Lu *et al*. *Hepatoma Research* 2018;4:21
DOI: 10.20517/2394-5079.2018.44

Hepatoma Research

Review

Open Access

Pathway analysis provides insight into the genetic susceptibility to hepatocellular carcinoma and insight into immuno-therapy treatment response

Yih-Kuang Lu[1], Jacob Morris Brill[1], Ardesher Aghili[1], Kenneth Howard Buetow[1,2,3]

[1]Computational Sciences and Informatics, Complex Adaptive Systems, Arizona State University, Tempe, AZ 85287, USA.
[2]Center for Evolution and Medicine, Arizona State University, Tempe, AZ 85287, USA.
[3]School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA.

Correspondence to: Dr. Kenneth Howard Buetow, Computational Sciences and Informatics, Complex Adaptive Systems, Arizona State University, Tempe, AZ 85287, USA. E-mail: kenneth.buetow@asu.edu

How to cite this article: Lu YK, Brill JM, Aghili A, Buetow KH. Pathway analysis provides insight into the genetic susceptibility to hepatocellular carcinoma and insight into immuno-therapy treatment response. *Hepatoma Res* 2018;4:21. http://dx.doi.org/10.20517/2394-5079.2018.44

Received: 26 Apr 2018 **First Decision:** 28 May 2018 **Revised:** 11 Jun 2018 **Accepted:** 12 Jun 2018 **Published:** 22 Jun 2018

Science Editor: Guang-Wen Cao **Copy Editor:** Jun-Yao Li **Production Editor:** Huan-Liang Wu

Abstract

Clear evidence exists for genetic susceptibility to hepatocellular carcinoma (HCC). Genome-wide association studies have identified multiple candidate susceptibility loci. These loci suggest that genetic variation in the immune system may underpin HCC susceptibility. Genes for the antigen processing and presentation pathway have been observed to be significantly enriched across studies and the pathway is identified directly through genome-wide studies of variation using pathway methods. Detailed analysis of the pathway indicates both variation in the antigen presenting loci and in the antigen processing are different in cases in controls. Pathway analysis at the transcriptional level also shows difference between normal liver and liver in individuals with HCC. Assessing differences in the pathway may prove important in improving immune therapy for HCC and in identifying responders for immune checkpoint therapy.

Keywords: Hepatocellular carcinoma, genetic susceptibility, genome-wide association study, pathway analysis, antigen presentation and processing, immune checkpoint therapy

INTRODUCTION

Hepatocellular carcinoma (HCC), the most common form of primary liver cancer, is ranked 5th in global incidence and 2nd in mortality[1]. With the exception of East Asia, the incidence of HCC is increasing in almost all regions of the world and has doubled in the USA since the early 1980s[2]. This increase is attributable to increases in obesity and type II diabetes[3,4]. Liver cancer's 5-year survival is the second worst among all cancers (18.1%)[5].

In this manuscript, the role of genetic susceptibility to HCC is examined. Novel tools that evaluate genetic data using collections of genes and their interactions within biologic networks are used to identify key biologic processes driving susceptibility. The relationship of germline and somatic variation is explored. The importance of these findings is assessed in the context of current therapeutic interventions for HCC.

SOMATIC GENETIC ETIOLOGY OF HCC

Like other solid tumors, at a somatic level, HCC appears to arise via alterations in numerous genes that modify multiple biologic processes. An early whole-genome sequencing effort identified an average of 9718 nucleotide alternations, 271 insertion/deletions, and 41 structural variations per tumor, with substantial variability from tumor to tumor[6]. Within coding sequences, it has been reported that there are an average of 21 synonymous and 64 non-synonymous mutations per tumor[7]. Tumors of larger size are observed to have greater numbers of point mutations, which are speculated to contribute to heterogeneity within the tumors. The Cancer Genome Atlas (TCGA) Research network's evaluation of HCC[8] finds alterations over-represented in the RAS pathway, WNT pathway, cell cycle regulation pathways and chromatin modification pathways with high mutation rates in TP53 (31%), CTNNB1 (27%), AXIN1 (8%), ARID1A (7%), ARID2 (5%), RB1 (4%), PIK3CA (4%), CDKN2A (2%), KRAS (1%), NRAS (1%), high deletion frequencies of RB1 (19%), CDKN2A (13%), PTEN (7%) and amplification of CCND1 (6%). The most commonly mutated locus was TERT with promoter mutations found in 44% of tumors[8]. The TCGA data unexpectedly also showed high mutation rates in ALB (13%) and APOB (10%).

GENETIC SUSCEPTIBILITY TO HCC

In contrast to other common tumors, genetic susceptibility to HCC remains poorly characterized. Studies have identified evidence for familiality of HCC, over and above familial exposures such as HBV infection[9-14]. For example, after accounting for HBV infection, individuals with a family history of HCC have a rate ratio of 2.4[10]. To date, these studies have examined only hepatitis virus associated HCC and have yet to explore the role of obesity and diabetes related susceptibility.

A limited number of studies have been conducted to identify the loci underpinning this familiality. Original studies focused on candidate genes whose observed single nucleotide polymorphisms (SNPs) could plausibly modify known environmental risk factors for HCC including aflatoxin, alcohol, or tobacco. A meta-analysis of these studies found associations with 5 genes *HFE*, *IL-1B, MnSOD, MDM,* and *2UGT1A7*[15].

HCC has had a small number of genome wide association studies (GWAS) conducted with modest success in identifying risk loci. The NHGRI-EBI Catalog lists a total of 11 studies that have identified 22 loci[16]. These studies examine East Asian populations and have included HCC associated with hepatitis B virus (HBV), hepatitis C virus (HCV), and non-alcoholic steatohepatitis (NASH) etiologies. The studies have identified SNPs in the genomic proximity (intronic, upstream and/or downstream) of twenty protein coding loci.

Clues to the biologic basis of HCC susceptibility across GWAS studies can be identified by looking for non-random enrichment. Using the resources of the Gene Ontology consortium (GO) (http://geneontology.org), the twenty protein coding loci were examined for biologic process enrichment in Homo sapiens. This enrichment analysis uses the tools of Panther (http://pantherdb.org/webservices/go/overrep.jsp). Four high level GO processes were observed to be significantly enriched "T cell receptor signaling pathway" ($P = 0.0366$), "interferon-gamma-mediated signaling pathway" ($P = 0.0026$), "T cell costimulation" ($P = 0.0020$), and "antigen processing and presentation of exogenous peptide antigen via MHC class II" ($P = 0.0001$).

We have previously looked for inherited susceptibility using genome-wide genotyping and a novel analytic approach that uses biologic networks - Pathways of Distinction Analysis (PoDA)[17]. In PoDA, the network is

Table 1. Updated significant networks identified through pathway of distinction analysis

PoDA pathway name	Source	DS	OR	No. of genes	No. of SNPs
Axon guidance	KEGG	1.888	3.1699	245	13,044
GPCR downstream signaling	REACTOME	1.706	2.4122	695	16,949
Focal adhesion	KEGG	0.802	2.3329	197	7999
Pathways in cancer	KEGG	0.570	2.2487	284	10,406
MAPK signaling pathway	KEGG	0.620	2.1152	245	7368
PI3K-Akt signaling pathway	KEGG	-0.339	2.0837	314	10,409
Calcium signaling pathway	KEGG	-1.030	1.8479	163	8684
Regulation of actin cytoskeleton	KEGG	-1.004	1.8207	195	5681
Glycerolipid metabolism	KEGG	2.003	1.7607	55	1590
Mechanism of gene regulation by peroxisome proliferators via ppara	BIOCARTA	2.371	1.7272	49	1076
Interleukin-3, 5 and GM-CSF signaling	REACTOME	2.969	1.7235	41	1188
Glycerophospholipid biosynthesis	REACTOME	2.201	1.7208	70	1714
T cell receptor signaling pathway	BIOCARTA	2.493	1.6792	55	1500
Dopaminergic synapse	KEGG	-1.348	1.6651	116	5396
Stabilization and expansion of the E-cadherin adherens junction	NCI/NATURE	2.142	1.6630	40	1449
Eicosanoid metabolism	BIOCARTA	3.026	1.6620	16	800
Netrin-mediated signaling events	NCI/NATURE	1.965	1.6620	28	2400
Pre-NOTCH expression and processing	REACTOME	3.240	1.6343	45	1451
Purine metabolism	KEGG	-1.190	1.6284	150	4726
Toxoplasmosis	KEGG	2.470	1.5901	110	2088
Angiopoietin receptor Tie2-mediated signaling	NCI/NATURE	2.163	1.5806	47	1331
Circadian entrainment	KEGG	-1.498	1.5738	88	5919
Systemic lupus erythematosus	KEGG	3.873	1.5688	82	1185
Bioactive peptide induced signaling pathway	BIOCARTA	2.276	1.5677	42	1260
Role of mef2d in t-cell apoptosis	BIOCARTA	2.138	1.5522	30	946
Herpes simplex infection	KEGG	2.816	1.5388	170	1994
Glycosphingolipid biosynthesis - lacto and neolacto series	KEGG	2.756	1.5285	23	535
Multi-step regulation of transcription by pitx2	BIOCARTA	2.935	1.5253	22	526
Retrograde endocannabinoid signaling	KEGG	-1.990	1.5208	94	4960
TCR signaling	REACTOME	3.001	1.4913	51	1226
TPO signaling pathway	BIOCARTA	2.556	1.4896	23	635
Growth hormone signaling pathway	BIOCARTA	2.144	1.4813	28	768
Rheumatoid arthritis	KEGG	2.895	1.4801	84	978
Huntington's disease	KEGG	2.156	1.4658	152	1647
Inactivation of gsk3 by akt causes accumulation of b-catenin in alveolar macrophages	BIOCARTA	2.467	1.4628	32	709
Chaperones modulate interferon signaling pathway	BIOCARTA	2.486	1.4615	18	313
Phospholipase c signaling pathway	BIOCARTA	2.886	1.4577	10	849
GnRH signaling pathway	KEGG	-1.259	1.4488	84	3622
Oocyte meiosis	KEGG	-1.285	1.4371	102	2727
Biosynthesis of unsaturated fatty acids	KEGG	1.927	1.4342	19	495
GMCSF-mediated signaling events	NCI/NATURE	1.843	1.4339	30	841
p75 NTR receptor-mediated signalling	REACTOME	-1.195	1.4335	76	2466
E-cadherin signaling in keratinocytes	NCI/NATURE	2.123	1.4326	21	477
Signaling events mediated by HDAC Class III	NCI/NATURE	1.927	1.4323	26	565
Keratan sulfate/keratin metabolism	REACTOME	1.962	1.4251	28	447
Morphine addiction	KEGG	-3.158	1.4243	86	4524
IL3-mediated signaling events	NCI/NATURE	2.199	1.4233	22	399
Intestinal immune network for IgA production	KEGG	3.740	1.4224	45	506
lectin induced complement pathway	BIOCARTA	2.541	1.4167	11	359
Leishmaniasis	KEGG	2.734	1.4130	68	927
Alternative complement pathway	BIOCARTA	2.196	1.4075	11	236
Autoimmune thyroid disease	KEGG	2.835	1.4056	39	513
Graft-versus-host disease	KEGG	3.079	1.4025	33	240
Activation of pkc through g-protein coupled receptors	BIOCARTA	1.733	1.3968	11	892
Allograft rejection	KEGG	3.327	1.3952	30	253
Costimulation by the CD28 family	REACTOME	2.648	1.3931	62	1270
Eicosanoid ligand-binding receptors	REACTOME	2.798	1.3899	11	174

Staphylococcus aureus infection	KEGG	3.410	1.3766	52	504
Serotonergic synapse	KEGG	-1.854	1.3733	73	3128
N-glycan antennae elongation in the medial/trans-Golgi	REACTOME	2.122	1.3729	14	396
Integrins in angiogenesis	NCI/NATURE	-1.929	1.3622	74	2110
Tandem pore domain potassium channels	REACTOME	2.299	1.3589	4	206
Fatty acid elongation in mitochondria	REACTOME	2.226	1.3579	12	170
IL5-mediated signaling events	NCI/NATURE	2.080	1.3568	12	304
Antigen processing and presentation	KEGG	3.506	1.3397	65	400
Asthma	KEGG	3.713	1.3246	31	200
Neurotransmitter release cycle	REACTOME	1.846	1.3233	9	326
Classical complement pathway	BIOCARTA	2.682	1.3164	12	239
Antigen processing and presentation	BIOCARTA	2.938	1.2857	9	52
Interferon gamma signaling	REACTOME	3.080	1.0558	61	2598
Antigen processing-cross presentation	REACTOME	2.187	1.0371	59	1962

the unit of analysis and accounts for interactions among features within the network. In this analysis "antigen processing and presentation" was identified as having significant differences in variability in a population of Korean HBV associate HCC cases and controls. Consistent with the results of the enrichment analysis, re-analysis of this dataset with an extended set of 1200 pathways again identified "antigen processing and presentation", but also "interferon gamma signaling", "TCR signaling", and "T cell receptor signaling pathway" [Table 1] suggesting that immune response may be a key driver of HCC susceptibility.

THE ROLE OF ANTIGEN PROCESSING AND PRESENTATION IN HCC

To assess what might be the key factors within "antigen processing and presentation", we performed analysis utilizing a modified version of PoDA using the Korean HCC dataset. In this analysis, all 400 of the SNPs genotyped in the data set for the 65 genes in the pathway were contrasted in the cases and controls. After assessing significance of the odds ratio for the entire set of SNPs, each individual SNP was removed one at a time from the dataset and the significance was re-assessed. The SNP which least affected the significance of the odds ratio was then removed and the process was repeated. SNPs were progressively removed in this "stepdown" procedure until the significance of the odds ratio was no longer improved. Interestingly, it was observed that initial removal of SNPs substantially improved significance of the difference between cases and controls. When stepdown was completed, a total of 49 SNPs in 26 genes were observed [Table 2].

While the genes identified included key genes seen in the GWAS catalog, specifically members of HLA class II, other genes associated with antigen processing were also observed [Figure 1]. The design of Genome-wide association studies does not permit the specific etiologic effects of the variation. By design, the variation used in the studies is not chosen for function, but instead the ability to test differences between populations. The high linkage disequilibrium observed between variations in humans further complicates the capacity to interpret the molecular mechanisms of action.

Nevertheless, this study identifies variation of genes of potential significance in etiology. Of particular interest are the proteasome (HSPA2, HSPA4, HSPA5 HSP90AB1), endoplasmic reticulum TAP1, TAP2, CANX), and exosome (LGMN) genes associated with the processing of antigens so that they may be presented by HLA loci. The pathway also identifies genes on the surface of immune cells - NK cells (KIR2DL3, KIR2DL4, and KIR2DL5) and CD4 T cells (CD4) that may compromise immune surveillance and regulation.

It is possible to examine the intra-pathway associations of the variants. Using the analytic tool PLINK[18], one can estimate the association (r^2) between loci in cases and controls [Table 3]. As expected by the PoDA analysis, variants within the pathways are associated with one another. Both variants within loci and between loci are observed to be associated. Interestingly, the magnitude of associations differs between cases and

Table 2. Significant genes and SNPs within the KEGG antigen processing and presentation pathway

Gene symbol	Gene name	SNP (rs id)
CANX	Calnexin	rs7734102
CD4	CD4 molecule	rs1075835
CD74	CD74 molecule, major histocompatibility complex, class II invariant chain	rs2748249
CIITA	Class II, major histocompatibility complex, transactivator	rs6498122
CIITA	Class II, major histocompatibility complex, transactivator	rs7203275
CIITA	Class II, major histocompatibility complex, transactivator	rs11074934
CIITA	Class II, major histocompatibility complex, transactivator	rs6498119
CTSS	Cathepsin S	rs11204722
HLA-A	Major histocompatibility complex, class I, A	rs12202296
HLA-DMA	Major histocompatibility complex, class II, DM alpha	rs11539216
HLA-DMA	Major histocompatibility complex, class II, DM alpha	rs17617515
HLA-DMB	Major histocompatibility complex, class II, DM beta	rs3132132
HLA-DMB	Major histocompatibility complex, class II, DM beta	rs714289
HLA-DOA	Major histocompatibility complex, class II, DO alpha	rs3129304
HLA-DOA	Major histocompatibility complex, class II, DO alpha	rs3129303
HLA-DOA	Major histocompatibility complex, class II, DO alpha	rs3130602
HLA-DOA	Major histocompatibility complex, class II, DO alpha	rs3129302
HLA-DPB1	Major histocompatibility complex, class II, DP beta 1	rs9277378
HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2	rs9275356
HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2	rs9276427
HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2	rs9469266
HLA-DRA	Major histocompatibility complex, class II, DR alpha	rs7194
HLA-G	Major histocompatibility complex, class I, G	rs2517898
HSP90AB1	Heat shock protein 90kDa alpha (cytosolic), class B member 1	rs504697
HSPA2	Heat shock 70kDa protein 2	rs4313734
HSPA4	Heat shock 70kDa protein 4	rs7702889
HSPA5	Heat shock 70kDa protein 5	rs12009
HSPA8	Heat shock 70kDa protein 8	rs4936770
KIR2DL3	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3	rs9797797
KIR2DL3	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 3	rs13344915
KIR2DL4	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4	rs10500318
KIR2DL4	Killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 4	rs3865509
KIR2DS4	Killer cell immunoglobulin-like receptor, two domains, short cytoplasmic tail, 4	rs11673276
KLRD1	Killer cell lectin-like receptor subfamily D, member 1	rs17206564
LGMN	Legumain	rs8177528
LGMN	Legumain	rs2250672
LGMN	Legumain	rs716097
LGMN	Legumain	rs12885208
LGMN	Legumain	rs9791
LOC100509457	HLA class II histocompatibility antigen, DQ alpha 1 chain-like	rs2647015
LOC100509457	HLA class II histocompatibility antigen, DQ alpha 1 chain-like	rs2859090
LOC100509457	HLA class II histocompatibility antigen, DQ alpha 1 chain-like	rs9272219
RFXAP	Regulatory factor X-associated protein	rs6563500
TAP1	Transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	rs4148882
TAP2	Transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	rs3819720
TAP2	Transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	rs2228396
TAP2	Transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	rs241428
TAP2	Transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	rs9784758
TAP2	Transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)	rs241431

controls. This confirms that the pathway utilizes information (interactions between loci) that would not be observed in simple single locus GWAS assessments.

"ANTIGEN PROCESSING AND PRESENTATION" TRANSCRIPTIONAL ACTIVITY

It is possible to assess whether the germline variation in "antigen processing and presentation" translates into functionally significant difference in normal liver when contrasted to tumor adjacent liver and HCC.

Table 3. Association of case and control SNP variation with r^2 greater than 0.1 within the KEGG antigen processing and presentation pathway

SNP_A	SNP_B	Case r^2	Control r^2
SNP_A-4289896 - KIR2DL3	SNP_A-8561730 - KIR2DL3	0.88	0.95
SNP_A-8566010 - HLA-DQA1L	SNP_A-2200530 - TAP2	0.38	0.20
SNP_A-8515749 - HLA-G	SNP_A-8649593 - HLA-A	0.16	0.37
SNP_A-2214036 - HLA-DQA1L	SNP_A-4206711 - HLADQA1	0.16	0.14
SNP_A-8524421 - KIR2DL4	SNP_A-8613821 - KIR2DS4	0.14	< 0.1
SNP_A-1985650 - HLA-DOA	SNP_A-8430032 - KIR2DL3	0.12	< 0.1
SNP_A-2214036 - HLA-DQA1L	SNP_A-2200530 - TAP2	0.11	< 0.1
SNP_A-8451478 - TAP2	SNP_A-8415280 - TAP2	0.10	< 0.1
SNP_A-2305613 - CSTB	SNP_A-1944939 - CSTB	< 0.1	1.00
SNP_A-8566010 - HLA-DQA1L	SNP_A-1985650 - HLA-DOA	< 0.1	0.28
SNP_A-4223083 - HLA-DQA1L	SNP_A-8415280 - CIITA	< 0.1	0.18
SNP_A-4206711 - HLA-DQA1	SNP_A-8451478 - TAP2	< 0.1	0.16
SNP_A-4277940 - HLA-DQA1L	SNP_A-1985650 - HLA-DOA	< 0.1	0.14

This can be done by looking at the transcriptome of these tissues using publicly accessible data from the Gene Tissue Expression project (GTEx)[19-21] and the TCGA[8]. Data from both sources were processed with a common analytic pipeline that included realignment of sequencing reads to Hg38[22,23], uniform count scoring[24] and adjustment for over-dispersion[25,26].

The scored transcript data was then evaluated using the novel pathway analysis tool PathOlogist[27-29]. PathOlogist utilizes the logical information contained within networks to compute network scores. By utilizing the structure of a network, in this approach the conditional state of genes determines expectations for the state of other members of the network. Two different scores are provided. The first assesses whether the activity state of the network differs. In the second, an assessment of the logical state of the network is measured as consistency. Consistency determines whether the transcription patterns follow the expected logic of the network.

Examination of the transcriptional state of "antigen processing and presentation" provides additional insight into the susceptibility findings. First, "antigen processing and presentation" activity is observed to be significantly higher in normal liver (GTEx) compared to TCGA tumor-adjacent (adjusted $P < 0.0001$) and tumor (adjusted $P < 0.0001$) while no difference is observed between tumor adjacent and tumor (adjusted $P = 0.87$). This suggests that individuals with HCC have a different "antigen processing and presentation" profile in both their non-tumor and tumor than normal liver.

No significant difference is observed between the consistency scores of normal liver (GTEx) and TCGA tumor-adjacent (adjusted $P = 0.64$) and tumor adjacent and tumor (adjusted $P = 0.89$b) for "antigen processing and presentation". However, significant difference is observed between normal liver and tumor (adjusted $P < 0.0001$). This suggests that "antigen processing and presentation" may be a target of mutagenesis in HCC.

IMMUNE CHECKPOINT THERAPY AND "ANTIGEN PROCESSING AND PRESENTATION"

"Antigen processing and presentation" may be an important mediator of treatment response for HCC. Immune checkpoint therapy is dramatically altering the cancer therapeutic landscape[30]. Checkpoint therapy targets inhibitory signals to the immune system such as CTLA-4 and PD-1/PD-L1. These treatments show promising, durable response results in previously treatment resistant cancers such as melanoma[31] and non-small cell lung cancer[32]. The US FDA has approved checkpoint therapy for second line treatment of HCC. Numerous studies are in progress to assess the efficacy as 1st line treatment (clinicaltrials.gov).

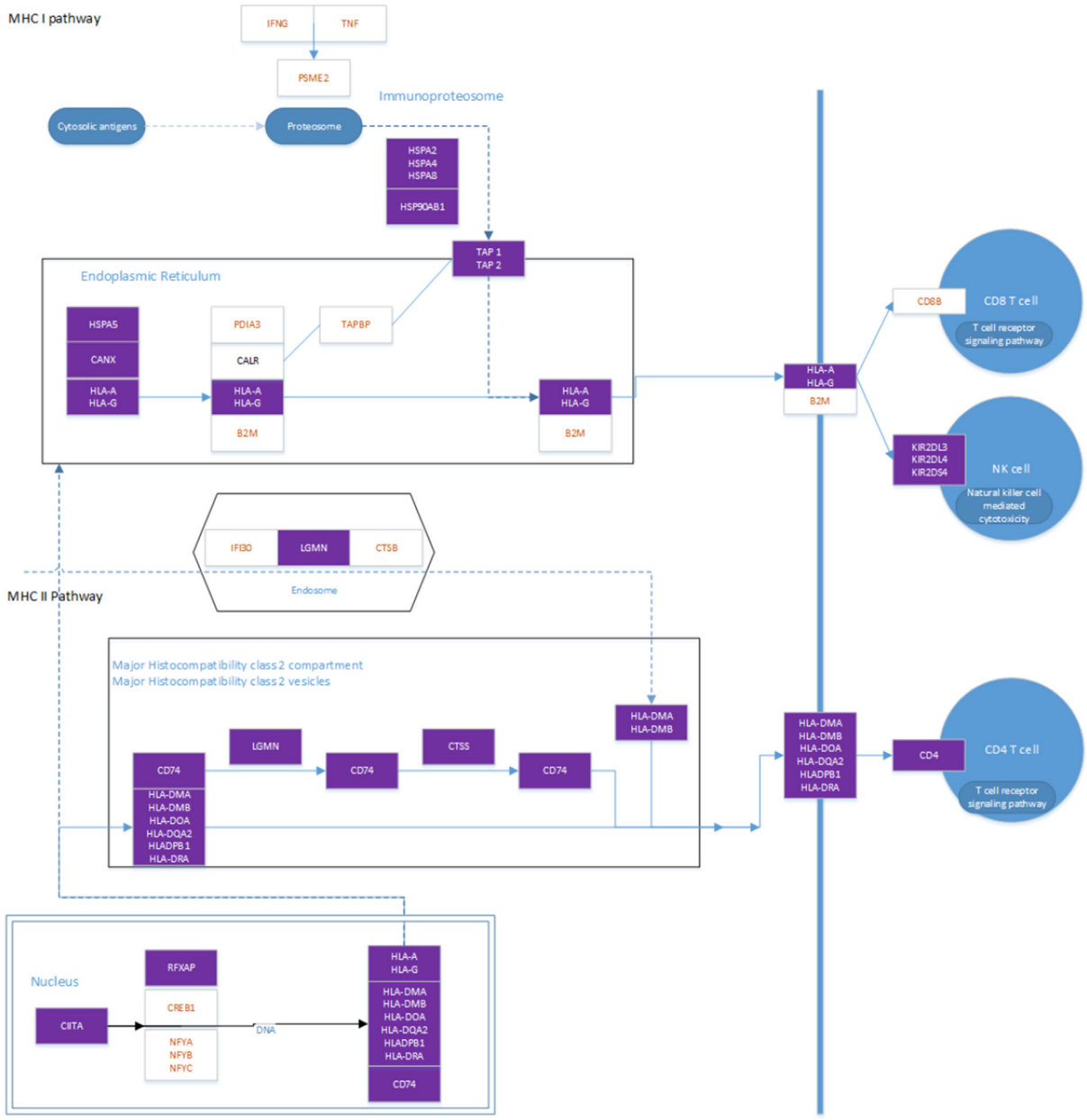

Figure 1. Gene-based SNPs associated with HCC in the antigen processing and presentation pathway. The genes and their relationships obtained from KEGG's antigen processing and presentation pathway. Purple boxes with white letters indicate genes SNP variations associated with HCC from the PoDA stepdown analysis. Removal of these loci reduced the overall threshold of significance below that observed for the entire pathway. Genes in open boxes (with orange letters) indicate gens which could be removed without altering significance of the pathway's association. HCC: hepatocellular carcinoma

Unfortunately only a minority of individuals respond to the treatments[33]. It is unknown what mediates response. Indicators of response include DNA mismatch repair capabilities[34] and tumor mutational burden[35]. But these have poor predictive capabilities.

For checkpoint therapy to work, an intact immune response is required. As implied from the indicators of response, the immune system must have the capacity to recognize tumor antigens as foreign. This recognition is mediated through antigen processing and presentation. Inherited variability may indicate individuals in which this capacity is compromised. Moreover, variation in these processes may indicate individual response to immune directed therapeutic interventions.

In conclusion, the results of the germline variation studies suggest that immune mediating processes are polymorphic in the population and systematically different in HCC. Individuals with HCC have significantly lower activity for these processes and HCC shows alterations in the "logic" of the processing and presentation pathways. As such, it may be possible to predict response to checkpoint therapy through the evaluation of the inherited genetic state of "antigen processing and presentation". Understanding these differences may provide opportunities designing new immune checkpoint modulators and provide a rational basis for combinatorial therapy.

DECLARATIONS

Acknowledgments

The Korean HCC case-control study was collected by Dr. Myung Lyu (NCI/NIH/DHSS) and Dr. Young-Hwa Chung (Asan Medical Center, Seoul, South Korea).

Authors' contributions

Data analysis: Lu YK, Brill JM, Aghili A

Design of the work, data analysis, manuscript drafting and revising, and final approval of the version to be published: Buetow KH

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

Buetow KH is an advisor for the Bristol Myers Squibb IO-ICON project.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

REFERENCES

1. Torre LA, Siegel RL, Ward EM, Jemal A. Global cancer incidence and mortality rates and trends - an update. Cancer Epidemiol Prev Biomark 2016;25:16-27.
2. Petrick JL, Braunlin M, Laversanne M, Valery PC, Bray F, McGlynn KA. International trends in liver cancer incidence, overall and by histologic subtype, 1978-2007. Int J Cancer 2016;139:1534-45.
3. McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. Clin Liver Dis 2015;19:223-38.
4. Makarova-Rusher OV, Altekruse SF, McNeel TS, Ulahannan S, Duffy AG, Graubard BI, Greten TF, McGlynn KA. Population attributable fractions of risk factors for hepatocellular carcinoma in the United States. Cancer 2016;122:1757-65.
5. Jemal A, Ward EM, Johnson CJ, Cronin KA, Ma J, Ryerson B, Mariotto A, Lake AJ, Wilson R, Sherman RL, Anderson RN, Henley SJ, Kohler BA, Penberthy L, Feuer EJ, Weir HK. Annual report to the nation on the status of cancer, 1975-2014, featuring survival. J Natl Cancer Inst 2017;109:djx030.
6. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, Tanaka H, Taniguchi H, Kawakami Y, Ueno M, Gotoh K, Ariizumi S, Wardell CP, Hayami S, Nakamura T, Aikata H, Arihiro K, Boroevich KA, Abe T, Nakano K, Maejima K, Sasaki-Oku A, Ohsawa A, Shibuya T, Nakamura H, Hama N, Hosoda F, Arai Y, Ohashi S, Urushidate T, Nagae G, Yamamoto S, Ueda H, Tatsuno K, Ojima H, Hiraoka N, Okusaka T, Kubo M, Marubashi S, Yamada T, Hirano S, Yamamoto M, Ohdan H, Shimada K, Ishikawa O, Yamaue H, Chayama K, Miyano S, Aburatani H, Shibata T, Nakagawa H. Whole-genome mutational landscape and characterization of noncoding and structural mutations in

liver cancer. Nat Genet 2016;48:500-9.
7. Schulze K, Imbeaud S, Letouzé E, Alexandrov LB, Calderaro J, Rebouissou S, Couchy G, Meiller C, Shinde J, Soysouvanh F, Calatayud AL, Pinyol R, Pelletier L, Balabaud C, Laurent A, Blanc JF, Mazzaferro V, Calvo F, Villanueva A, Nault JC, Bioulac-Sage P, Stratton MR, Llovet JM, Zucman-Rossi J. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet 2015;47:505-11.
8. Cancer Genome Atlas Research Network. Comprehensive and integrative genomic characterization of hepatocellular carcinoma. Cell 2017;169:1327-41.e23.
9. Shen FM, Lee MK, Gong HM, Cai XQ, King MC. Complex segregation analysis of primary hepatocellular carcinoma in Chinese families: interaction of inherited susceptibility and hepatitis B viral infection. Am J Hum Genet 1991;49:88-93.
10. Yu MW, Chang HC, Liaw YF, Lin SM, Lee SD, Liu CJ, Chen PJ, Hsiao TJ, Lee PH, Chen CJ. Familial risk of hepatocellular carcinoma among chronic hepatitis B carriers and their relatives. J Natl Cancer Inst 2000;92:1159-64.
11. Hassan MM, Spitz MR, Thomas MB, Curley SA, Patt YZ, Vauthey JN, Glover KY, Kaseb A, Lozano RD, El-Deeb AS, Nguyen NT, Wei SH, Chan W, Abbruzzese JL, Li D. The association of family history of liver cancer with hepatocellular carcinoma: a case-control study in the United States. J Hepatol 2009;50:334-41.
12. Dragani TA. Risk of HCC: genetic heterogeneity and complex genetics. J Hepatol 2010;52:252-7.
13. Turati F, Edefonti V, Talamini R, Ferraroni M, Malvezzi M, Bravi F, Franceschi S, Montella M, Polesel J, Zucchetto A, La Vecchia C, Negri E, Decarli A. Family history of liver cancer and hepatocellular carcinoma. Hepatology 2012;55:1416-25.
14. Caruso S, Calderaro J, Letouzé E, Nault JC, Couchy G, Boulai A, Luciani A, Zafrani ES, Bioulac-Sage P, Seror O, Imbeaud S, Zucman-Rossi J. Germline and somatic DICER1 mutations in familial and sporadic liver tumors. J Hepatol 2017;66:734-42.
15. Jin F, Xiong WJ, Jing JC, Feng Z, Qu LS, Shen XZ. Evaluation of the association studies of single nucleotide polymorphisms and hepatocellular carcinoma: a systematic review. J Cancer Res Clin Oncol 2011;137:1095-104.
16. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, Junkins H, McMahon A, Milano A, Morales J, Pendlington ZM, Welter D, Burdett T, Hindorff L, Flicek P, Cunningham F, Parkinson H. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res 2017;45:D896-901.
17. Braun R, Buetow K. Pathways of distinction analysis: a new technique for multi-SNP analysis of GWAS data. PLoS Genet 2011;7:e1002101.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-75.
19. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648-60.
20. GTEx Consortium; Laboratory, Data Analysis &Coordinating Center (LDACC)-Analysis Working Group; Statistical Methods groups-Analysis Working Group; Enhancing GTEx (eGTEx) groups; NIH Common Fund; NIH/NCI; NIH/NHGRI; NIH/NIMH; NIH/NIDA; Biospecimen Collection Source Site-NDRI; Biospecimen Collection Source Site-RPCI; Biospecimen Core Resource-VARI; Brain Bank Repository-University of Miami Brain Endowment Bank; Leidos Biomedical-Project Management; ELSI Study; Genome Browser Data Integration & Visualization-EBI; Genome Browser Data Integration & Visualization-UCSC Genomics Institute, University of California Santa Cruz; Lead analysts; Laboratory, Data Analysis & Coordinating Center (LDACC); NIH program management; Biospecimen collection; Pathology; eQTL manuscript working group; Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. Nature 2017;550;204-13.
21. Saha A, Kim Y, Gewirtz ADH, Jo B, Gao C, McDowell IC; GTEx Consortium; Engelhardt BE, Battle A. Co-expression networks reveal the tissue-specific regulation of transcription and splicing. Genome Res 2017;27:1843-58.
22. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015;12:357-60.
23. Pertea M, Kim D, Pertea GM, Leek JT, Salzberg SL. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat Protoc 2016;11:1650-67.
24. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 2014;30:923-30.
25. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 2014;15:R29.
26. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
27. Efroni S, Schaefer CF, Buetow KH. Identification of key processes underlying cancer phenotypes using biologic pathway analysis. PLoS One 2007;2:e425.
28. Efroni S, Carmel L, Schaefer CG, Buetow KH. Superposition of transcriptional behaviors determines gene state. PLoS One 2008;3:e2901.
29. Greenblum SI, Efroni S, Schaefer CF, Buetow KH. The PathOlogist: an automated tool for pathway-centric analysis. BMC Bioinformatics 2011;12:133.
30. Sharma P, Allison JP. The future of immune checkpoint therapy. Science 2015;348:56-61.
31. Garbe C, Peris K, Hauschild A, Saiag P, Middleton M, Bastholt L, Grob JJ, Malvehy J, Newton-Bishop J, Stratigos AJ, Pehamberger H; Eggermont AM; European Dermatology Forum (EDF); European Association of Dermato-Oncology (EADO); European Organisation for Research and Treatment of Cancer (EORTC). Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline - update 2016. Eur J Cancer 2016;63:201-17.
32. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, Patnaik A, Aggarwal C, Gubens M, Horn L, Carcereny E, Ahn MJ, Felip E, Lee JS, Hellmann MD, Hamid O, Goldman JW, Soria JC, Dolled-Filhart M, Rutledge RZ, Zhang J, Lunceford JK, Rangwala R, Lubiniecki GM, Roach C, Emancipator K, Gandhi L; KEYNOTE-001 Investigators. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med 2015;372:2018-28.

33. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. Cell 2017;168:707-23.
34. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, Laheru D, Donehower R, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Greten TF, Duffy AG, Ciombor KK, Eyring AD, Lam BH, Joe A, Kang SP, Holdhoff M, Danilova L, Cope L, Meyer C, Zhou S, Goldberg RM, Armstrong DK, Bever KM, Fader AN, Taube J, Housseau F, Spetzler D, Xiao N, Pardoll DM, Papadopoulos N, Kinzler KW, Eshleman JR, Vogelstein B, Anders RA, Diaz LA Jr. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 2017;357:409-13.
35. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA, Kurzrock R. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. Mol Cancer Ther 2017;16:2598-608.