PROTEOMICS

Supporting Information for Proteomics

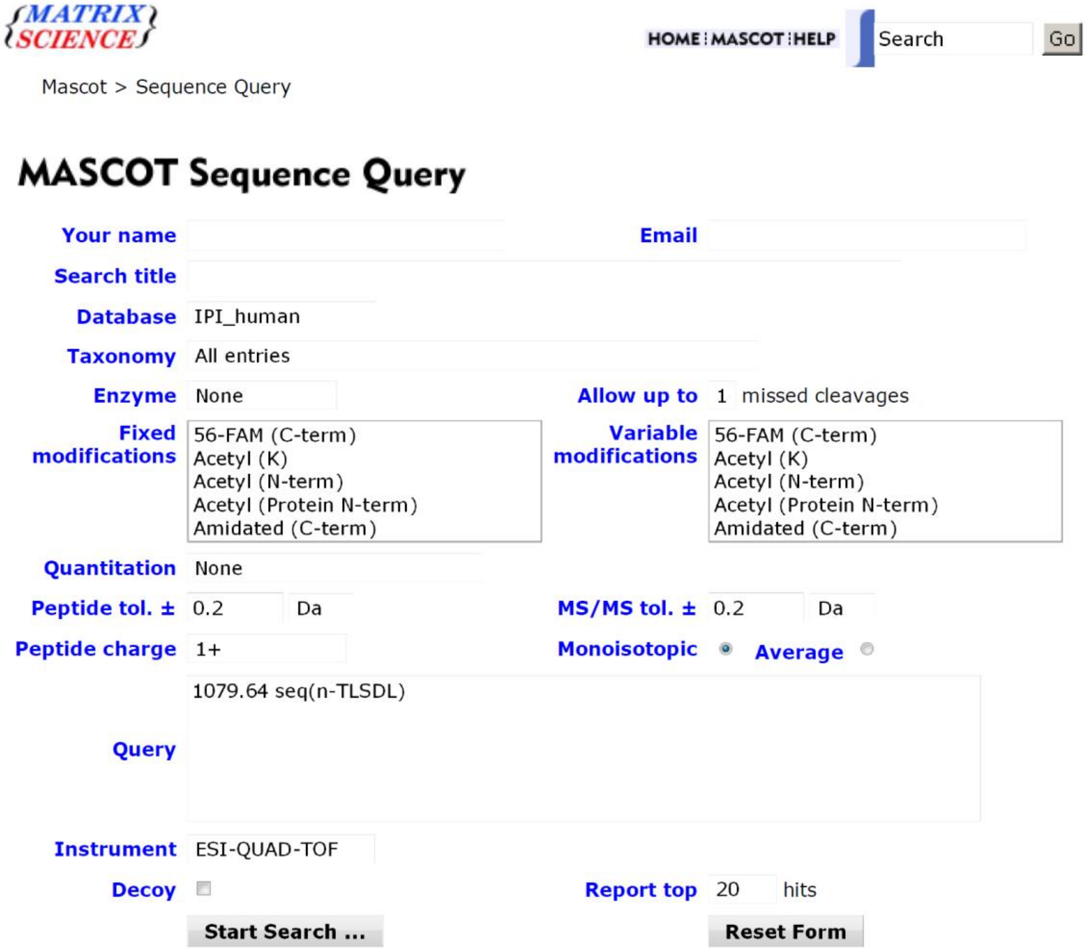
DOI 10.1002/pmic.201000486

Barbara Seliger, Sven P. Dressler, Chiara Massa, Christian V. Recktenwald, Florian Altenberend, Juergen Bukur, Francesco M. Marincola, Ena Wang, Stefan Stevanovic and Rudolf Lichtenfels

Identification and characterization of human leukocyte antigen class I ligands in renal cell carcinoma cells

peptide	RT [min]	Rtcalc.	(M+H)$^+$ calc.	(M+2H)2$^+$ calc.	(M+2H)2$^+$ obs.	∆M
VPATDRNAL	23.1	20.2	956.5084	478.7542	478.77	0.02
ALADGVQKV	27.8	13.9	900.5073	450.7537	450.74	0.01
VEVGRVAYV	29.5	15.3	991.5497	496.2749	496.26	0.01
NTDSPLRY	31.6	21.6	965.4611	483.2306	483.24	0.01
KEFDGKSLV	31.6	12.8	1022.544	511.7721	511.77	0.00
RPELVRPAL	39.3	29.2	1050.634	525.8172	525.79	0.03
RVMAPRALL	45.7	35.7	1026.617	513.8083	513.77	0.04
RPQKVCSFL	46.5	25.6	1077.58	539.2899	539.31	0.02
TEITHAVVI	49.0	32.9	982.5492	491.7746	491.75	0.02
AGLPRPDLSL	51.3	40.9	1038.587	519.7933	519.78	0.01
TAHLMVVVL	53.7	45.7	982.5679	491.7840	491.75	0.03
LTYEALGLCP	54.7	43.9	1079.537	540.2683	540.29	0.00
FPNIPGKSL	59.4	49.2	972.5436	486.7718	468.75	0.02
GEHTLLVTV	63.0	28.9	968.5335	484.7668	484.77	0.00
LPKQPPLAL	64.1	54.7	976.6112	488.8056	488.80	0.01
REAPSPLMI	66.5	35.6	1013.537	507.2687	507.25	0.02
GEYPKLLRL	68.6	38.3	1088.639	544.8193	544.83	0.01
SIFEVVWAI	69.6	57.4	1063.575	532.2874	532.29	0.01
MPRGVVVTL	69.7	38.0	971.5631	486.2816	486.28	0.00
LVLMVLYLI	70.6	80.2	1076.671	538.8356	538.79	0.05
IYAHGALPII	70.7	68.2	1067.617	534.3086	534.31	0.00
APRAVFPSI	71.9	43.1	957.5441	479.2721	479.26	0.01
REANLQALI	73.2	22.0	1027.582	514.2910	514.25	0.04
RIGSGLKALL	75.3	35.8	1027.655	514.3273	514.28	0.05
APRPGLLSL	76.3	43.7	923.5596	462.2798	462.30	0.02
RESFSLVQV	76.3	15.1	1064.566	532.7830	532.81	0.03
STDHIPILY	80.2	52.1	1058.544	529.7720	529.77	0.00
IPRSITVLV	82.6	48.0	997.6328	499.3164	499.31	0.01
IPRAALLPL	94.6	59.6	963.6273	482.3137	482.31	0.00
RPTLWAAAL	95.6	49.3	998.5706	499.7853	499.79	0.00
ISVGISLLLL	101.3	65.5	1027.668	514.3342	514.31	0.02
IMLEALERV	103.0	29.1	1073.595	537.2975	537.31	0.01
LEVIPRTLI	104.4	50.9	1053.659	527.3295	527.33	0.00
GEFIIGRVI	107.2	46.1	1003.586	502.2930	502.29	0.00
REFLFNAI	113.8	42.4	1009.539	505.2696	505.28	0.01
TLSDLRVYL	114.7	38.5	1079.602	540.3010	540.32	0.02
IEIERILSV	114.7	32.0	1071.633	536.3167	536.33	0.02
REMIPFAVV	114.7	41.8	1061.574	531.2869	531.29	0.00
LLMVLSPRL	122.2	56.7	1041.641	521.3207	521.34	0.02
PSRDSLPLPV	123.1	39.1	1080.597	540.7986	540.84	0.04
VLIPKLPQL	124.3	61.9	1020.674	510.8369	510.84	0.00
GPRAVFVLL	125.5	49.9	971.5961	486.2981	486.28	0.02
LSLENLEKI	127.6	24.8	1058.602	529.8008	529.82	0.02
VMLDVPIRL	129.5	52.7	1055.621	528.3103	528.30	0.01
AEFIKFTVI	129.6	51.9	1067.606	534.3030	534.31	0.01
APRTLVLLL	130.8	44.7	995.6535	498.3268	498.34	0.01
RESFPMILV	134.6	44.9	1091.584	546.2922	546.31	0.02
RLLEVPVML	135.8	46.2	1069.636	535.3182	535.35	0.03
RLLDVLAPL	138.1	50.4	1009.633	505.3164	505.34	0.02

1st example, TLSDLRVYL



MASCOT Sequence Query

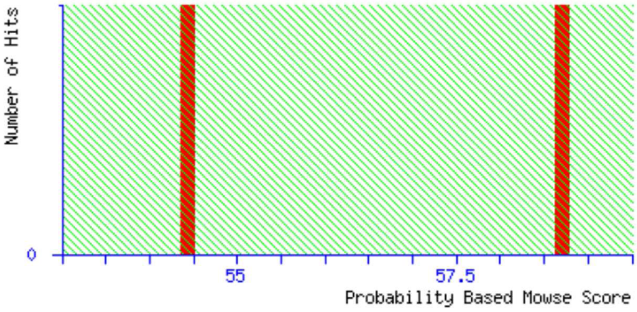
Your name		Email	
Search title			
Database	IPI_human		
Taxonomy	All entries		
Enzyme	None	Allow up to	1 missed cleavages
Fixed modifications	56-FAM (C-term) Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term)	Variable modifications	56-FAM (C-term) Acetyl (K) Acetyl (N-term) Acetyl (Protein N-term) Amidated (C-term)
Quantitation	None		
Peptide tol. ±	0.2 Da	MS/MS tol. ±	0.2 Da
Peptide charge	1+	Monoisotopic ●	Average ○
Query	1079.64 seq(n-TLSDL)		
Instrument	ESI-QUAD-TOF		
Decoy	☐	Report top	20 hits

Mascot Search Results

```
User             :
Email            :
Search title     :
Database         : IPI_human  (86635 sequences; 35084983 residues)
Timestamp        : 3 Mar 2011 at 12:06:45 GMT
Top Score        : 59 for IPI00168554, Tax_Id=9606 Gene_Symbol=SRXN1 Sulfiredoxin-1
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 62 are significant (p<0.05).



Protein Summary Report

Format As | Protein Summary | Help

Significance threshold p< 0.05 | Max. number of hits 20

Re-Search All | Search Unmatched

Index

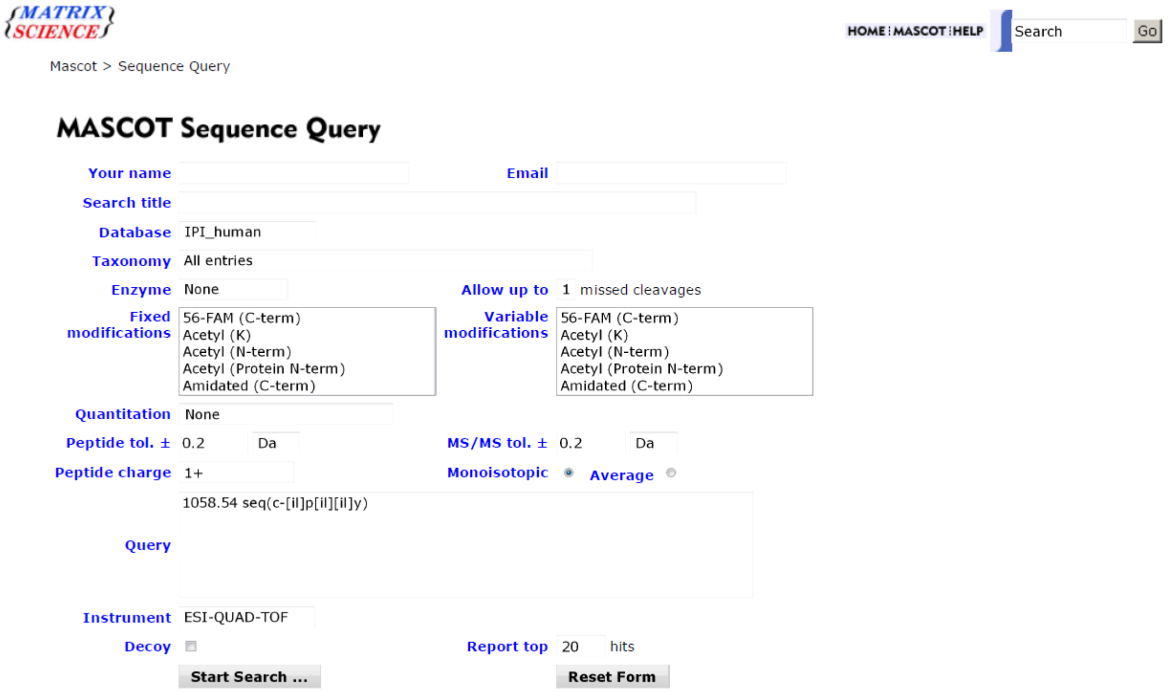
```
   Accession     Mass   Score  Description
1. IPI00168554   14250     59  Tax_Id=9606 Gene_Symbol=SRXN1 Sulfiredoxin-1
2. IPI00646354   39902     54  Tax_Id=9606 Gene_Symbol=PTAR1 40 kDa protein
```

Results List

```
1.     IPI00168554     Mass: 14250     Score: 59     Expect: 0.12  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=SRXN1 Sulfiredoxin-1
        Observed   Mr(expt)   Mr(calc)   Delta Start   End Miss Ions   Peptide
      1079.6400  1078.6327  1078.6022  0.0305   121 - 129    0  ---    S.TLSDLRVYL.G

2.     IPI00646354     Mass: 39902     Score: 54     Expect: 0.31  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=PTAR1 40 kDa protein
        Observed   Mr(expt)   Mr(calc)   Delta Start   End Miss Ions   Peptide
      1079.6400  1078.6327  1078.6022  0.0305   267 - 276    0  ---    G.TLSDLHLIPA.G
```

2nd example, STDHIPILY



Mascot > Sequence Query

MASCOT Sequence Query

Your name

Email

Search title

Database IPI_human

Taxonomy All entries

Enzyme None

Allow up to 1 missed cleavages

Fixed modifications
56-FAM (C-term)
Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)

Variable modifications
56-FAM (C-term)
Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (C-term)

Quantitation None

Peptide tol. ± 0.2 Da

MS/MS tol. ± 0.2 Da

Peptide charge 1+

Monoisotopic ● **Average** ○

Query 1058.54 seq(c-[il]p[il][il]y)

Instrument ESI-QUAD-TOF

Decoy ☐

Report top 20 hits

Mascot Search Results

```
User          :
Email         :
Search title  :
Database      : IPI_human  (86635 sequences; 35084983 residues)
Timestamp     : 3 Mar 2011 at 12:11:21 GMT
Top Score     : 47 for IPI00299506, Tax_Id=9606 Gene_Symbol=GFPT1 Isoform 2 of Glucosamine--fructose-6-phosphate aminotra
```

Probability Based Mowse Score

Protein score is -10*Log(P), where P is the probability that the observed match is a random event.
Protein scores greater than 62 are significant (p<0.05).

Protein Summary Report

Significance threshold p< 0.05

Max. number of hits 20

Index

```
    Accession     Mass    Score  Description
 1. IPI00299506   76710     47   Tax_Id=9606 Gene_Symbol=GFPT1 Isoform 2 of Glucosamine--fructose-6-phosphate aminotransfera
 2. IPI00217952   78756     47   Tax_Id=9606 Gene_Symbol=GFPT1 Isoform 1 of Glucosamine--fructose-6-phosphate aminotransfera
 3. IPI00953492   135891    44   Tax_Id=9606 Gene_Symbol=BAT2L2 Protein
 4. IPI00902970   291012    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 2 of Protein BAT2-like 2
 5. IPI00941970   295881    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 3 of Protein BAT2-like 2
 6. IPI00946618   301388    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Uncharacterized protein
 7. IPI00148247   301485    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 6 of Protein BAT2-like 2
 8. IPI00902809   312542    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 5 of Protein BAT2-like 2
 9. IPI00941219   312516    41   Tax_Id=9606 Gene_Symbol=BAT2L2 cDNA FLJ52084, highly similar to Homo sapiens BAT2 domain co
10. IPI00852708   316718    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 4 of Protein BAT2-like 2
11. IPI00902898   316815    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 1 of Protein BAT2-like 2
12. IPI00739117   316888    41   Tax_Id=9606 Gene_Symbol=BAT2L2 317 kDa protein
13. IPI00083708   316985    41   Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 7 of Protein BAT2-like 2
```

Results List

```
1.     IPI00299506     Mass: 76710     Score: 47      Expect: 1.7  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=GFPT1 Isoform 2 of Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 1
        Observed   Mr(expt)   Mr(calc)    Delta Start   End Miss Ions  Peptide
       1058.5400  1057.5327  1057.5444  -0.0117   218 - 226   0  ---   L.STDHIPILY.R

2.     IPI00217952     Mass: 78756     Score: 47      Expect: 1.8  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=GFPT1 Isoform 1 of Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 1
        Observed   Mr(expt)   Mr(calc)    Delta Start   End Miss Ions  Peptide
       1058.5400  1057.5327  1057.5444  -0.0117   218 - 226   0  ---   L.STDHIPILY.R

3.     IPI00953492     Mass: 135891    Score: 44      Expect: 3.4  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=BAT2L2 Protein
        Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
       1058.5400  1057.5327  1057.5808  -0.0481   998 - 1006   0  ---   S.NTQPIPILY.E

4.     IPI00902970     Mass: 291012    Score: 41      Expect: 6.9  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 2 of Protein BAT2-like 2
        Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
       1058.5400  1057.5327  1057.5808  -0.0481  2272 - 2280   0  ---   S.NTQPIPILY.E

5.     IPI00941970     Mass: 295881    Score: 41      Expect: 7  Queries matched: 1
       Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 3 of Protein BAT2-like 2
        Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
       1058.5400  1057.5327  1057.5808  -0.0481  2516 - 2524   0  ---   S.NTQPIPILY.E
```

```
6.   IPI00946618    Mass: 301388   Score: 41     Expect: 7  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Uncharacterized protein
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2450 - 2458   0  ---   S.NTQPIPILY.E


7.   IPI00148247    Mass: 301485   Score: 41     Expect: 7  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 6 of Protein BAT2-like 2
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2450 - 2458   0  ---   S.NTQPIPILY.E

8.   IPI00902809    Mass: 312542   Score: 41     Expect: 7.4  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 5 of Protein BAT2-like 2
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2468 - 2476   0  ---   S.NTQPIPILY.E

9.   IPI00941219    Mass: 312516   Score: 41     Expect: 7.4  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 cDNA FLJ52084, highly similar to Homo sapiens BAT2 domain containing 1 (BAT2D1), mRNA
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2469 - 2477   0  ---   S.NTQPIPILY.E

10.  IPI00852708    Mass: 316718   Score: 41     Expect: 7.5  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 4 of Protein BAT2-like 2
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2515 - 2523   0  ---   S.NTQPIPILY.E

11.  IPI00902898    Mass: 316815   Score: 41     Expect: 7.5  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 1 of Protein BAT2-like 2
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2515 - 2523   0  ---   S.NTQPIPILY.E

12.  IPI00739117    Mass: 316888   Score: 41     Expect: 7.5  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 317 kDa protein
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2517 - 2525   0  ---   S.NTQPIPILY.E

13.  IPI00083708    Mass: 316985   Score: 41     Expect: 7.5  Queries matched: 1
     Tax_Id=9606 Gene_Symbol=BAT2L2 Isoform 7 of Protein BAT2-like 2
      Observed   Mr(expt)   Mr(calc)    Delta Start    End Miss Ions  Peptide
     1058.5400  1057.5327  1057.5808  -0.0481  2517 - 2525   0  ---   S.NTQPIPILY.E
```

Search Parameters

```
Type of search         : Sequence Query
Enzyme                 : None
Mass values            : Monoisotopic
Protein Mass           : Unrestricted
Peptide Mass Tolerance : ± 0.2 Da
Fragment Mass Tolerance: ± 0.2 Da
Max Missed Cleavages   : 1
Instrument type        : ESI-QUAD-TOF
Query1 (1058.5400,1+)  : <no title>
                       : seq(C-[IL]P[IL][IL]Y)
```

Supplementary Table 1: List of independently defined HLA ligand source proteins, their HLA-restriction and chromosomal localization

protein	previously defined HLA ligands (SYFPEITHI DB)	chromosomal localisation
Sterol carrier protein 2	negative	1p32
Glutamine-fructose-6-phosphate transaminase 1	HLA-A*01 [72], HLA-A*2601; HLA-Cw*0402 [73]	2p13
Ribosomal protein SA	HLA-A*01	3p22.2
Apolipoprotein L, 1	HLA-A*0201 [13], HLA-B*08 [18]	22q13.1
Calcium channel, voltage- dependent, alpha 1E subunit	negative	1q25-q31
Clathrin, heavy chain	HLA-A*01, HLA-A*0201, HLA-B*44 [9], HLA-A*2601	17q11-qter
Collagen, type XVIII, alpha 1	HLA-A*0201 [9]	21q22.3
hypothetical LOC400614	negative	17q24.2
Chromosome 2 open reading frame 54	negative	2q37.3
LSM8 homolog, U6 small nuclear RNA associated	negative	19q13.42
ORM1-like 3 (S. cerevisiae)	HLA-A*0201 [9], HLA-B*5201 [74]	17q12-q21.1
Phosphatidylinositol glycan anchor biosynthesis, class M	HLA-B*07 [9]	1q23.2
RAS protein activator like 2	HLA-A*0201 [9]	1q24
Small nuclear ribonucleoprotein polypeptide G	HLA-A*0201 [9]	2p13.3
Sulfiredoxin 1	HLA-A*0201 [9]	20p13
Major histocompatibility complex, class I, C	HLA-DRB1*0404 (DR4Dw14) [75], HLA-DRB1*1302 [76], HLA-B*07 [9, 77], HLA-B*0702	6p21.3
Olfactory receptor, family 13, subfamily F, member 1	negative	9q31.1
Actin, alpha 1, skeletal muscle	HLA-A*1101, HLA-B*45 [9], HLA-B*0702 [78], HLA-B*4402 [52], HLA-DRB1*0404 (DR4Dw14) [79]	1q42.13-q42.2
Aldo-keto reductase family 1, member C3	HLA-B*07 [9, 25], HLA-B*0702 [24]	10p15-p14
Degenerative spermatocyte homolog 1, lipid desaturase	negative	1q42.12
Eukaryotic translation initiation factor 3, subunit 6 48kDa	HLA-Cw*0401 [73]	8q22-q23
G-protein signaling modulator 1 (AGS3-like, C. elegans)	HLA-B*07 [9]; HLA-A*07 [9], HLA-B*27	9q34.3
Guanine nucleotide binding protein (G protein), gamma 5 subunit	HLA-A*0201 [80], HLA-A*03, HLA-A*07 [9], HLA-B*1516 [53]	1p22
HECT domain containing 1	negative	14q12
Insulin-like growth factor binding protein 3	HLA-B*0707 [9]	7p13-p12
Major histocompatibility complex, class I, A	HLA-DRB1*0701 [81]	6p21.3
Minichromosome maintenance complex component 7	HLA-A*0201 [52]	7q21.3-q22.1
HtrA serine peptidase 1	HLA-B*07 [9]	10q26.3
RAN binding protein 9	HLA-A*25 [9], HLA-B*1801[82]	6p23
Transcription factor binding to IGHM enhancer 3	negative	Xp11.22
Widely interspaced zinc finger motifs	negative	19p13.1

GRB2-binding adaptor protein, transmembrane	negative	5q11.2
Inositol polyphosphate-5-phosphatase F	negative	10q26.11-q26.12
Bromodomain containing 4	negative	7q11.23
Chaperonin containing TCP1, subunit 3 γ	negative	19p13.1
Component of oligomeric golgi complex 5	HLA-B*1801 [82], HLA-B*1501 (B62) [9], HLA-B*4901 [28]	1q23
Eukaryotic translation initiation factor 3, subunit 8, 110kDa	HLA-B*4001 [28], HLA-B*4901 [28]	7q22-q31
GTF2IRD2 beta	HLA-B*1801 [82], [9], HLA-B*4901 [28]	16p11.2
Heat shock protein 90kDa α, class B member 1	HLA-A*0201 [9], HLA-B*4901 [28]	6p12
HLA-B associated transcript 3	negative	6p21.3
LIM and senescent cell antigen-like domains 2	negative	2q14.3
Muscle RAS oncogene homolog	negative	3q22.3
Myosin IE	HLA-A*1101, HLA-B*08 [9]; HLA-B*08 [52], HLA-B*4901 [28]	15q21-q22
Ribonucleotide reductase M2 B (TP53 inducible)	negative	8q23.1
Ribosomal protein L14	HLA-A*25, HLA-B*1501 (B62) [9], HLA-B*4901 [28]	3p22-p21.2
Septin 9	HLA-B*1509 [78], HLA-B*1510 [46], HLA-Cw*0102 [83]	17q25
similar to CG14977-PA	negative	7q22.1
Ubiquilin 1	negative	9q22 | 9q21.2-q21.3
Progestin and adipoQ receptor family member IV	HLA-A*2601	16p13.3

Table 2 lists the names of the identified source proteins previously defined HLA ligands and their chromosomal localization. The references cited in column 2 were extracted from the SYFPEITHI data base (http://www.syfpeithi.de/scripts/MHCServer.dll/home.htm).

Supplementary Table 2: GO annotation for source proteins defined by the ligandomics approach

A

	GO:0005515 protein binding	GO:0003824 catalytic activity	GO:0003676 nucleotide acid binding	GO:0000166 nucleotide binding	GO:0046872 metal ion binding	GO:0005198 structural molecule activity	GO:0030234 enzyme regulator activity	GO:0008289 lipid binding	GO:0004871 signal transducer activity	GO:0005215 transporter activity	others
EIF3C		■	■	■							■
MYO1E	■	■		■							
SCP2	■	■						■			
RASAL2	■	■			■						
DEGS1	■	■									■
RPL14	■		■			■					
TFE3	■		■								■
RPSA	■		■	■							
HSP90AB1	■			■			■				
ACTA1	■			■		■					
IGFBP3	■				■		■				
APOL1	■							■		■	
MCM7		■				■					■
AKR1C3	■	■									
OR13F1	■	■									
BRD4	■		■								
EIF3E	■		■								
LSM8	■		■								
RANBP9	■		■								
SNRPG	■		■								
CCT3	■			■							
CLTC	■					■					
MRAS		■		■							
RRM2B		■			■						
GNG5		■							■		
PAQR4		■							■		
GFPT1		■									■
GTF2IRD2B			■		■						
COL18A1					■	■					
GAPT	■										
GPSM1	■										
HECTD1	■										
HLA-A	■										
HLA-C	■										
LIMS2	■										
UBQLN1	■										
BAT3		■									
CACNA1E		■									
WIZ		■									
SEPT9					■						
PIGM							■				
SRXN1										■	

B

	GO:0005737 cytoplasm	GO:0016020 membrane	GO:0005634 nucleus	GO:0005856 cytoskeleton	GO:0005783 endoplasmic reticulum	GO:0005615 extracellular space	GO:0005739 mitochondrion	others
RPSA	■	■	■					
HECTD1	■		■		■			
HSP90AB1	■						■	■
LIMS2	■	■	■					
UBQLN1	■	■						■
RPL14	■							
SNRPG	■		■					
WIZ	■		■	■				
EIF3E	■		■					
ACTA1	■			■				
CCT3	■			■				
SCP2	■						■	
RRM2B	■		■					
AKR1C3	■							
BRD4	■		■					
MCM7	■		■					
MRAS		■						
DEGS1		■			■			
GPSM1		■			■			
SRXN1		■			■			
APOL1		■				■		
GNG5		■					■	
CACNA1E		■						■
GTF2IRD2B			■					
LSM8			■					
IGFBP3			■			■		
MYO1E				■				■
COG5	■							
EIF3C	■							
GFPT1	■							
RANBP9	■							
CLTC		■						
GAPT		■						
HLA-A		■						
HLA-C		■						
ORMDL3		■						
PAQR4		■						
BAT3			■					
TFE3			■					
RASAL2								■
SEPT9								■
COL18A1						■		
HTRA1						■		
OR13F1						■		
						■		

50 distinct proteins were classified based on GO annotations as outlined in the Experimental Procedures according to (A) their putative function and (B) their subcellular localization. For C2orf54, hypothetical LOC400614, similar to CG14977-PA, COG5, HTRA1, INPP5F, ORMDL3 no functional annotations exist and for C2orf54, hypothetical LOC400614, similar to CG14977-PA, INPP5F, PIGM no subcellular localization are known.

Supplementary Table 3: Colored heatmap of those targets which were also identified as members of the ligandome of the cell line MZ3733RC

Gene symbol	**2733**	**2820**	**2825**	**2826**	**2865**	**2875**
ACTA1	3.47	-0.29	-0.10	-0.52	0.53	
AKR1C3	0.47		0.69	2.41	2.64	-0.03
APOL1		0.26	0.84	-0.23	1.73	-0.23
BRD4	0.48	-0.77	-0.28	-0.02	-0.12	-0.34
CCT3	-1.36	-0.29	-0.77	1.40	0.66	
CLTC	2.83	0.96	0.51	1.87	-1.75	-0.51
COL18A1	-0.81	0.84	3.58	0.80	-1.97	1.92
COG5	0.64	1.06	0.17	0.89	-0.41	0.20
DEGS1	1.21	0.53	0.56	1.35	-0.17	1.13
HSP90AB1	-1.24	0.72	0.04	0.74	0.22	0.72
HECTD1	-1.26	-0.23	0.80	1.31	-2.20	-0.29
BAT3	-0.84	-0.13	-0.08	-0.30	0.45	0.61
INPP5F	-0.90	0.16	0.66	-0.31	-1.14	0.21
IGFBP3	2.43	1.44	2.71	3.16	1.99	1.22
HLA-A	1.01	0.57	0.57	0.14	0.11	0.58
HLA-C	-0.19	1.34	0.53	0.06	0.26	1.78
MCM7	-0.37	0.39	0.11	2.66	0.94	1.30
MRAS	-0.22	0.15	0.08	-0.03	-0.26	0.00
MYO1E	0.50	0.18	0.30	0.68	1.23	-0.38
PIGM		0.28	0.44	0.05		-0.06
RASAL2	-0.38	0.19	1.13	0.19	0.53	0.03
RRM2B	4.89		1.14	1.06	0.28	0.11
RPSA	-0.06	-0.08	-0.58	0.43	1.45	0.85
SEPT9	-0.10	-0.06	-0.43	-0.06	-0.87	-0.43
SCP2	-0.38	0.02	0.31	-0.16	-0.71	-0.80
TFE3	-0.33	0.29	0.45	-0.11	0.60	0.20
UBQLN1		-0.44	-0.39	-0.79	0.00	-0.35
WIZ	-1.16	0.54	0.66	0.71	-0.90	0.37

🟩 -2	⬛ 0	🟥 +2

The table lists the gene symbols along with the log fold changes defined in a subset of 6 previously analysed tissue systems comprised of renal cancer vs. tumor adjacent normal renal epithelium [44]. The bottom line defines the color coding of the regulation mode. Data which were not interpretable on the array are marked grey**.**

Supplementary Table 4: Target verification for source proteins of identified HLA class I ligands

protein	UniProtKB ID	gene ID	Human Protein Atlas		cDNA microarray data
			ab code	validation score	
Sterol carrier protein 2	P22307	SCP2	HPA027317 HPA027317 HPA027317	uncertain	yes
Glutamine-fructose-6-phosphate transaminase 1	Q06210	GFPT1	no ab in depository	no scoring	no probes on array
Ribosomal protein SA	P08865	RPSA	CAB009561	uncertain	yes
Apolipoprotein L, 1	**O14791**	**APOL1**	**HPA018885**	**supportive**	**yes**
Calcium channel, voltage-dependent, alpha 1E subunit	Q15878	CACNA1E	no ab in depository	no scoring	no probes on array
Clathrin, heavy polypeptide (HC)	**Q00610**	**CLTC**	**CAB010389** **CAB011571** **CAB017155**	**supportive**	**yes**
Collagen, type XVIII, alpha 1	P39060	COL18A	CAB001961	uncertain	yes
Hypothetical LOC400614			no ab in depository	no scoring	no probes on array
Chromosome 2 open reading frame 54 (hypothetical protein FLJ22671)	Q53QU5	C2orf54	no ab in depository	no scoring	no probes on array
Isochorismatase domain containing 2	**O95777**	**LSM8**	**HPA020116**	**supportive**	**no probes on array**
ORM1-like 3 (S. cerevisiae)	Q8N138	ORMDL3	no ab in depository	no scoring	no probes on array
Phosphatidylinositol glycan, class M	Q9H3S5	PIGM	no ab in depository	no scoring	yes
RAS protein activator like 2	Q9UJF2	RASAL2	HPA018805 HPA020453	low uncertain	yes
Small nuclear ribonucleoprotein polypeptide G	P62308	SNRPG	no ab in depository	no scoring	no probes on array
Sulfiredoxin 1 homolog (S. cerevisiae)	Q9BYN0	SRXN1	no ab in depository	no scoring	no probes on array
Major histocompatibility complex, class I, C	P10321	HLA-C	no ab in depository	no scoring	yes
Actin, alpha 1, skeletal muscle	**P68133**	**ACTA1**	**CAB000045**	**supportive**	**yes**

Aldo-keto reductase family 1, member C3 (3-alpha hydroxy-steroid dehydrogenase, type II)	P42330	AKR1C3	CAB010874	uncertain	yes
Degenerative spermatocyte homolog 1, lipid desaturase (Drosophila)	O15121	DEGS1	HPA014763	uncertain	yes
Eukaryotic translation initiation factor 3, subunit 6 48kDa	**P60228**	**EIF3E**	**HPA023973**	**supportive**	**no probes on array**
G-protein signalling modulator 1 (AGS3-like, C. elegans)	Q86YR5	GPSM1	no ab in depository	no scoring	no probes on array
Guanine nucleotide binding protein (G protein), gamma 5 subunit	P63218	GNG5	CAB032623	uncertain	no probes on array
HECT domain containing 1	Q9ULT8	HECTD1	HPA002929	uncertain	yes
Insulin-like growth factor binding protein 3	P17936	IGFBP3	CAB010360 HPA013357	uncertain	yes
Major histocompatibility complex, class I, A	P16189	HLA-A	no ab in depository	no scoring	yes
MCM7 minichromosome maintenance deficient 7 (S. cerevisiae)	P33993	MCM7	CAB002163 HPA003898 CAB016312	supportive	yes
Protease, serine, 11 (IGF binding)	**Q92743**	**HTRA1**	**HPA036655**	**supportive**	**no probes on array**
RAN binding protein 9	**Q96S59**	**RANBP9**	**CAB033767**	**supportive**	**no probes on array**
Transcription factor binding to IGHM enhancer 3	**P19532**	**TFE3**	**HPA023881**	**supportive**	**yes**
Widely-interspaced zinc finger motifs	Q96IG5	WIZ	HPA022923	uncertain	yes
Growth factor receptor-bound protein 2-binding adapter protein (protein GAPT; (hypothetical protein FLJ33641)	Q8N292	GAPT	HPA011070	uncertain	no probes on array
Inositol polyphosphate-5-phosphatase F	Q5W135	INPP5F	no ab in depository	no scoring	yes
General transcription factor II-I repeat domain-containing protein 2B	Q6EKJ0	GTF2IRD 2	no ab in depository	no scoring	no probes on array
Bromodomain containing 4	O60885	BRD4	HPA015055	uncertain	yes
Chaperonin containing TCP1, subunit 3 (gamma)	P49368	CCT3	HPA006543	uncertain	yes

Component of oligomeric Golgi complex 5	**Q9UP83**	**COG5**	**HPA020300**	**supportive**	**yes**
Eukaryotic translation initiation factor 3, subunit 8	Q99613	EIF3C	no ab in depository	no scoring	no probes on array
Heat shock 90kDa protein 1, beta	P08238	HSP90AB1	CAB005230	uncertain	yes
HLA-B associated transcript 3	**P46379**	**BAT3**	**CAB020704**	**supportive**	**yes**
LIM and senescent cell antigen-like domains 2	Q7Z4I7	LIMS2	no ab in depository	no scoring	no probes on array
Muscle RAS oncogene homolog	O14807	MRAS	no ab in depository	no scoring	yes
Myosin IE	Q12965	MYO1E	HPA023886	uncertain	yes
Ribonucleotide reductase M2 B (TP53 inducible)	Q7LG56	RRM2B	HPA028812	supportive	yes
60S ribosomal protein L14 (ribosomal protein L14)	P50914	RPL14	HPA004136	medium supportive	no probes on array
Septin 9	Q9UHD8	SEPT9	no ab in depository	no scoring	yes
Similar to CG14977-PA			no ab in depository	no scoring	no probes on array
Ubiquilin 1	**Q9UMX0**	**UBQLN1**	**CAB037256**	**supportive**	**yes**
Olfactory receptor, F13, member 1	Q8NGS4	OR13F1	no ab in depository	no scoring	no probes on array
Progestin and adipoQ receptor family member IV	Q8N4S7	PAQR4	no ab in depository	no scoring	no probes on array

Supplementary Table 2 lists the names of the source proteins from the identified HLA ligands along with their UniProtKB ID, gene ID, the antibody code and the IHC validation score extracted from the Human Protein Atlas as well as their representation on the cDNA microarray. The shadings indicate that the respective source proteins were either represented on both the cDNA microarray as well as in the antibody depository of the Human Protein Atlas (dark grey) or only in one of the two next to the qPCR-verified SRXN1 (light grey). The bold lettering indicates that the respective IHC staining pattern stored in the Human Protein Atlas is supportive. For the remaining source proteins currently no further validation data are available.

Supplementary Figure legends

Supplementary Fig. 1

Schematic diagram of the ligandomics analysis work flow

Briefly, peptide loaded HLA class I molecules presented on the cell surface of MZ2733RC cells were purified from lysates by immunoprecipitation with the antibody W6/32 followed by protein affinity chromatography. The HLA ligands were eluted by acid treatment and subsequently concentrated by ultrafiltration. The resulting pool of peptides was further fractionated using a microcapillary reverse phase high performance liquid chromatography (HPLC) column online coupled to a Q-TOF mass spectrometer equipped with a micro-ESI source. Peptide sequence analysis was performed manually by running sequence tag searches using Mascot 2.0 software. However, relevant hits were manually assessed and not solely by evaluation of the Mascot score but by applying further selection criteria as outlined in the experimental procedures section.

Supplementary Fig. 2

MS chromatograms of HLA-A, -B and -C ligands eluted from the cell line MZ2733RC.

Total ion current (A) and one of four MS/MS chromatograms (B) to detect doubly charged ions at m/z 300-550.

Supplementary Fig. 3

Immunohistochemical analysis of normal renal epithelium and renal cancer tissue sections for the expression of HLA-B-associated transcript 3 (BAT3).

Randomly selected sections of TMA representing normal renal epithelium (left segment) and renal cancer tissue (right segment), all taken from the deposit available via the Human Protein Atlas (HPA, Version 7) were selected to document the immunohistochemical staining with the

HLA-B-associated transcript 3-specific antibody CAB020704

HLA-B-associated transcript 3 shows moderate to strong nuclear and slightly weaker cytoplasmic immunoreactivity in most normal and malignant tissues. Renal cancer tissues frequently show strong cytoplasmic positivity with high intensity.

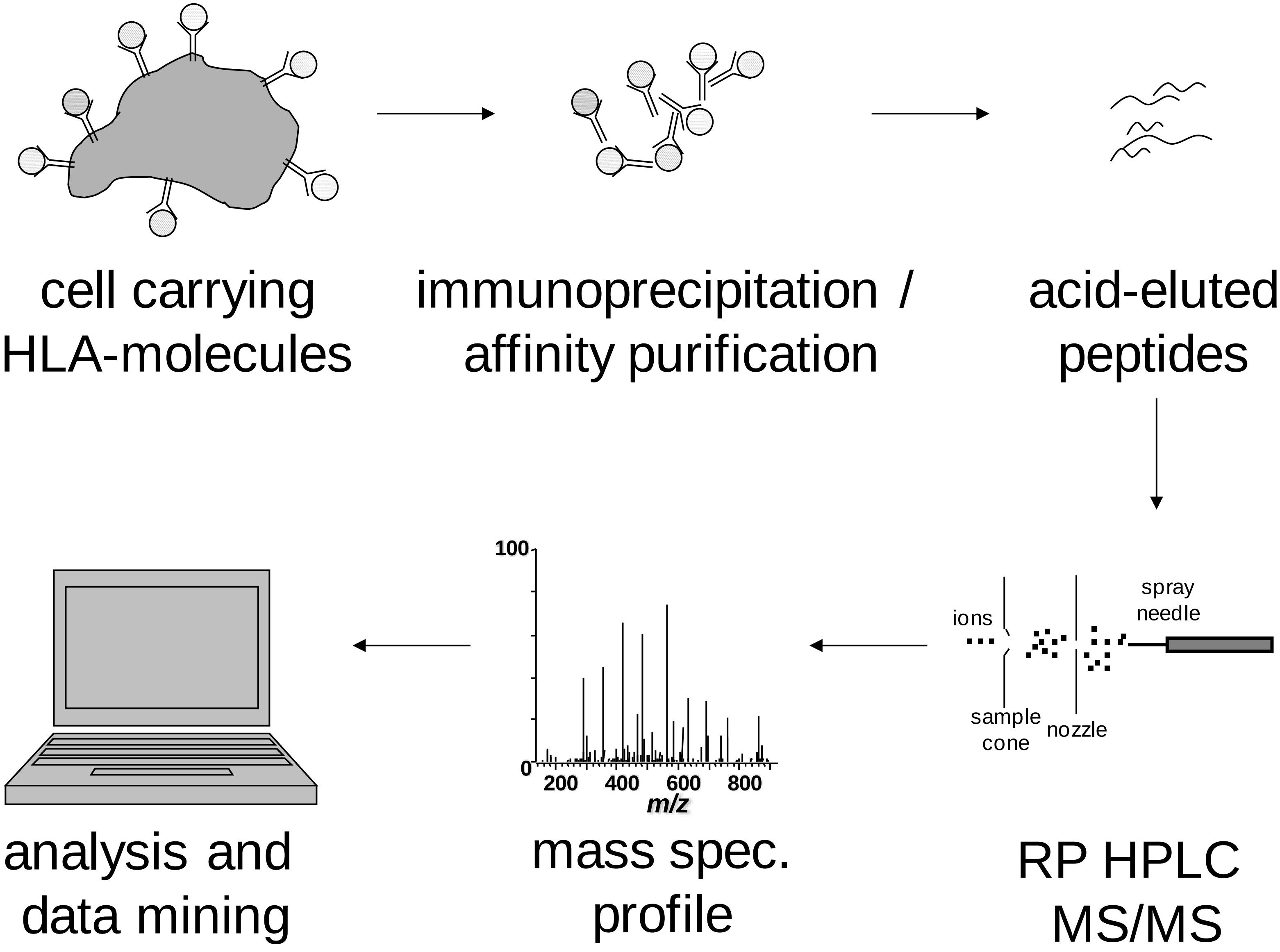
Supplementary Fig. 4

Immunohistochemical analysis of normal renal epithelium and renal cancer tissue sections for the expression of minichromosome maintenance complex component 7 (MCM7)

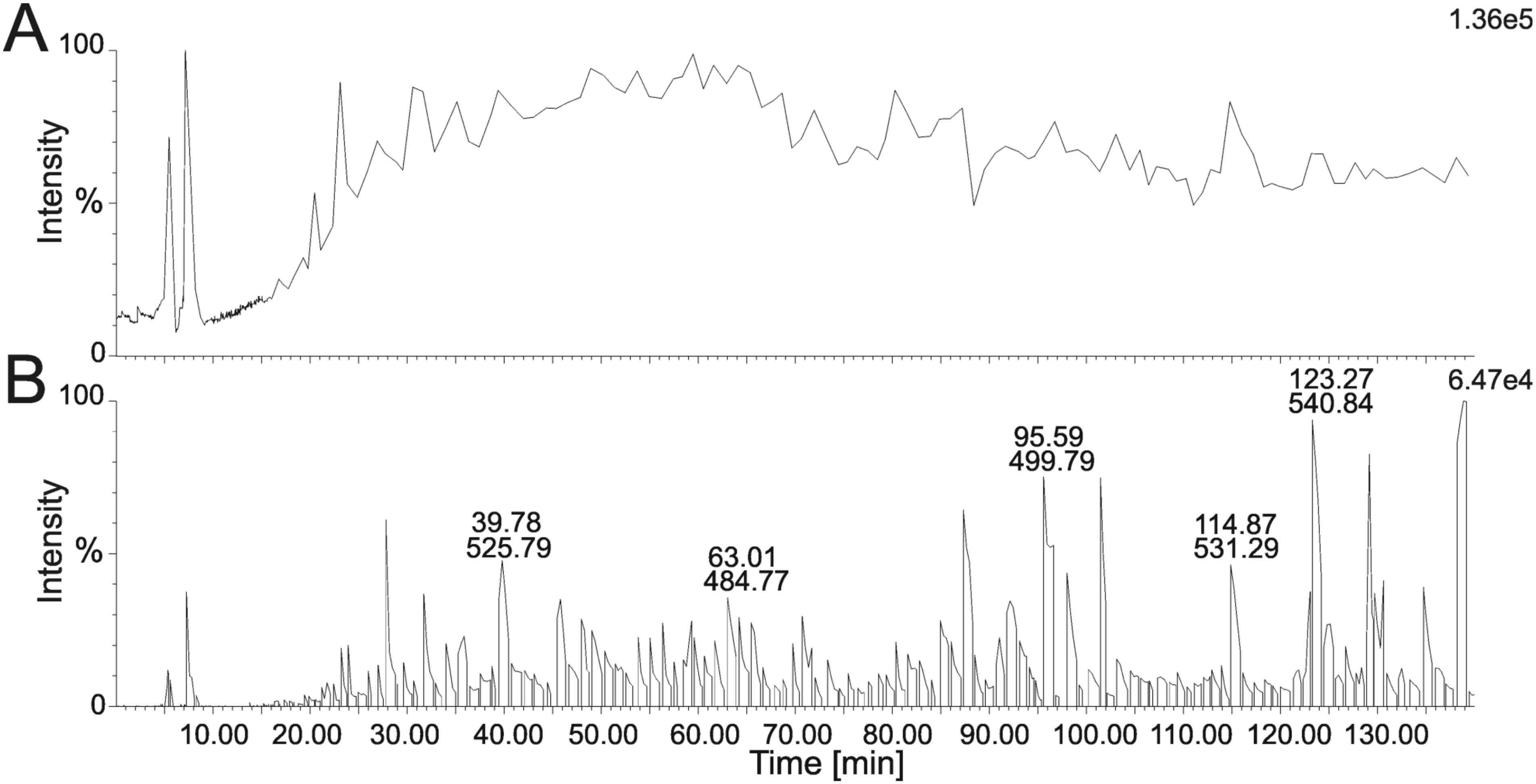
Randomly selected sections of TMA representing normal renal epithelium (left segment) and renal cancer tissue (right segment), all taken from the deposit available via the Human Protein Atlas (HPA, Version 7) were selected to document the immunohistochemical staining pattern of the minichromosome maintenance complex component 7-specific Ab (CAB016312).

Minichromosome maintenance complex component 7 is not expressed in glomeruli and only rarely at moderate levels in tubules showing nuclear localization. In contrast renal cancer tissues show both a higher frequency and higher staining intensity of MCM7 positive cells.

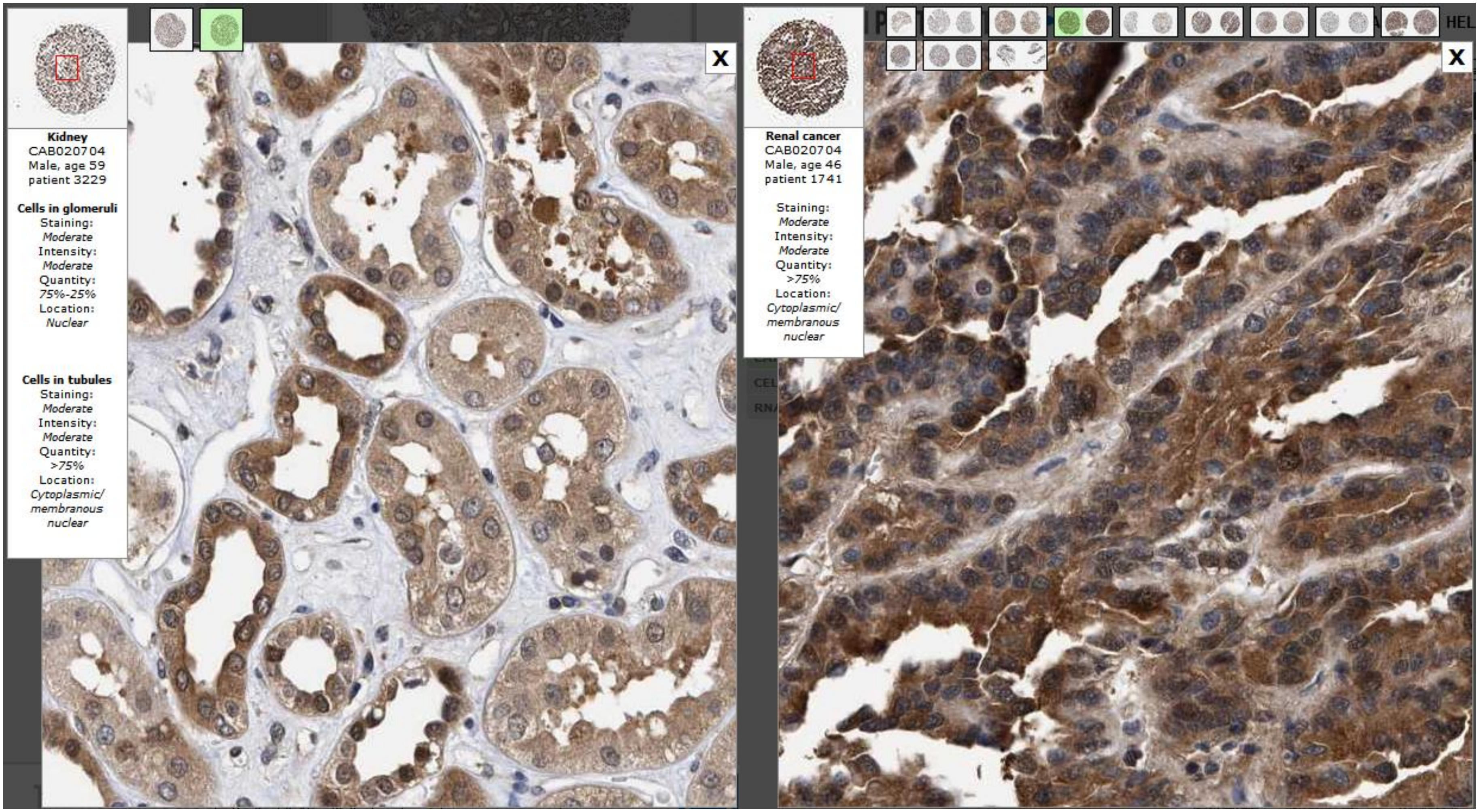
Supplementary Fig. 1



Supplementary Fig. 2



Supplementary Fig. 3



Kidney
CAB020704
Male, age 59
patient 3229

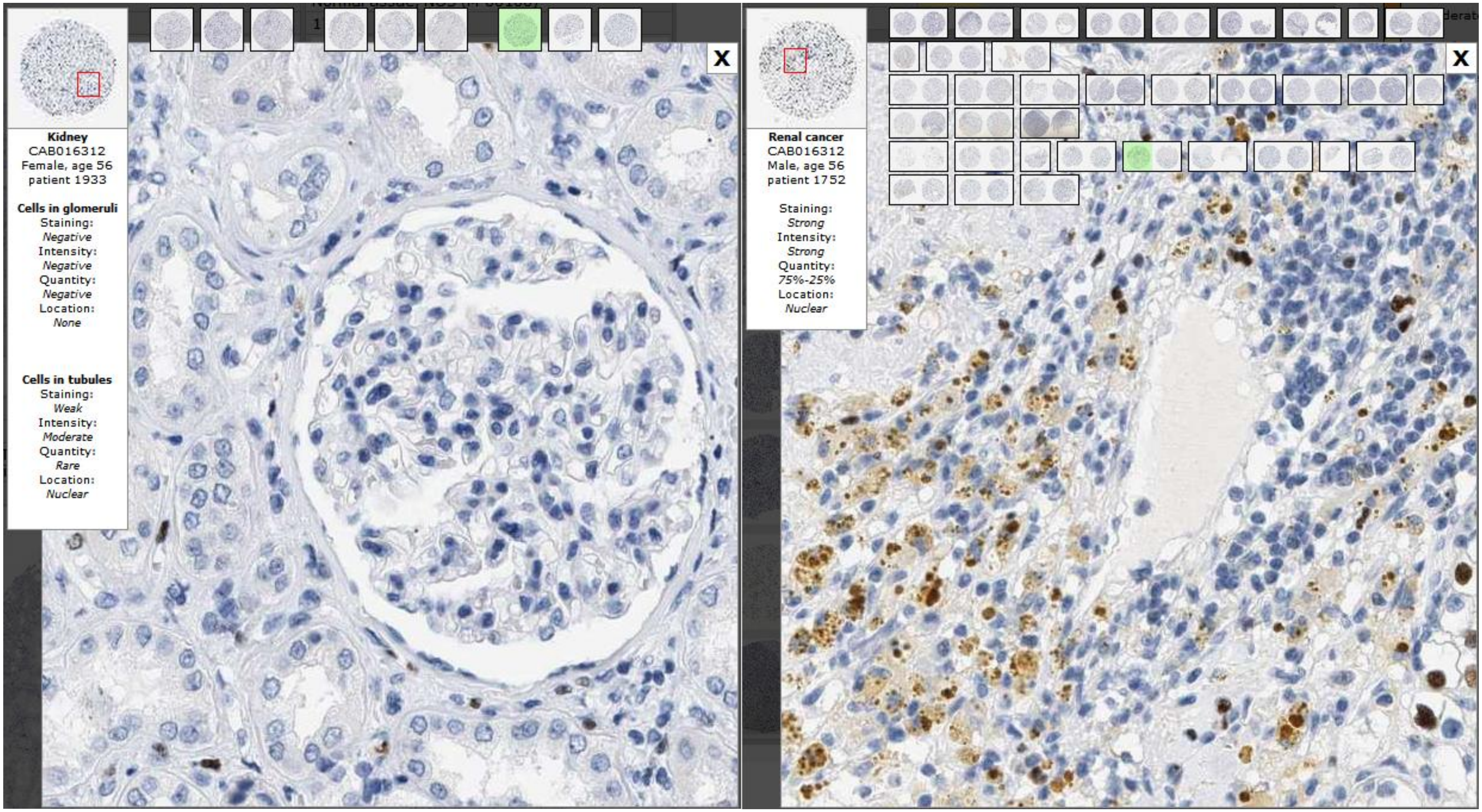
Cells in glomeruli
Staining:
Moderate
Intensity:
Moderate
Quantity:
75%-25%
Location:
Nuclear

Cells in tubules
Staining:
Moderate
Intensity:
Moderate
Quantity:
>75%
Location:
Cytoplasmic/
membranous
nuclear

Renal cancer
CAB020704
Male, age 46
patient 1741

Staining:
Moderate
Intensity:
Moderate
Quantity:
>75%
Location:
Cytoplasmic/
membranous
nuclear

Supplementary Fig. 4



Kidney
CAB016312
Female, age 56
patient 1933

Cells in glomeruli
Staining:
Negative
Intensity:
Negative
Quantity:
Negative
Location:
None

Cells in tubules
Staining:
Weak
Intensity:
Moderate
Quantity:
Rare
Location:
Nuclear

Renal cancer
CAB016312
Male, age 56
patient 1752

Staining:
Strong
Intensity:
Strong
Quantity:
75%-25%
Location:
Nuclear