The pattern of gene copy number variations (CNVs) in hepatocellular carcinoma; in silico analysis

Hossein Ansari

Islamic Azad University

Arman Shahriza

Tarbiat Modares University Faculty of Biological Sciences

Maryam Tahmaseby (✉ maryam_tahmaseby@yahoo.com)

Ahvaz Jondishapour University of Medical Sciences

Zahra Mohammadi

Ahvaz Jondishapour University of Medical Sciences

Vinicio Carloni

Florence University

Javad Mohammadi Asl

Ahvaz Jondishapour University of Medical Sciences

Research

Keywords: Hepatocellular carcinoma, copy number alteration, RNA dysregulation, 8p, 1q

Posted Date: April 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-20514/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Cancer-associated death is the second leading cause of death worldwide. Study of the involved molecular networks of cancers can identify the potential target for early diagnosis, efficient therapies, and predictive prognosis of patients with cancer. Copy number variations are one type of DNA mutations which has been connected with human cancers. The CNVs can be used to find the regions of the genome involved in cancer phenotypes. This study is aimed to perform genome-wide chromosomal CNVs in HCC samples to find hotspot regions of disease using in silico analysis. The obtained data showed that gain of chromosome 1q and loss of 8p were frequently observed in target cancerous tissues. All the gains and losses were associated with tumor grade and metastasis. However, the amplification of YY1AP1 (Yin Yang-1 Associated Protein 1) and deletion of CHMP7 (Charged Multivesicular Body Protein 7) were observed in most of patients. These expression levels of YY1AP1 and CHMP7 were also upregulated and downregulated in cancerous samples respectively. Additionally, these two genes interact with critical oncogene and tumor suppressor genes like MDM2 (Mouse double minute 2 homolog) and VHL (von Hippel-Lindau tumor suppressor) showing the potential of these genes in HCC pathogenesis. Based on the observed data we suggest the 1q and 8p as candidate regions for HCC for researches especially YY1AP1 and CHMP7 loci.

Background

Hepatocellular carcinoma (HCC) is one the most lethal cancers which is mostly induced by hepatitis B virus infection, alcoholism and metabolic disorders[1, 2]. Although the liver lesion is usually seen on computed tomography, many HCC tumors are asymptomatic and most patients at risk for HCC will not be diagnosed in time, so the incidence of HCC is still rising and prognosis of the disease remains poor [3]. Multiple line of evidences showed that the incidence of HCC increases progressively with the age proposing that the accumulation of genetic alterations over the time associates with HCC development. These genetic alterations are included with variety of genetic gains and losses which are totally termed as gene copy number alteration (CNAs) [3]. Actually, chromosomal aberrations are one the common feature of malignant cells in which the number or structure of chromosomes deviated from normal diploid state [4]. Growing body of evidences has been demonstrated the significance of such karyotyping changes on stability of genome and facing the cells toward tumorigenesis [5]. Additionally, it has revealed that chromosomal aberrations are contributed to tumor age of onset, tumor metastatic state, drug resistant and tumor failure phenotypes [6]. Recent advances in functional genomics techniques like comparative genomic hybridization (CGH) or microarray open a new window to characterize the cytogenetic signatures of the malignant cells as genomic loss and gains can be simply diagnosed in medical samples. The issue that seems to be promising in diagnostic and therapeutic area, these days[7]. Here, by means of in silico analysis we have introduced the most detected CNVs in 370 HCC samples which could be an interesting result to define the experimental job on HCC

Results

As illustrated in Fig. 1, chromosome 1 and chromosome 8 loci carry the hotspot loci of gene amplification and deletion which will be discussed below.

Cytobands 1q and 8q were strongly associated with gene amplification in HCC samples

From 24776 examined genes, 1024 genes (41%) were amplified in studied samples with linear copy number values cut-off ≥ 0.5 (Supplementary file 1). These genes were mapped on 1p (1.3%; n = 13), 1q (86%; n = 880) and 8q (13%; n = 131) of chromosome arms (Fig. 2A). The highest scores were belonged to 35 genes including ADAM15, FLAD1, ZBTB7B, PYGO2, DCST2, LENE, SHC1, PBXIP1, PMVK, DCST1, KCNN3, SLC50A1, DPM3, SCAMP3, FAM189B, MTX1, KRTCAP2, GBA, GBAP1, MUC1, THBS3, TRIM46, ASH1L, RUSC1, CLK2, HCN3, FDPS, PKLR, YY1AP1, MSTO1, GON4L, RIT1, SYT11 and KIAA0907.

The genes encode the proteins which act as enzyme, transcription factor, signaling molecule, receptor, cytoskeletal protein, cell adhesion extra cellular matrix protein in different signaling pathways including Wnt, inflammation mediated by chemokine and cytokine, EGFR, Integrin, FGF, angiogenesis, Ras, VEGF, Glycolysis, Apoptosis, p53, TGF- β , PI3K, Hedgehog, FAS and JAK/STAT signaling pathways (Fig. 2B) (Supplementary File 2). Most of the amplified genes were located in 1q22 (65.7%; n = 23), and the remaining genes were belonged to 1q21.3 (34.3%; n = 11) (Fig. 2A). The amplification of most of these genes was associated with tumor grade (Table 1). However, FLAD1 gene was associated with reduced overall survival in patients too (Logrank P-value = 0.01) (Fig. 3). The amplifications of 3 genes on 1q including YY1AP1, SCAMP3 and GBA were found in more than 40% studied patients. Although, all the 35 target genes were amplified in HCC samples, However, the expression of only YY1AP1 showed the strong correlation with their corresponding CNVs ($r > 0.6$) in 72% of samples. Extraction the interactions and pathways of YY1AP1 in UCSC genome browser showed that this protein interact with oncogene MDM2, tumor suppressor gene VHL, tyrosine kinase receptor MAPK, proteasome protein UBC, EGLN3 and SMURF1 and transcription factor CEBPA (Fig. 4).

Table 1

The association of Copy number variations with clinicopathologic parameters of 370 HCC tissues. Data was extracted from cBioPortal dataset (<https://www.cbioportal.org>)

Gene	CNV type	Stage	Metastasis	Grade	Gene	CNV type	Stage	Metastasis	Grade
YY1AP1	Amp	-	-	+	LENEP	Amp	-	-	+
FLAD1	Amp	-	-	+	MTX1	Amp	-	-	-
ADAM15	Amp	-	-	-	MSTO1	Amp	-	-	+
FDPS	Amp	-	-	-	KIAA0907	Amp	-	-	+
FAM189B	Amp	-	-	+	MUC1	Amp	-	-	-
DCST1	Amp	-	-	-	ASH1L	Amp	-	-	-
DCST2	Amp	-	-	+	CLK2	Amp	-	-	-
ZBTB7B	Amp	-	-	+					
TRIM46	Amp	-	-	+					
THBS3	Amp	-	-	-					
SYT11	Amp	-	-	+					
SLC50A1	Amp	-	-	-					
SHC1	Amp	-	-	+					
SCAMP3	Amp	-	-	+					
RUSC1	Amp	-	-	-					
RIT1	Amp	-	-	+					
PYG02	Amp	-	-	+					
PMVK	Amp	-	-	+					
PKLR	Amp	-	-	-					
PBXIP1	Amp	-	-	+					
GBA	Amp	-	-	-					
GBAP1	Amp	-	-	-					

** Use of color for printed figures is not needed.

Gene	CNV type	Stage	Metastasis	Grade	Gene	CNV type	Stage	Metastasis	Grade
GON4L	Amp	-	-	+					
HCN3	Amp	-	-	-					
KRTCAP2	Amp	-	-	-					
** Use of color for printed figures is not needed.									

Cytobands 8p was strongly associated with gene deletion in HCC samples

From 24776 examined genes, 172 genes (7%) were lost in studied samples (n = 370) with linear copy number values cut-off ≥ 0.5 which were mostly located on 8p21-23. The most deletion scores were belonged to 17 genes including TNFRSF10B, RHOTB2, PEBP4, TNFRSF10C, CHMP7, TNFRSF10A, ENTPD4, EGR3, BIN3, PDLIM2, R3HCC1, LOXL2, STC1, PIWIL2, SLC25A37, TNFRSF10D and CSMD1 and considered for more analysis as target genes. Among target genes, 16 genes were mapped on 8p21.3 and 1 one gene 8p23.2 (Fig. 5A). These genes were transcription factor and cytoskeletal proteins act in apoptosis, EGF, FGF and P53 signaling pathways (Fig. 5B) (Supplementary file 2). The deletions of none of mentioned genes were associated with tumor metastasis or grade (Table 1). Although all of mentioned genes were downregulated in cancerous tissues, the correlation with CNV was not significant and only CHMP7 showed the moderate correlation ($r = 0.5$) in 70% of studied samples. Interaction analysis showed that CHMP7 can interact with growth inhibitory protein VPS37A, proteasome protein TSG101 and variety of transporter proteins (Fig. 6).

Discussion

Liver cancer is one of human cancers with poor prognosis and has a lot to do with its genetic signature. Identification of key genetic components and new therapeutic targets is important issue specially for patients of hepatocellular carcinoma (HCC) patients who are ineligible for surgical resection or liver transplant. Although the heterogenous nature of HCC has been complicated this approach [8]. It is truly amazing that exponential advances in genomic sequencing during the past 10 years along with the entrance of bioinformatics tools to translate such high-throughput data, speed up the discovery of hundreds of potential new targets in human diseases specially cancers and draw a perspective to propose new prognostic, diagnostic and therapies approaches [9]. However, there is an urgent need to work on such putative targets experimentally to reveal the druggable candidates [8]. Copy number variations are one kind of DNA mutations seem have high impact in cancer pathogenesis. These genetic variations are also valuable tools for finding the hotspot regions in cancers [10]. In this study, the authors did genome-wide chromosomal CNVs in HCC patients. This was a bioinformatics analysis using available microarray mega data of HCC microarray in highly cited cancer portals. Around 24776 genes were screened in this study for the presence of CNVs. As

chromosomal CNVs may contribute to their transcript expression in HCC, the RNA sequencing profile of target genes were also considered. The observed data showed that the chromosome 1q and chromosome 8p were the most regions for gene amplification and deletion respectively. In study by Takafumi Nishimura et al, they showed that chromosomal gain at 1q is one of the most common features of HCC [11, 12]. In other study by Chen and coworker 1q was introduced as host spot regions of gene amplification in HCC the regions of 1q12–q22, 1q23.3–q25.3 and 1q23.1–1q43 were reported as minimal amplified region on chromosome 1q [13]. The 8p region was also highlighted in a study by Roessler [14] and Qin [15], separately. It has also demonstrated that HCC may develop from cirrhotic cells carrying 8p loss [16]. Our candidate genes were associated with tumor grade and metastasis, mostly. We also find out that amplification of gene FLAD1 can reduced the overall survival of patients harboring this CNV. Growing body of evidences showed that amplifications of 1q and 8q have been strongly connected with tumor grade and size [3]. We also found that all of selected genes were involved in pathways which previously reported as critical rout in HCC cells. these included with Wnt signaling pathway, angiogenesis signaling pathway, apoptosis signaling pathway, Ras signaling pathway, TGF- β signaling pathway, PI3K signaling pathway, P53 signaling pathway, EGF/EGFR signaling pathway, FAS signaling pathway, JAK/STAT signaling, Hedgehog signaling pathway or inflammatory signaling pathways[17]. We also found the gain of YY1AP1 and lose of CHMP7 were observed in most patients and resulted in upregulation and downregulation of corresponding transcript respectively. it has also found that YY1AP1 interact with oncogene MDM2, tumor suppressor gene VHL, tyrosine kinase receptor MAPK, proteasome protein UBC,EGLN3 and SMURF1 and transcription factor CEBPA while CHMP7 can interact with growth inhibitory protein VPS37A, proteasome protein TSG101 and variety of transporter proteins. The YY1AP1 is a component of the INO80 chromatin remodeling complex, which is responsible for transcriptional regulation, DNA repair, and replication[18]. In a recent study by Zhao X, they found that YY1AP1 may serve as a key molecular target for EpCAM(+) AFP(+) HCC subtype Which was attributed with poor prognosis and stem cell-like phenotype[8]. They also showed that YY1AP1 is associated with stem cell features of this subtype and silencing of YY1AP1 eliminates the oncogenic feature of the cells through the altering the chromatin landscape and triggering massive apoptosis in vitro and vivo[8]. However, no publication was found for CHMP7 gene showing its potency for further analysis in HCC samples. Altogether, the observed data we suggest the 1q and 8p as candidate regions for HCC for researches especially YY1AP1 and CHMP7 loci.

Methods

The TCGA CNA raw data of 24777 genes in 370 samples of Liver Hepatocellular Carcinoma (TCGA, Provisional) was extracted from cBioPortal (<http://www.cbioportal.org>) and analyzed in R v3.5 using the cgdsr extension package (cran.rproject.org/web/packages/cgdsr/). The linear copy number values cut-off ≥ 0.5 and cut-off ≥ -0.5 were considered as threshold of gene amplification and deletion respectively. The filtered genes were selected as target genes for more analysis. The CNVs frequency of target genes was also calculated in HCC samples. Using the PANTHER classification system (<http://www.pantherdb.org>), the protein classes and the involved pathways of target proteins were estimated. The association of CNV variants with pathological tumor grade and stage was also examined. Consequently, we extracted the TCGA RNA-seq raw data of Liver Hepatocellular Carcinoma (TCGA, Provisional) for these candidate genes to examine if any correlation was exist between CNA of target genes and their corresponding expression value. As in strong

positive correlation, the linear correlation coefficient (r) is close to + 1, the results were filtered based on the $r > 0.7$. Using the UCSC genome browser, the interaction of target protein was traced (<https://genome.ucsc.edu>).

Declarations

Acknowledgements

Not applicable

Authors' contributions

All authors have participated in the study and manuscript preparation. All authors read and approved the final manuscript.

Funding

Not applicable

Availability of data

The datasets used and/or analyzed during the current study are available on cbioportal.org and are also accessible through `cgdsr` R package.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interest.

References

1. Tahmasebi Birgani M, Carloni V. Tumor Microenvironment, a Paradigm in Hepatocellular Carcinoma Progression and Therapy. *Int J Mol Sci.* 2017;18(2).
2. Mazzocca A, Birgani MT, Sabba C, Carloni V. Tetraspanin-enriched microdomains and hepatocellular carcinoma progression. *Cancer Lett.* 2014;351(1):23-29.
3. Kim TM, Yim SH, Shin SH, Xu HD, Jung YC, Park CK *et al.* Clinical implication of recurrent copy number alterations in hepatocellular carcinoma and putative oncogenes in recurrent gains on 1q. *Int J Cancer.* 2008;123(12):2808-2815.
4. Albertson DG, Collins C, McCormick F, Gray JW. Chromosome aberrations in solid tumors. *Nat Genet.* 2003;34(4):369-376.
5. Vargas-Rondon N, Villegas VE, Rondon-Lagos M. The Role of Chromosomal Instability in Cancer and Therapeutic Responses. *Cancers (Basel).* 2017;10(1).
6. Sansregret L, Swanton C. The Role of Aneuploidy in Cancer Evolution. *Cold Spring Harb Perspect Med.* 2017;7(1).
7. Monni O, Hyman E, Mousses S, Barlund M, Kallioniemi A, Kallioniemi OP. From chromosomal alterations to target genes for therapy: integrating cytogenetic and functional genomic views of the breast cancer genome. *Semin Cancer Biol.* 2001;11(5):395-401.
8. Zhao X, Parpart S, Takai A, Roessler S, Budhu A, Yu Z *et al.* Integrative genomics identifies YY1AP1 as an oncogenic driver in EpCAM(+) AFP(+) hepatocellular carcinoma. *Oncogene.* 2015;34(39):5095-5104.
9. Wu D, Rice CM, Wang X. Cancer bioinformatics: a new approach to systems clinical medicine. *BMC Bioinformatics.* 2012;13:71.
10. Shlien A, Malkin D. Copy number variations and cancer. *Genome Med.* 2009;1(6):62.
11. Nishimura T, Nishida N, Itoh T, Komeda T, Fukuda Y, Ikai I *et al.* Discrete breakpoint mapping and shortest region of overlap of chromosome arm 1q gain and 1p loss in human hepatocellular carcinoma detected by semiquantitative microsatellite analysis. *Genes Chromosomes Cancer.* 2005;42(1):34-43.
12. Rao UN, Gollin SM, Beaves S, Cieply K, Nalesnik M, Michalopoulos GK. Comparative genomic hybridization of hepatocellular carcinoma: correlation with fluorescence in situ hybridization in paraffin-embedded tissue. *Mol Diagn.* 2001;6(1):27-37.
13. Chen YJ, Yeh SH, Chen JT, Wu CC, Hsu MT, Tsai SF *et al.* Chromosomal changes and clonality relationship between primary and recurrent hepatocellular carcinoma. *Gastroenterology.* 2000;119(2):431-440.
14. Roessler S, Long EL, Budhu A, Chen Y, Zhao X, Ji J *et al.* Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. *Gastroenterology.* 2012;142(4):957-966 e912.
15. Qin LX, Tang ZY, Sham JS, Ma ZC, Ye SL, Zhou XD *et al.* The association of chromosome 8p deletion and tumor metastasis in human hepatocellular carcinoma. *Cancer Res.* 1999;59(22):5662-5665.
16. Kishimoto Y, Shiota G, Wada K, Kitano M, Nakamoto K, Kamisaki Y *et al.* Frequent loss in chromosome 8p loci in liver cirrhosis accompanying hepatocellular carcinoma. *J Cancer Res Clin Oncol.* 1996;122(10):585-589.

17. Villanueva A, Newell P, Chiang DY, Friedman SL, Llovet JM. Genomics and signaling pathways in hepatocellular carcinoma. *Semin Liver Dis.* 2007;27(1):55-76.

18. Guo DC, Duan XY, Regalado ES, Mellor-Crummey L, Kwartler CS, Kim D *et al.* Loss-of-Function Mutations in YY1AP1 Lead to Grange Syndrome and a Fibromuscular Dysplasia-Like Vascular Disease. *Am J Hum Genet.* 2017;100(1):21-30.

Figures

Figure 1

Chromosome ideograms with hotspot regions of amplification (green) and deletion (red).

Figure 2

(A) the frequency of gene amplification in chromosome 1 and 8 in HCC cells. Data was extracted from cBioPortal dataset (<https://www.cbioportal.org>). (B) The signaling pathways associated to amplified genes. Data was extracted from PANTHER dataset (<http://www.pantherdb.org>).

Figure 3

The association of FLAD1 gene alteration(s) with reduced overall survival of HCC patients. Data was extracted from cBioPortal dataset (<https://www.cbioportal.org>)

Figure 4

The top 15 interaction protein with YY1AP1. Data was extracted from UCSC genome browser (<https://genome.ucsc.edu>).

Figure 5

(A) the frequency of gene deletion in chromosome 8 in HCC cells. Data was extracted from cBioPortal dataset (<https://www.cbioportal.org>). (B) The signaling pathways associated to deleted genes. Data was extracted from PANTHER dataset (<http://www.pantherdb.org>).

Figure 6

The top 15 interaction protein with CHMP7. Data was extracted from UCSC genome browser (<https://genome.ucsc.edu>)