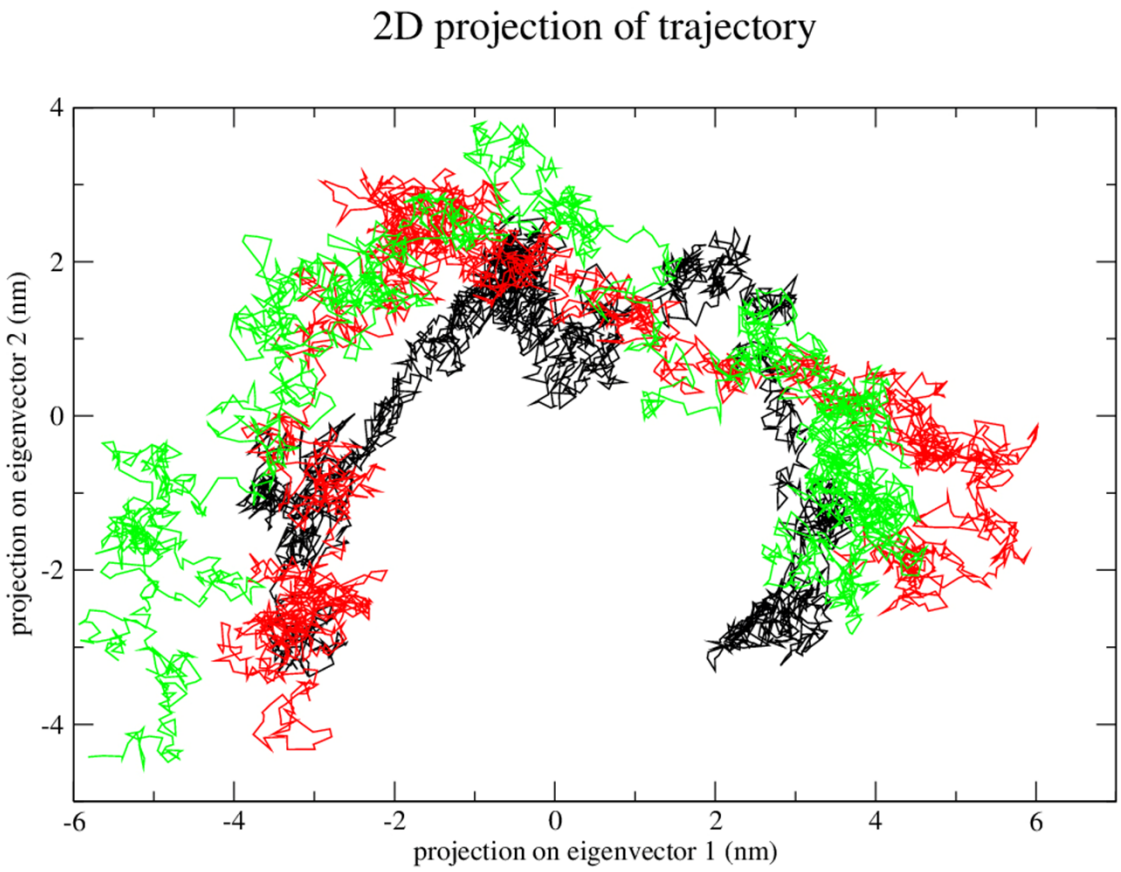
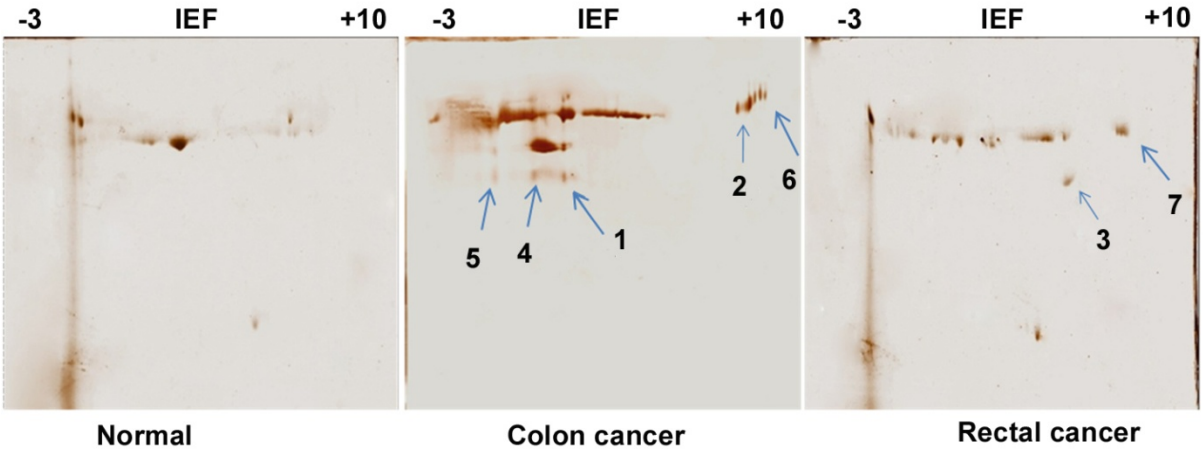


Fig S10: The cluster obtained from wt structure is stable (black) where as Mu –Q190L (red) is slightly expanded compared to wt. However mu-R345C (green) covers a large area and therefore is considered to be highly unstable.



Spot ID	Protein Name	Score	Remarks
1	Vimentin	93	Confident
2	KRT1	64	Confident
3	PP16B	70	Confident
4	ERG1	40	Not Significant
5	ZN619	58	Not Significant
6	NKG21	64	PI/MW mismatch
7	LIN37	44	Not Significant

Fig S11: Two-Dimensional gels display differential protein spots precipitated after incubation of pooled sera with anti-Sc antibody (Affinity Reagent).

MATRIX SCIENCE **Mascot Search Results**

```
User            : uok
Email           : shoiabb@gmail.com
Search title    :
Database        : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Timestamp       : 26 Feb 2014 at 19:31:11 GMT
Top Score       : 58 for ZN619_HUMAN, Zinc finger protein 619 OS=Homo sapiens GN=ZNF619 PE=2 SV=1
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 70 are significant (p<0.05).

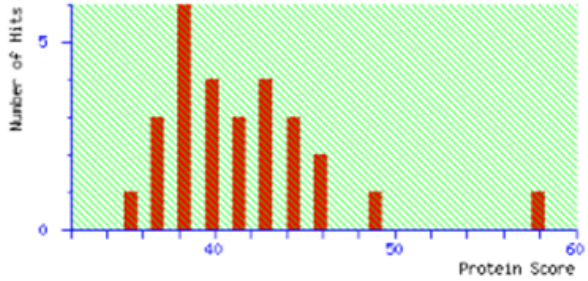


Fig S12a: Identification of ZN619 as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS.

MATRIX SCIENCE **Mascot Search Results**

```
User            : uok
Email           : shoiabb@gmail.com
Search title    :
Database        : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Taxonomy        : Homo sapiens (human) (20271 sequences)
Timestamp       : 25 Feb 2014 at 18:24:28 GMT
Top Score       : 40 for ERGI1_HUMAN, Endoplasmic reticulum-Golgi intermediate compartment protein 1 OS=Homo sapiens GN=ERGIC1 PE=1 SV=1
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 56 are significant (p<0.05).

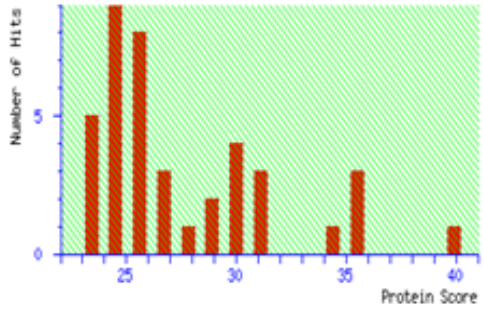


Fig S12b: Identification of ERGI1 as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS.

MATRIX SCIENCE Mascot Search Results

```
User            : uok
Email           : shoiabb@gmail.com
Search title    :
Database        : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Taxonomy        : Homo sapiens (human) (20271 sequences)
Timestamp       : 25 Feb 2014 at 18:29:07 GMT
Top Score       : 44 for LIN37_HUMAN, Protein lin-37 homolog OS=Homo sapiens GN=LIN37 PE=1 SV=1
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 56 are significant (p<0.05).

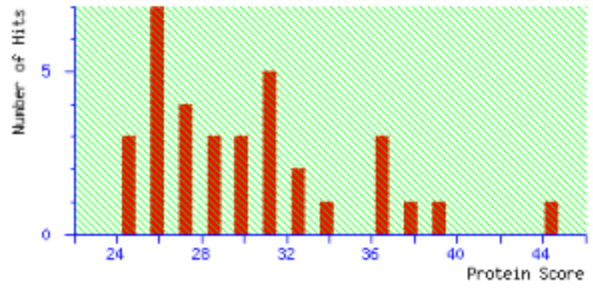


Fig S12c: Identification of LIN37 as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS.

MATRIX SCIENCE Mascot Search Results

```
User            : uok
Email           : shoiabb@gmail.com
Search title    :
Database        : SwissProt 2014_02 (542503 sequences; 192888369 residues)
Taxonomy        : Homo sapiens (human) (20271 sequences)
Timestamp       : 25 Feb 2014 at 19:06:49 GMT
Top Score       : 64 for NKG2A_HUMAN, NKG2-A/NKG2-B type II integral membrane protein OS=Homo sapiens GN=KLRC1 PE=1 SV=2
```

Mascot Score Histogram

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 56 are significant (p<0.05).

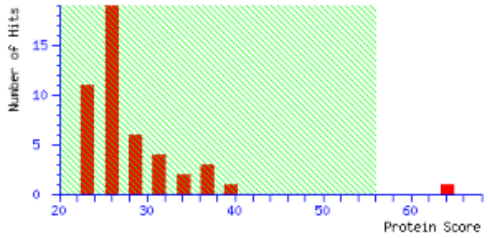


Fig S12d: Identification of NKG2A as an immunogenic cancer biomarker from pooled sera of rectal patients using combined strategy of Affinity- 2-DE - MS.

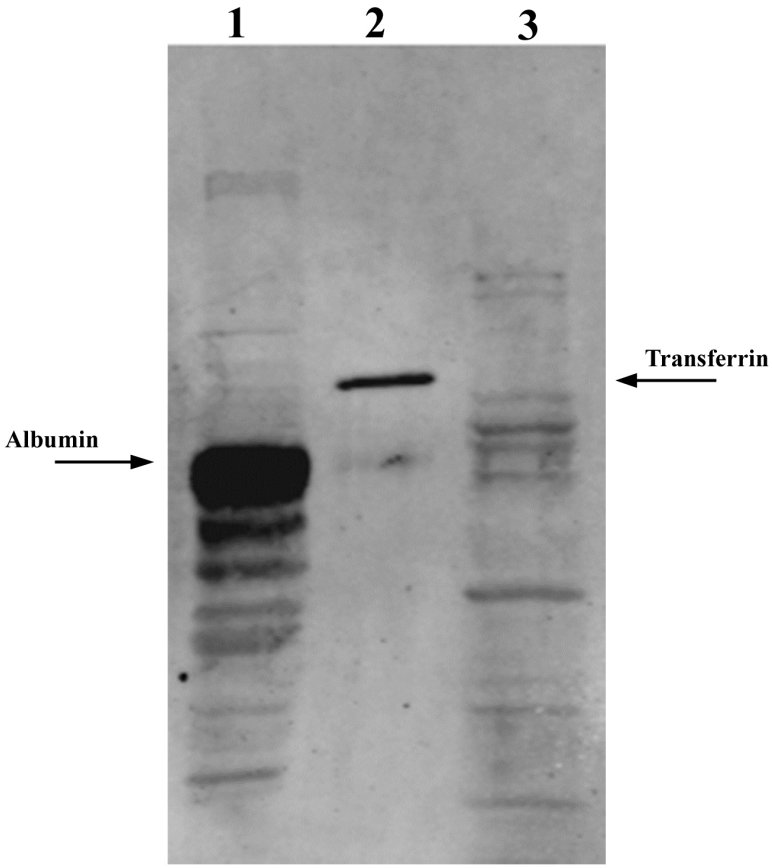


Fig: S13 Western blot analysis for evaluating the affinity reagent specificity:(1) Precipitated HT29 secretory fraction (2) Transferrin supplemented media and (3) precipitated HT29 secretory fraction precleared from abundant proteins were probed with anti-secretome antisera at 1:200 dilution. The blot clearly indicates that the abundant proteins like Transferrin, Albumin etc were removed from pre-cleared HT29 secretory fraction.