Carcinogenesis of renal cell carcinoma reflected in HLA ligands: A novel approach for synergistic peptide vaccination design

Martin G. Klatt, Daniel J. Kowalewski, Heiko Schuster, Moreno Di Marco, Jörg Hennenlotter, Arnulf Stenzl, Hans-Georg Rammensee, and Stefan Stevanovic

Department of Hematology and Oncology, University of Tübingen, Tübingen, Germany; Institute for Cell Biology, Department of Immunology, University of Tübingen, Tübingen, Germany; German Cancer Consortium (DKTK), DKFZ Partner Site, Tübingen, Germany; Department of Urology, University of Tübingen, Tübingen, Germany; Institute for Cell Biology, Department of Immunology, University of Tübingen, Tübingen, Germany

ABSTRACT
Despite recent advances in immunotherapy of renal cell carcinoma (RCC), peptide vaccination strategies still lack an approach for personalized peptide vaccination that takes intra- and inter-tumoral heterogeneity and biological characteristics into account. In this study, we use an immunoprecipitation and mass spectrometry-based approach supplemented by network analysis of HLA ligands to target this goal. By analyzing HLA-presented peptides for HLA class I and II of 11 malignant and 6 non-malignant kidney tissue samples, more than 2,700 peptides and 1,600 corresponding source proteins were identified. A high overlap with HLA ligands derived from peripheral blood mononuclear cells (PBMCs) was detected most likely due to tumor-infiltrating inflammatory cells and therefore excluded from network analysis. Subsequent biological function analysis of HLA ligands by the GeneMANIA online platform showed enrichment for well established, but also novel, pathways and biological processes involved in carcinogenesis of RCC almost exclusively in malignant tissue samples. By exploring source proteins involved in these overrepresented pathways, we verified various known tumor-associated antigens (TAAs) for RCC (e.g., CA9, EGLN3, IGFBP3, MMP7, PAX2, VEGFA, or EGFR) but could also detect novel TAAs for RCC (e.g., PLOD2, LOX, ENPEP, or TGFB1). Some of these new TAAs had already been shown to elicit a T cell response in cancer patients. Thus, network analysis of HLA ligands provided a new platform for implementing personalized, multipeptide vaccines with potentially synergistic antitumor effects.

INTRODUCTION
Based on case reports describing spontaneous tumor regression that date back to the early 1960s, renal cell carcinoma (RCC) was first hypothesized to be immunogenic almost 20 years ago. Over the years, this hypothesis has been confirmed and has led to various studies that focus on immunotherapies for RCC. These include unspecific interleukin-2 (IL-2) therapy and specific strategies such as autologous vaccines and dendritic cell-based as well as RNA- or peptide-based approaches. Recently, in a multipeptide vaccination phase II trial, overall patient survival was prolonged by the induction of specific T cell response. Although a corresponding phase III trial did not confirm these results (data not published) other vaccination trials, e.g., in head and neck squamous cancer, showed induction of T cell response and prolonged overall survival. These results demonstrate the potency of this strategy but also reveal the need for an improved vaccine design for RCC.

Striking results in the field of checkpoint inhibition for other immunogenic tumors, and for RCC, underscore the power of immunotherapy and the potential of a two-sided approach that...
combines specific and unspecific immunotherapy for the treatment of RCC.11-13

However, peptide vaccine design still lacks a sustainable approach to generate individualized vaccination cocktails that consist of several immunogenic tumor-associated antigens (TAAs) which is essential for adequately addressing intra- and intertumoral heterogeneity.14,15 The most elegant approach to target this goal is the identification of mutated, immunogenic HLA-presented peptides, which have been discovered in mice transgenic for human MHC16 and are shown to induce tumor rejection in sarcoma-bearing mice17 but have not yet been identified in any other species. Alternative strategies are therefore required.

Here, we isolated HLA-presented peptides mostly from triplets of primary kidney tissue samples consisting of normal, tumorous, and metastatic kidney tissue by immunoprecipitation and subsequent analysis by liquid chromatography tandem mass spectrometry (LC/MS-MS). Subsequently, network analyses with source proteins derived from HLA ligands were performed. The analyses revealed several tumor-specific, enriched signaling pathways, and biological processes in the HLA ligandome of RCC tissue samples and their metastasis that are known to be essential in tumorigenesis of RCC. They include the cellular response to hypoxia, epidermal growth factor receptor, platelet-derived growth factor, and fibroblast growth factor receptor signaling pathways.18-24 Therefore, we provided first-time evidence that key proteins representing the carcinogenesis of RCC give rise to MHC-binding peptides and are therefore targetable by the adaptive specific T cell response.

Furthermore, we used this insight to develop a novel strategy for identifying new TAAs and corresponding HLA-restricted peptides. For those TAAs and HLA ligands which had not already been described to be immunogenic, we performed in silico predictions to gauge their potential for eliciting T-cell responses. However, further confirmatory in vitro assays corroborating the in silico experiments will be required prior to the implementation of these new HLA ligands in peptide vaccination studies. Nevertheless, with this approach we were able to engineer models for individualized peptide vaccines with at least 15 TAAs per patient that are not only tumor specific but also synergistically directed against crucial biological processes in carcinogenesis of RCC.

Results

**LC-MS/MS identifies a high amount of naturally presented HLA class I and II ligands in tissue and blood samples**

Broadly, in this study our main goal was to depict a new approach for defining TAAs and developing peptide vaccines for HLA class I and II that focuses on biological processes enriched in RCC and metastases by using network analysis (Fig. 1 and Fig. S1). Four triplets consisting of normal, tumorous, and metastatic tissue samples were included in this study. To check whether our findings are transferable to individuals for whom no triplets were available, one patient with only normal and tumorous tissue samples and one with two resectable metachronous metastases were also included in this study. After isolation of HLA ligands and analysis via mass spectrometry, we confirmed these results by matching them with already published HLA ligands or with peptides derived from a large in-house database containing tumorous and normal but no metastatic kidney tissue samples generated by independent experiments. Data was then processed by network analysis via GeneMania and enriched pathways were checked for TAAs.

In our first step that involved mapping, the HLA class I ligandomes of 17 different tissue samples derived from either normal, tumorous, or metastatic (and in one case adenomatous) kidney tissue from six different clear cell renal cell carcinoma (ccRCC) patients (Table 1), we were able to identify a total of 1,910 different peptides representing 1,411 source proteins. The number of validated distinct peptides per patient ranged from 415 to 633 (mean 509 peptides; Fig. 2). For healthy volunteers (30 peripheral blood mononuclear cell (PBMC) donors, five bone marrow mononuclear cell (BMMC) donors), a total of 17,940 unique peptides were identified (17,322 peptides/7,207 source proteins on PBMCs; 1,738 peptides/1,384 source proteins on BMMCs; Table S1) for HLA class I.

In the case of HLA class II, a total of 805 unique peptides (range 96–497 peptides/patient, mean 252 peptides; Fig. 2) representing 222 source proteins were detected in ccRCC patients. The HLA class II healthy control cohort (20 PBMC, 10 BMMC donors) yielded 4,240 different peptides (2,482 peptides/667 source proteins on PBMCs; 1,758 peptides/932 source proteins on BMMCs; Table S2). Matching HLA alleotypes of HLA-A and HLA-B alleles of healthy volunteers with the RCC patient cohort showed 100% coverage. No correlation of sample size and number of peptide identifications was found for HLA class I (Spearman’s r = 0.21) and only a trend for HLA class II (r = 0.79; Fig. S2).

**HLA ligandome-derived source proteins are equally distributed among tumorous, metastatic, and normal kidney tissues and show a large overlap with PBMC antigens**

To assess the distribution of 1,633 identified source proteins between normal, tumorous, and metastatic kidney tissues, we used the BioVenn tool. It revealed no significant differences between the three tissue subtypes for HLA class I or II (Fig. 3A and B). Furthermore, overlap analysis from pooled tissue samples with PBMC-derived source proteins showed high identity with 78.6% (1,109 proteins) for HLA class I and 53.9% (111 proteins) for HLA class II, respectively (Fig. 3C and D). Thus, to prevent biased data we repeated the initial analysis with 302 or 90 remaining proteins after sorting out proteins found on both tissue and PBMC samples. These samples are later defined as reassessed. Again, no major differences could be discovered between proportions of identified proteins in tumorous, metastatic, or normal tissue (Fig. 3E and F). Summarized, the vast overlap of HLA ligands of PBMCs and kidney tissues could be explained by a high infiltration of inflammatory cells in all tissue samples analyzed or shared proteins between PBMCs and cells derived from kidney tissue samples. In contrast, the high amount of shared proteins in all three tissue subtypes of reassessed kidney samples might represent peptides derived from house-keeping proteins.

**Network analyses of source proteins derived from HLA class I but not class II ligandome reveal high functional similarity of tumorous and metastatic but not tumorous and normal kidney tissue**

Since carcinogenesis is accompanied by changes of the intra- and extra-cellular proteome, we further investigated whether
these aberrations are indirectly represented in the HLA ligand profile. We performed network analyses for each reassessed tissue sample individually with source proteins derived from HLA ligands to determine characteristics and distribution of enriched biological processes in tumorous, metastatic, and normal kidney tissue. After demonstrating a rather uniform overlap of source proteins between these three tissue subtypes, we first focused on distribution of biological processes identified by network analysis.

HLA ligands of tissue samples were matched with PBMC data and duplicates sorted out as described above, taking into account the possible high infiltration by inflammatory cells. These input lists were processed by the GeneMANIA software tool.

Analysis of all tissue samples combined yielded 217 unique enriched pathways and biological processes for healthy kidney tissue, 209 for RCC and 557 for metastases. We compared the distribution of these pathways with HLA class I derived peptides, which like source proteins showed no significant differences for overlaps between tumorous and metastatic, metastatic and normal, as well as tumorous and normal tissue samples. This is in contrast to patterns reported in a previous study in

Table 1. Samples characteristics.

| ISN    | ST | TNM   | HLA typing, class I | Sample weight [g] |
|--------|----|-------|---------------------|-------------------|
| RCC 70 | N, T, M | pT3a pN1 M1 G2 | A’01, A’02, B’07, B’08 | 2.64 ; 1.19 ; 0.14 |
| RCC 329| N, T, M | pT3c pN0 M1 G2 | A’02, A’32, B’14, B’35 | 0.53 ; 0.86 ; 0.23 |
| RCC 377| N, T, M | pT3b pN1 M1 G3 | A’01, A’03, B’07 | 0.17 ; 0.17 ; 0.74 |
| RCC 399| N, T, M | pT3b pN2 M1 G3 | A’01, A’02, B’08, B’50 | 0.31 ; 0.43 ; 1.02 |
| RCC 210| N, T, A | pT1b pN0 M0 G2 | A’01, A’24, B’08, B’51 | 1.35 ; 3.00 ; 0.99 |
| RCC 297| M1, M2 | pT3b pN1 M1 G2 | A’03, A’29, B’35, B’44 | 0.38 ; 0.25 |

**Abbreviations:** ISN, internal sample number; ST, subtypes; N, normal tissue; T, tumorous tissue; M, metastatic tissue; A, adenomatous tissue. Samples were provided by the Department of Urology, University Hospital Tuebingen and then stored at – 80°C until further use.
two RCC patients analyzed on mRNA and peptide level\(^{25}\) (Fig. 4A and B). HLA class I comparative analysis revealed first, the highest overlap of enriched pathways between tumorous and metastatic tissue, representing a similar biological behavior (Fig. 4C); second, a similarity between metastatic and normal tissue that was significantly lower \((p < 0.03)\) compared to tumor and metastasis when analyzed in samples from the same patient by paired t-test; and third, a complete mismatch of enriched pathways in tumorous and normal tissue without a single-shared biological function. Thus, functional overlap of tumor and metastasis was significantly higher \((p = 0.02)\) than between tumorous and normal tissue samples.

Furthermore, grouped pathway analysis of all HLA class I antigens confirmed the disparity of tumorous and normal tissue (Fig. 4D), which demonstrates superior sensitivity of pathway analysis compared to peptide and source protein comparison.
Together, this illustrates perfectly the differences of tumorous and normal kidney tissue as well as the similarity of tumor and metastasis in regard to biological features. Remarkably, these functional characteristics from intracellular processes are maintained in HLA ligands despite the high overlap of source proteins used for analysis.

In contrast, for HLA class II no definite tendencies for a distinct distribution of the three different tissue modalities could be found, neither for HLA ligands nor for enriched pathways and biological functions (Fig. S3A–C). The most likely explanation for this discrepancy of HLA class I and II is that even though it is functionally different between tumorous, normal, and metastatic tissue, extracellular space is built from similar components all represented in the three different tissues which leads to less plasticity in class II antigen presentation in malignant versus non-malignant tissues compared to HLA class I. This is also demonstrated by the biggest overlap representing 187 pathways (44.6% of theoretical complete overlap) in Venn diagram between normal, tumorous, and metastatic tissue samples when combining all enriched pathways for HLA class II (Fig. S3D). Nevertheless, tumor associated and specific pathways can be detected for both HLA class I and II as described later.

**Tumor-associated pathways are enriched in the HLA ligandomes of malignant tissue samples**

After identifying distinct differences in the distribution of pathways between malignant and non-malignant tissue samples for HLA class I, we scanned enriched pathways detected in tumors and metastases for processes involved in carcinogenesis. Network analysis revealed highly significant and strong enrichment for several tumor-associated pathways and biological functions for both HLA class I and II. Especially for HLA class I, these tumor pathways were often among the most enriched processes in these samples (Fig. 5A–C, Table 2). Results included well-established mechanisms of tumorigenesis in ccRCC such as cellular response to hypoxia as well as epidermal growth factor, platelet-derived growth factor, and fibroblast growth factor signaling pathways. Additionally, besides involvement of several growth factor receptor pathways, extracellular matrix organization was found in 72.7% of all malignant specimens and was often accompanied by collagen catabolism and positive regulation of cellular motility, all of which are involved in promoting metastasis. For HLA class I, platelet degranulation was also considered as a tumorigenic process leading to pro-angiogenic signaling, as shown for ovarian cancer.

For HLA class II, blood vessel development which is one of the most important hallmarks of cancer, especially in ccRCC, was detected in 72.7% of analyzed malignant tissues (Table 2). Epithelial cell proliferation, which is another basic process in tumorigenesis, was also identified. Further mechanisms, including collagen catabolic processes, growth factor, and platelet-derived growth factor binding plus positive regulation of cell motility, were also found in HLA class I analysis, thus underlining the consistency of our findings.

As to the distribution of these pathways: all processes shown in Table 2 which were found more than once were present in both tumorous and metastatic kidney tissues, which emphasizes the importance of these pathways in both tumor formation and metastasis. Only two processes (blood vessel development and growth factor binding) were found to be enriched once, and only for HLA class II in healthy kidney tissue (RCC399N). Nevertheless, this aberrant finding can be explained by tumor-infiltrating HLA class II positive cells such as macrophages, CD4+ T-cells or dendritic cells presenting TAA s. On the other hand, the patient corresponding to sample RCC399N suffered from diabetic nephropathy which leads to increased angiogenesis and elevated levels of transforming growth factor signaling in kidney tissue.

Together, HLA ligandome-derived peptides and their corresponding source proteins are able to represent biological
intracellular processes which are involved in tumorigenesis and metastasis, predominantly in malignant tissue samples.

Pathways enriched in malignant tissue samples contain high amounts of individual tumor-associated antigens and corresponding immunogenic HLA ligands suitable for vaccine design

Finally, we explored the possibility that tumor-associated and highly specific pathways found in RCC samples can lead to the verification of TAAs and the definition of a novel TAAs. By analyzing the source proteins of the pathways, we were able to identify several antigens with their corresponding peptides found almost exclusively on malignant tissue samples (Tables 2 and 3). Due to prior reassessment not a single antigen was found on PBMCs and only two antigens derived from HLA class II peptides were found once in normal tissue samples: COL1A1 and PAX2. However, as mentioned above, this can be due to HLA class II positive cells infiltrating normal kidney tissue after internalizing and processing antigens in adjacent cancerous tissue.

Nevertheless, HLA ligands representing well-established gene products involved in carcinogenesis and immunohistological
Table 2. Pathways associated with tumor or metastasis formation.

| Pathway or biological function | Number of malignant samples positive for pathway (frequency in %) | Best FDR for pathway | Best rank for pathway | Tumor-associated proteins involved in pathways |
|--------------------------------|---------------------------------------------------------------|----------------------|----------------------|-----------------------------------------------|
| HLA class I                    |                                                               |                      |                      |                                               |
| Extracellular matrix organization | 8 (72.7%)                                                      | 7.71E-22             | 1                    | COL1A1, F1N, LOX, MMP7, PLOD2                 |
| Cellular response to hypoxia   | 4 (36.4%)                                                      | 1.05E-34             | 3                    | CA9, EGLN3, VEGFA                            |
| Collagen catabolic process     | 4 (36.4%)                                                      | 4.14E-18             | 1                    | COL1A1, MMP7                                 |
| Platelet-derived growth factor binding | 4 (36.4%)                                                      | 3.2E-7              | 5                    | FN1, VEGFA                                   |
| Fibroblast growth factor receptor signaling pathway | 3 (27.3%)                                                      | 4.86E-9             | 3                    | COL1A1                                      |
| Epithelial growth factor receptor signaling pathway | 2 (27.3%)                                                      | 2.84E-6             | 2                    | BCAR1, EGFR, GRB2, ITPR3, TGFBI             |
| Growth factor binding          | 2 (27.3%)                                                      | 7.39E-5             | 9                    | COL1A1, IGFBP3                               |
| Positive regulation of cell motility | 3 (27.3%)                                                      | 1.43E-7             | 3                    | APC, ENPEP, BCAR1, EGFR, ENO1                |
| HLA class II                   |                                                               |                      |                      |                                               |
| Blood vessel development       | 8 (72.7%)                                                      | 2.96E-7             | 20                   | COL1A1, ENPEP, TGFBI, THBS1, VEGFA           |
| Growth factor binding          | 4 (36.4%)                                                      | 5.08E-5             | 19                   | CD36, COL1A1, NRP2, THBS1                   |
| Carbohydrate catabolic process | 3 (27.3%)                                                      | 1.12E-18            | 7                    | EN01                                         |
| Collagen catabolic process     | 3 (27.3%)                                                      | 1.85E-5             | 17                   | COL1A1                                      |
| Epithelial cell proliferation | 2 (27.3%)                                                      | 1.13E-3             | 28                   | PAX2, TGF1, THBS1, VEGFA, CALR, COL1A1, NRP2, PFN1, THBS1, VEGFA |
| Positive regulation of cell motility | 2 (27.3%)                                                      | 7.20E-7             | 17                   |                                             |
| Platelet-derived growth factor binding | 2 (27.3%)                                                      | 2.96E-6             | 21                   | THBS1, TGF1                                 |
| Epithelial to mesenchymal transition | 1 (9.1%)                                                      | 1.85E-2             | 26                   | THBS1, TGF1                                 |
| Mesenchymal to epithelial transition | 1 (9.1%)                                                      | 2.80E-2             | 14                   | PAX2                                         |

Pathways and biological functions showing significant enrichment in malignant tissue samples. Inclusion criteria for HLA class I are at least one FDR of < 0.01 in one sample and placement in the top 10 hits of at least one pathway analysis. For HLA class II FDR cut off of < 0.05 was chosen but placement in the top 30 hits for at least one analysis was mandatory.

Abbreviations: APC, adenomatous polyposis coli protein; BCAR1, breast cancer anti-estrogen resistance protein 1; CA9, carbonic anhydrase 9; CALR, Calreticulin; CD36, platelet glycoprotein 4; COL1A1, collagen α−1 (I) chain; EGFR, epidermal growth factor receptor; EGLN3, egl-9 homolog 3; EN01, Enolase 1; ENPEP, gliotamin aminopeptidase; FN1, fibronectin; GRB2, growth factor receptor-bound protein 2; IGFBP3, insulin-like growth factor-binding protein 3; ITPR3, inositol 1,4,5-triphosphate receptor type 3; LOX, protein-lysine 6-oxidase; MMP7, matrilysin; NRP2, neuropilin-2; PAX2, paired box gene Pax-2; PFN1, profilin-1; PLOD2, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2; TGFBI, transforming growth factor β−1; TGF2, transforming growth factor-β−related protein ig-h3; THBS1, thrombospondin-1; THBS2, thrombospondin-2; VEGFA, vascular endothelial growth factor A.

Table 3. Tumor-associated antigens and peptides derived from network analysis.

| Antigen | Peptide | HLA | Number of malignant samples positive for antigen (frequency in %) | Immune response to antigen published |
|---------|---------|-----|---------------------------------------------------------------|-------------------------------------|
| FN1     | RPAGVVTTL | B*07 | 7 (63.6%)                                                     | no                                  |
| IGFBP3  | RPTLWAAL | B*07, B*35 | 6 (54.5%)                                                   | yes⁴¹                                |
| COL1A1  | RLPIDVAPLVDGPAD | DR | 4 (36.4%)                                                     | no                                  |
| ENPEP   | AVHGDVSKRINSG | DR | 3 (27.3%)                                                     | no                                  |
| PAX2    | APGHTVIPSTASPPVSSA | DR | 3 (27.3%)                                                     | yes⁴⁴                                |
| TGFBI   | VSGIGALVRLKSLQGD | DR | 3 (27.3%)                                                     | no                                  |
| EGLN3   | NHQEVPSY | B*35 | 2 (18.2%)                                                     | yes⁴⁵                                |
| LOX     | QPRPILLI | B*07 | 2 (18.2%)                                                     | no                                  |
| PLOD2   | KVFAGYTYK | A*03 | 2 (18.2%)                                                     | No                                  |
| APC     | QIQEKDLI | B*08 | 1 (9.1%)                                                      | no                                  |
| BCAR1   | KPFDAPTPLVL | B*07 | 1 (9.1%)                                                      | no                                  |
| CA9     | SPPRAEPVQL | B*07 | 1 (9.1%)                                                      | yes⁴¹                                |
| CALR    | GGGYVKLPPNSLQDT | DR | 1 (9.1%)                                                      | no                                  |
| CD36    | VPYPVTITV | B*51 | 1 (9.1%)                                                      | no                                  |
| EGFR    | SPSTSRTPLL | DR | 1 (9.1%)                                                      | yes⁴¹                                |
| EN01    | TIAPALVSK | A*03 | 1 (9.1%)                                                      | yes⁴¹                                |
| GRB2    | QNWYKAEL | B*08 | 1 (9.1%)                                                      | no                                  |
| ITPR3   | DLIRTEL | B*08 | 1 (9.1%)                                                      | no                                  |
| MMP7    | YPFDDGPGNTL | B*35 | 1 (9.1%)                                                      | yes⁴¹                                |
| NRP2    | EGQYALISPPVH | DR | 1 (9.1%)                                                      | no                                  |
| PFN1    | VPGKTFTNITAEGVL | DR | 1 (9.1%)                                                      | no                                  |
| TGFBI   | TPLERAQHL | B*08 | 1 (9.1%)                                                      | no                                  |
| THBS1   | SPAFRIEANLIPP | DR | 1 (9.1%)                                                      | no                                  |
| THBS2   | STGILITALEGGLSQDR | DR | 1 (9.1%)                                                      | no                                  |
| VEGFA   | APAARAPQQL | B*07 | 1 (9.1%)                                                      | yes⁴¹                                |

HLA class I and II antigens and corresponding peptides derived from renal cell carcinoma and their metastases found in significantly enriched pathways in malignant tissue samples. Abbreviations: see Table 2.

characterization of RCC were indeed detected, such as EGFR, VEGFA, IGFBP3, EGLN3, CA9, PAX2, and MMP7.20,21,30-35 Furthermore, a vast array of new TAAs with corresponding peptides from HLA class I and II were discovered (Table 3). These included LOX and PLOD2, with high potential in engaging metastasis,36-38 GRB2, ITPR3, and BCAR1, all of which are part of the EGFR pathway39 and other growth factor receptor pathways, plus ENPEP, NRP2, TGFBI, THBS1, and THBS2, which are essential for angiogenesis.40

It is notable that 16 out of 24 peptides listed in Table 3 were only found in one tissue sample each. Network analysis suggests that other proteins involved in the pathway could also be active in the same tissue sample but were not detected due to lack of sensitivity or the selectivity of antigen processing. For example, network analysis of sample RCC297M2 (Fig. 5C) showed enrichment for the extracellular matrix organization process and used (as described above) genes not included in the input list to complete the network. These genes, which are not striped in the visualization of the network, incorporated LOX, for example, which was detected later in samples RCC377M and RCC377T. Similar events were observed for BCAR1, CA9, COL1A1, EGFR, ENPEP, FN1, NRP2, PAX2, PLOD2, TGFBI, THBS1, THBS2, and VEGFA.

Additionally, to determine the immunogenicity of these potential antigens and their corresponding peptides, we used different in silico methods to predict the immunogenicity.

Second, we used two in silico methods to predict the immunogenicity.
of tumor-specific proteins and peptides identified by our approach: the VaxiJen server for proteins and the IEDB immunogenicity prediction tool for HLA class I nonamer-peptides. To improve comparability of our results 53 immunogenic tumor-specific peptides (39 for HLA class I and 14 for HLA class II) derived from 46 antigens (34 for HLA class I and 12 for HLA class II) were defined as positive controls in contrast to our experimental results. All of these HLA ligands have been published by our lab and have been identified in tissue samples of patients with multiple myeloma, chronic lymphatic leukemia, acute myeloid leukemia, ovarian cancer, and RCC (Table S3).9,46–49 T cell activation in these experiments was validated by ELISpot or intracellular cytokine staining.

The analysis of source proteins by the VaxiJen server revealed that 83.8% of our experimental data were classified as antigens in the tumor prediction model which was even higher than the results for the positive controls with 71.7% of all source proteins reaching the antigen threshold. For the IEDB T-cell immunogenicity prediction tool nonamer-peptides of our data were predicted to be immunogenic in 37.5% of all cases when choosing an arbitrary threshold of 0.1 for T-cell recognition score. In contrast, 22.2% of positive controls reached this threshold, but again the ratio of positive hits of our experimental data did exceed results for well-established immunogenic HLA ligands (Table S3). Additionally the mean T cell recognition score for all potential T cell epitopes was 0.088 which meets the results for a group of immunogenic peptides (Table S3).9,46–49 A possible explanation for this disparity might be the selection of suitable peptides for vaccination design. Therefore, new algorithms for vaccine design are needed, taking into account tumor- and metastasis-specific HLA-peptide presentation. Additionally, personalized approaches should be given priority: these can lead to individualized vaccine cocktails that encompass small but distinct differences in the biological characteristics of carcinomas due to intra- and inter-tumoral heterogeneity of tumor sub-clones.14,15

To achieve this, our primary aim was to develop a new approach using network analysis of source proteins that represent HLA-restricted peptides, thereby enabling the identification of biological processes and pathways enriched in and exclusively presented on ccRCC and their metastasis. Subsequently, these network analyses should provide the basis for defining new TAAs and peptide vaccines.

Our study was based on analysis of HLA ligands by LC-MS/MS isolated from 17 normal, tumorous and metastatic kidney tissues and 38 blood plus 10 bone marrow samples from healthy donors with over 1,500 unique, source proteins for tissue samples and over 8,500 antigens for healthy blood donors, respectively. Thereby HLA class I analysis of RCC tissue samples yielded considerably more HLA ligand and source proteins identifications compared to HLA class II. This could be explained by lower efficiency of HLA class II peptide isolation or detection by mass spectrometry since results for PBMCs and BMMCs did show similar differences in detection of HLA class I or II peptides (Table S1 and 2). On the other hand, the lower expression of HLA class II molecules compared to HLA class I. Subtracting PBMC-derived peptides from results of tissue analysis allowed for a focused analysis of kidney-associated antigens in malignant versus normal kidney tissue. In performing analysis for distribution of HLA ligands between tumors, metastases, and normal kidney tissues in contrast to previous research which was performed mostly on RNA but partly on peptide level as well,25 no major differences could be detected for these tissue subtypes. This discrepancy is probably due to the higher sensitivity in detection of HLA ligands in our study and underscores the advancement of mass spectrometry based HLA ligandome analysis over the last years. In our earlier studies on the HLA ligandome of RCC, the effectiveness on chromatographic separation was limited by the relatively slow scan rates of time-of-flight mass spectrometers. Therefore, a lot of peptides were eluted from the separation column without detection by mass spectrometry. The implementation of orbitrap hybrid mass spectrometers for HLA peptidomics resulted in vastly improved identification rates due to high mass accuracy and resolution, improved sensitivity and speed of peptide detection.

**Discussion**

In recent years, specific and unspecific immunotherapies have become increasingly important in cancer therapy. Checkpoint inhibitors targeting programmed-death 1 (PD-1), programmed death-ligand 1 (PD-L1), or cytotoxic T-lymphocyte-associated antigen 4 (CTLA-4) in particular have already been established for treatment of RCC and other immunogenic malignancies.10–12 On the other hand, peptide vaccination strategies, as a specific immunotherapy, were able to elicit a peptide specific T cell response and prolong patient survival but not as sustainable as shown for checkpoint blockade.9,10,41 A possible explanation for this disparity might be the selection of suitable peptides for vaccination design. Therefore, new algorithms for vaccine design are needed, taking into account tumor- and metastasis-specific HLA-peptide presentation. Additionally, personalized approaches should be given priority: these can lead to individualized vaccine cocktails that encompass small but distinct differences in the biological characteristics of carcinomas due to intra- and inter-tumoral heterogeneity of tumor sub-clones.14,15

In summary, for each patient, independent of the number of tissue samples available, we were able to identify 6 to 11 pathways or biological processes found almost exclusively on malignant tissue samples. This enabled detection of 15 to 20 well established or newly identified TAAs and a similar number of representing HLA class I and II peptides, of which three to six are already known to elicit T cell responses in vitro or in vivo (Tables 3 and 4), whereas other HLA ligands according to prediction models are very likely to do so as well. This novel approach can directly be implemented to inform the generation of individual peptide vaccines (Table S4) targeting crucial biological functions in carcinogenesis.

**Table 4.** Quantity of tumor-associated pathways and proteins found in RCC samples.

| Samples  | RCC70 | RCC210 | RCC297 | RCC329 | RCC377 | RCC399 |
|----------|-------|--------|--------|--------|--------|--------|
| TA pathways HLA class I | 6     | 4      | 4      | 2      | 6      | 4      |
| TA pathways HLA class II | 2     | 2      | 7      | 5      | 1      | 7      |
| TA pathways combined | 8     | 6      | 11     | 7      | 7      | 11     |
| TAA HLA class I | 12    | 10     | 8      | 7      | 14     | 11     |
| TAA HLA class II | 6     | 6      | 9      | 8      | 1      | 9      |
| TAA combined | 18    | 16     | 17     | 15     | 15     | 20     |

Number of identified tumor-associated pathways and corresponding tumor-associated antigens in renal cell carcinoma samples.

**Abbreviations:** TA, tumor associated; TAA, tumor-associated antigen; RCC, renal cell carcinoma.

---

9. [Calis et al.](#)
10. [CTLA-4](#)
11. [PD-L1](#)
12. [Checkpoint Blockade](#)
If not for HLA ligands per se, we were nevertheless able to show high functional similarity of tumorous and metastatic compared to normal and metastatic and especially tumorous and normal kidney tissue by network analysis for HLA class I using HLA ligand-derived source proteins. Analysis for HLA class II revealed no differences in pathway distribution between the three tissue modalities which can be explained by a high amount of broadly used extracellular structure proteins all present in different tissue samples. Otherwise as mentioned before the absence of clear differences in pathway analysis for HLA class II might be due to a biased analysis of HLA class II peptides by detection of only abundant proteins. In light of recent findings, implementation of multiple fragmentation strategies may help overcome this problem in future studies.

Nevertheless, for both HLA class I and II we identified several tumor-associated enriched pathways conserved in malignant tissue samples. This shows for the first time that key proteins of intra- and extra-cellular biological processes of carcinogenesis are reflected in HLA ligands and can be used for targeting carcinomas in a more functional manner.

The two networks with the lowest FDR detected exclusively in malignant tissue samples include first the main factor of carcinogenesis in ccRCC which is cellular response to hypoxia, and second extracellular matrix organization which most probably encodes for the process of metastasis considering the high functional similarity of tumorous and metastatic compared to normal tissue samples. This shows for the general tumor-associated enriched pathways conserved in malignancy.

The two networks with the lowest FDR detected exclusively in malignant tissue samples include first the main factor of carcinogenesis in ccRCC which is cellular response to hypoxia, and second extracellular matrix organization which most probably encodes for the process of metastasis considering the high functional similarity of tumorous and metastatic compared to normal tissue samples. This shows for the general tumor-associated enriched pathways conserved in malignancy.

Additionally, multiple growth factor receptor pathways, e.g., epidermal growth factor receptor (EGFR), fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptor (PDGFR), and growth factor receptor pathway, were also found to be overrepresented. With angiogenesis being among the enriched processes, the vascular endothelial growth factor receptor (VEGFR) pathway likewise indirectly plays a role in the HLA ligand-dependent functional description of RCC.

All of these receptor pathways with the exception of the EGFR and growth factor receptor pathway are well-established targets of tyrosine kinase inhibitors (TKIs) in the first and second line treatment of metastatic RCC. This suggests the use of a twofold approach that targets these pathways by peptide vaccination in combination with established TKIs and other anti-angiogenic agents.

Finally, by reviewing the biological processes and their source proteins mentioned above, we were able to define new TAAs first, by their involvement in tumor specific, enriched networks and second, by their HLA-restricted presentation on malignant tissue samples. With only one or two malignant tissue samples of small size for each of the six patients, using this approach we were able to engineer an individual peptide vaccine for every patient that targets essential biological tumor functions with at least 15 peptides from HLA class I and II, of which three to six peptides are already known to be immunogenic in vivo. For those proteins, which are to this point not published to be immunogenic, in silico prediction models highly suggest their immunogenicity especially by comparison to well-established T-cell epitopes derived from different cancer entities. This is also corroborated by our previous studies which showed a high correlation of tumor association on the HLA ligandome level and immunogenicity in patients with chronic lymphatic leukemia. In light of the clear in silico predictions of immunogenicity and the high cancer restriction of the identified HLA ligands, no additional T cell assays were carried out. Notably, predictions are in line with our previous studies, which demonstrated that non-mutated ligands, though derived from ubiquitous source proteins, are able to elicit T cell responses in cancer patients and therefore able to overcome central and peripheral tolerance if they are not detected on healthy PBMCs. However, since many of the identified proteins listed in Table 3 are known to be broadly expressed among normal tissues and immunogenicity has not been experimentally proven in this study, it is mandatory that all HLA ligands corresponding to TAAs identified in this study have to be analyzed by in vitro T cell assays to determine their immunogenicity before they can be used as basis for clinical peptide vaccination studies.

In conclusion, this study demonstrates for the first time that carcinogenesis for RCC is reflected in HLA ligands and can be used via a network analysis-based approach to inform the design of individualized peptide vaccines for renal cell carcinoma. By representing a specific immunotherapy that targets distinct biological processes enriched in carcinogenesis, our strategy emphasizes its potential for combined treatment with TKIs or other antiangiogenic small molecules as well as checkpoint inhibitors for synergistic effects.

Materials and methods

Patients, tissue, and blood samples

For analysis of HLA ligands, tissue samples from RCC patients were provided by the Department of Urology, Tuebingen, Germany, after written informed consent had been obtained according to the Declaration of Helsinki protocol. We received tissue samples from primary clear cell RCC, from corresponding normal kidney, and from synchronous or metachronous metastases of different anatomical sites (Patient characteristics are shown in Table 1). Due to rareness of these triplets only four complete sample sets could be provided (RCC70, RCC329, RCC377, RCC399). In one additional case (RCC210), assumed metastasis had been identified as adrenal adenoma. From the sixth patient (RCC297) only two metastases were provided without primary tumor or healthy kidney tissue. The two incomplete sets were nevertheless included in the study to validate if the results are transferable to other sample sets as well. Specimens were frozen in liquid nitrogen immediately after resection and stored at −80°C until further use. For detecting HLA-restricted peptides from PBMCs of healthy blood donors, cells were isolated by density gradient centrifugation. Informed consent had also been obtained in accordance with Declaration of Helsinki protocol and cells were stored at −80°C until further use. Histological and TNM classification were performed by the Department of Pathology, Tuebingen, Germany. HLA typing was carried out by the Department of Hematology and Oncology, Tuebingen, Germany. This study has been approved by the local ethics committee.

Isolation of HLA ligands from primary samples

Tissue and PBMC samples were prepared as described by Weinzierl et al. and peptides bound to HLA class I and II molecules were subsequently isolated using standard immunoaffinity purification as published elsewhere using the pan-HLA class I specific monoclonal antibody W6/32 and the pan-HLA class II specific monoclonal antibody Tü39, respectively. Samples were stored at −20°C until analysis by LC-MS/MS.
Analysis of HLA ligands by LC-MS/MS

Due to optimization procedures, analysis was performed with two different settings. Initial parameters for tissue samples (primary RCC, normal tissue, and metastases) are specified first, followed by settings for blood samples in squared brackets: peptide samples were separated by reversed-phase liquid chromatography (nano-UHPLC, UltiMate 3000 RSLCnano, Dionex) and subsequently analyzed in an on-line coupled LTQ Orbitrap XL hybrid mass spectrometer (ThermoFisher). Samples were analyzed in five technical replicates. Sample volumes of 5 μL (sample shares of 20%) were injected onto a 300 μm × 1 cm [75 μm × 2 cm] trapping column (Acclaim PepMap RSLC, Dionex) at 4 μL/min for 5.75 min. Peptide separation was subsequently performed at 50°C and a flow rate of 300 nL/min [175 nL/min] on a 75 μm × 25 cm [50 μm × 50 cm] separation column (Acclaim PepMap RSLC, Dionex) applying a gradient ranging from 2.0 to 44.9% [2.4 to 32.0%] of acetonitrile (ACN) over the course of 120 min [140 min]. Eluting peptides were ionized by nanospray ionization and analyzed in the mass spectrometer implementing a top five CID (collision-induced dissociation) method generating fragment spectra for the five most abundant precursor ions in the survey scans. Resolution was set to 60,000. For HLA class I ligands the mass range was limited to 400–800 m/z [400–650 m/z] with charge states 2+ and 3+ selected for fragmentation. For HLA class II, a mass range of 400–1,200 m/z [300–1,500 m/z] was analyzed with charge states ≥ 2 selected for fragmentation.

Database search and spectral annotation

For data processing, the software Proteome Discoverer (v1.3, ThermoFisher) was used to integrate the search results of the Mascot search engine (Mascot 2.2.04, Matrix Science) against the human proteome as comprised in the Swiss-Prot database (www.uniprot.org, release: September 27th, 2013; 20,279 reviewed protein sequences contained). The search combined data of technical replicates and was not restricted by enzymatic specificity. Precursor mass tolerance was set to 5 ppm, and fragment mass tolerance was set to 0.5 Da. Oxidized methionine was allowed as a dynamic modification. False discovery rates (FDR) were determined by the Percolator algorithm60 based on processing against a decoy database consisting of the shuffled target database. FDR was set at a target value of q ≤ 0.05 (5% FDR). Peptide-spectrum matches (PSMs) with B ≤ 0.05 were filtered according to additional, orthogonal parameters to ensure spectral quality and validity. Mascot scores were filtered to ≥20. For HLA class I, peptide lengths were limited to 8–12 aa. For HLA class II peptides, these were limited to 10–25 aa of length. Protein grouping was disabled, allowing for multiple annotations of peptides (e.g., conserved sequences mapping into multiple proteins). To improve sensitivity of peptide analysis for tissue samples considering small sample size no FDR target value or Mascot score cut off was set resulting in a higher yield of annotated peptide spectra. To ensure validity of these results, peptides were matched either with a large in-house database consisting of tumorous and normal RCC tissue sample and PBMC derived peptides from independent experiments fulfilling initially mentioned criteria (5% FDR, PSM with B ≤ 0.05, and mascot score ≥20) or with HLA ligands already published in peer reviewed journals.

For quality control, yield thresholds of ≥400 unique HLA class I ligands and ≥100 unique HLA class II ligands per triplet (consisting of tumorous, normal, and metastatic tissue) were applied. HLA annotation was performed using SYFPEITHI (www.syfpeithi.de) or an extended in-house database.

Pathway analysis of proteins derived from reassessed tissue samples

Pathway analyses of tissue samples were carried out with the online resource GeneMania (http://www.genemania.org). Input lists were generated as follows: HLA ligands of tissue samples and healthy blood donor PBMCs were compared and overlapping results discarded resulting in “reassessed,” more kidney-specific samples. Unique entry identifier from the uniprot database corresponding to source proteins of these HLA ligands was then used to create final input lists for each tissue sample separately.

Pathway analysis was performed with all available networks and four possible weighting methods (automatically, biological process based, molecular function based, and cellular component based) at the GeneMania platform. Additionally 20 related genes and 10 attributes were displayed according to GeneMania standard settings. Results of all four analyses were grouped for each tissue sample, ranked by FDR, and duplicates were singularized by selecting the one with the best FDR.

In silico prediction of immunogenicity of source proteins and HLA class I nonamer-peptides

For prediction of immunogenicity two different tools were used. Vaxijen is an online available software tool which uses an alignment-free approach for antigen prediction (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html).62 We chose the tumor prediction model with a threshold of 0.45 with slightly higher sensitivity than the pre-set threshold of 0.5. Every potential antigen was checked and the prediction value calculated. According to the threshold source proteins were predicted to be an “antigen” or “non-antigen.” Immune responses to HLA class I ligands were evaluated by the IEDB immunogenicity prediction tool (http://tools.iedb.org/immunogenicity/).50 Non-antigen-peptides were tested exclusively according to the implementation of this tool. For potential T-cell epitopes, the T-cell recognition score was determined and a threshold of 0.1 was set to decide if HLA-ligands were predicted to be “immunogenic” or “non-immunogenic.” Additionally, the mean T-cell recognition score was calculated as published by Calis et al. to determine how the mean score of a whole group of HLA ligands relates to established mean T cell recognition scores of established T cell epitopes.

Software and statistical analysis

Microcal Origin 9.0 software (OriginLab, Northampton, MA) was used for statistical analysis. Overlap analysis was performed using the BioVenn online tool (http://www.cmbi.ru.nl/cdd/biovenn).
Disclosure of potential conflicts of interest

Hans-Georg Rammmensee declares to be shareholder of immatics biotechnologies, Tübingen, and CureVac GmbH, Tübingen. The other authors disclose no potential conflicts of interest. The authors have no relevant financial relationship to disclose.

Acknowledgments

We thank Patricia Hrstic, Beate Pömmler, Claudia Falkenburger, Katharina Graf, and Nicole Zuschke for providing excellent technical support and Mathias Walzer for expert programming of in-house analytical scripts.

Funding

This work was supported by the Deutsche Forschungsgesellschaft (DFG, SFB 685 and GRK 794) and the German Cancer Consortium (DKTK).

References

1. Everson TC. Spontaneous regression of cancer. Ann N Y Acad Sci 1964; 114:721-35; PMID:14131914; http://dx.doi.org/10.1111/j.1749-6632.1964.tb00991.x
2. Spontaneous regression of cancer. Lancet 1966; 2:896; PMID:4162303; http://dx.doi.org/10.1016/S0140-6736(66)91989-1
3. Lokich J. Spontaneous regression of metastatic renal cancer - Case report and literature review. Am J Clin Oncol-Canc 1997; 20:416-8; PMID:9256902; http://dx.doi.org/10.1007/978-007-00020-0
4. Hughes T, Klairmont M, Broucek J, Iodice G, Basu S, Kaufman HL. The prognostic significance of stable disease following high-dose interleukin-2 (IL-2) treatment in patients with metastatic melanoma and renal cell carcinoma. Cancer Immunol Immunother 2015; 64:459-65; PMID:25603775; http://dx.doi.org/10.1007/s00262-014-1652-6
5. Joachim D, Richter A, Hoffmann L, Iwig K, Fahlenkamp D, Zakrewski G, Schmitt E, Dannenberg T, Lehmanner W, von Wietersheim J et al. Adjuvant autologous renal tumour cell vaccine and risk of tumour progression in patients with renal-cell carcinoma after radical nephrectomy: phase III, randomised controlled trial. Lancet 2004; 363:594-9; PMID:14987883; http://dx.doi.org/10.1016/S0140-6736(04)15590-6
6. Amato RJ, Hawkins RE, Kaufman HL, Thompson JA, Tomczak P, Szczylik C, McDonald M, Easly S, Shinger WH, de Belin J et al. Vaccination of metastatic renal cancer patients with MVA-ST4: a randomized, double-blind, placebo-controlled phase III study. Clin Cancer Res 2010; 16:5539-47; PMID:20881001; http://dx.doi.org/10.1158/1078-0432.CCR-09-3209
7. Amin A, Dudek AZ, Logan TF, Lance RS, Holzbeierlein JM, Knox J, Master VA, Pal SK, Miller WH, Jr, Karsh LI et al. Survival with AGS-003, an autologous dendritic cell-based immunotherapy, in combination with sunitinib in unfavorable risk patients with advanced renal cell carcinoma (RCC): Phase 2 study results. J Immunother Cancer 2015; 3:14; PMID:25901286; http://dx.doi.org/10.1186/s40425-015-0055-3
8. Combe P, de Guillebon E, Thibault C, Granier C, Tartour E, Oudard S, Amato RJ, Hawkins RE, Kaufman HL, Thompson JA, Tomczak P, Szczylik C, Stenzl A, Dzio R, Iranzo A, Procopio L et al. Nivolumab plus ipilimumab in advanced melanoma. N Engl J Med 2013; 369:122-33; PMID:23724867; http://dx.doi.org/10.1056/NEJMoa1302369
9. Marusyk A, Tabassum DP, Altrick PM, Almedro V, Michor F, Polak Y. Non-cell-autonomous driving of tumour growth supports subclonal heterogeneity. Nature 2014; 514:54-8; PMID:25079331; http://dx.doi.org/10.1038/nature13556
10. Schumacher T, Busse L, Pusch S, Sahn F, Wiestler B, Quandt J, Menn O, Osuald M, Oen I, Ott M et al. A vaccine targeting mutant IDH1 induces antitumour immunity. Nature 2014; 512:324-7; PMID:25043408; http://dx.doi.org/10.1038/nature13837
11. Wolchok JD, Kluger H, Callahan MK, Postow MA, Rizvi NA, Lesokhin AM, Segal NH, Ariyan CE, Gordon RA, Reed K et al. Nivolumab...
58. Falk K, Rotzschke O, Stevanovic S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. Nature 1991; 351:290-6; PMID:1709722; http://dx.doi.org/10.1038/351290a0

59. Kowalewski DJ, Stevanovic S. Biochemical large-scale identification of MHC class I ligands. Methods Mol Biol 2013; 960:145-57; PMID:23329485; http://dx.doi.org/10.1007/978-1-62703-218-6_12

60. Kall L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods 2007; 4:923-5; PMID:17952086; http://dx.doi.org/10.1038/nmeth1113

61. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. Nucleic Acids Res 2010; 38:W214-20; PMID:20576703; http://dx.doi.org/10.1093/nar/gkq537

62. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007; 8:4; PMID:17207271; http://dx.doi.org/10.1186/1471-2105-8-4