Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE MEDICHE GENERALI E SCIENZE DEI SERVIZI

Ciclo 32º
Settore Concorsuale: 06/B1
Settore Scientifico Disciplinare: MED/09

ANALYSIS AND CHARACTERIZATION OF MITOCHONDRIAL DNA MUTATIONS IN THE CANCER GENOME ATLAS HEPATOCELLULAR CARCINOMA COHORT

Presentata da ***Camelia Alexandra Coada***

Coordinatore Dottorato

Prof. Fabio Piscaglia

Supervisore

Prof. Fabio Piscaglia

Esame finale anno 2020

Table of contents

Table of contents 2

Abstract 5

Introduction 7

Hepatocellular carcinoma 7

The role of mitochondrial metabolism in cancer 10

Mitochondrial genetics 14

Mitochondrial mutations in cancer 18

The Cancer Genome Atlas (TCGA): a valuable resource for 'omic studies 20

Aim 21

Methodology 22

Bioinformatics analysis of mtDNA mutations 22

Immunohistochemistry staining 24

Results 25

Evaluation of the TCGA LIHC cohort revealed 104 cases eligible for mtDNA mutation analysis 25

Non-tumor specific mutations in HCC 27

Tumor specific mutations in HCC 30

Adjusting for the microenvironment component of the HCC tumor increases mtDNA variant detection 35

Germline mtDNA variants show heteroplasmic change under positive or negative selective pressure in HCC tumors 37

mtDNA mutations correlate with immune infiltration in HCC tumors 39

Pathogenic mtDNA mutations predict overall survival of TCGA LIHC patients 41

TP53 status predict overall survival independent of mitochondrial DNA mutations burden 49

Sources of mtDNA damage in the LIHC-TCGA cohort 54

HCC patients present also mutations in OXPHOS genes encoded by the nuclear genome 60

Livers with HCC present a heterogenous staining for mitochondrial complexes 62

Discussion 68

Conclusion 77

Acknowledgements 78

Bibliography 79

Annex 1 94

"I put my **heart and my soul into my work,**
and have lost my mind in the process."

Vincent van Gogh

Abstract

Background and Aims

Hepatocellular carcinoma (HCC) is the most common primary hepatic malignancy and represents the second cause of cancer related death worldwide, characterized by high recurrence rates and poor survival, even when detected and treated at its early stages.

Mitochondrial mutations have been known to play a role in carcinogenesis, but to date, few studies correctly prioritize and interpret the variants discovered. Thus, we aimed to identify and analyze the occurrence and clinical impact of mtDNA mutations in the HCC dataset from The Cancer Genome Atlas (TCGA) consortium - National Cancer Institute.

Method

Whole exome sequencing *fastq* files from 377 TCGA-HCC patients were downloaded from the TCGA database. Paired tumor, non- tumor tissues originating from each patient were processed to reconstruct the mtDNA genomes using the MToolBox automated pipeline. Pairwise comparison between blood/normal solid tissue and tumor was performed in order to identify the potentially germline and tumor-specific somatic mtDNA variants. Information regarding the variability and pathogenicity of the variants were obtained from HmtVar database.

Results

The assembly of the mitochondrial reads showed an adequate coverage and quality for 104 patients. Variants were classified as pathogenic based on the allele frequency and disease score using the HmtVar criteria. After discarding the germline variants used in haplogroup classification, fixing the heteroplasmic fraction (HF) at 0.4 and prioritizing the variants we found 13 pathogenic/likely-pathogenic missense mutations and three tRNA pathogenic mutations in tRNA genes. HCC tumors presented a total of 302 somatic variants. After applying the same criteria, we found 24 pathogenic mutations in 22 patients. The burden of pathogenic mtDNA mutations resulted independently associated with a poorer survival of these patients ($p<0.05$ for different heteroplasmy thresholds).

Conclusion

We found 21% of HCC patients to harbor somatic pathogenic mtDNA mutations in their tumors. We found that these patients had a poorer survival than those harbouring non-pathogenic variants. mtDNA mutations could cause mitochondrial dysfunction and impact the prognosis and survival of HCC patients.

Introduction

Hepatocellular carcinoma

Hepatocellular carcinoma (HCC) is the most frequent form of primary liver cancer, ranking sixth in worldwide incidence. It is characterized by a poor prognosis and a high mortality rate even when detected and treated at its early stage. [1–3]

With limited loco-regional therapeutic options, having a 5-year recurrence rate as high as 70% for the few patients that achieve complete response,[4] and only one systemic therapy approved, HCC remains one of the deadliest types of cancer worldwide.[5,6]

An underlying chronic liver disease is present in most cases and constitutes the background condition leading to liver cancer[7,8]. The most common risk factors include chronic hepatitis B (HBV) and C (HCV) infections, obesity and metabolic syndrome and chronic alcohol intake. Out of these, hepatitis B and C are the major causes of the morbidity and mortality, with HBV representing the main cause for HCC worldwide. Another risk factor which is increasing in incidence is non-alcoholic fatty liver disease (NAFLD), mainly due to the increase of obesity and metabolic syndrome associated with the western diet. All of these factors cause a chronic injury which induces progressive liver damage leading to cirrhosis and ultimately liver cancer.[6]

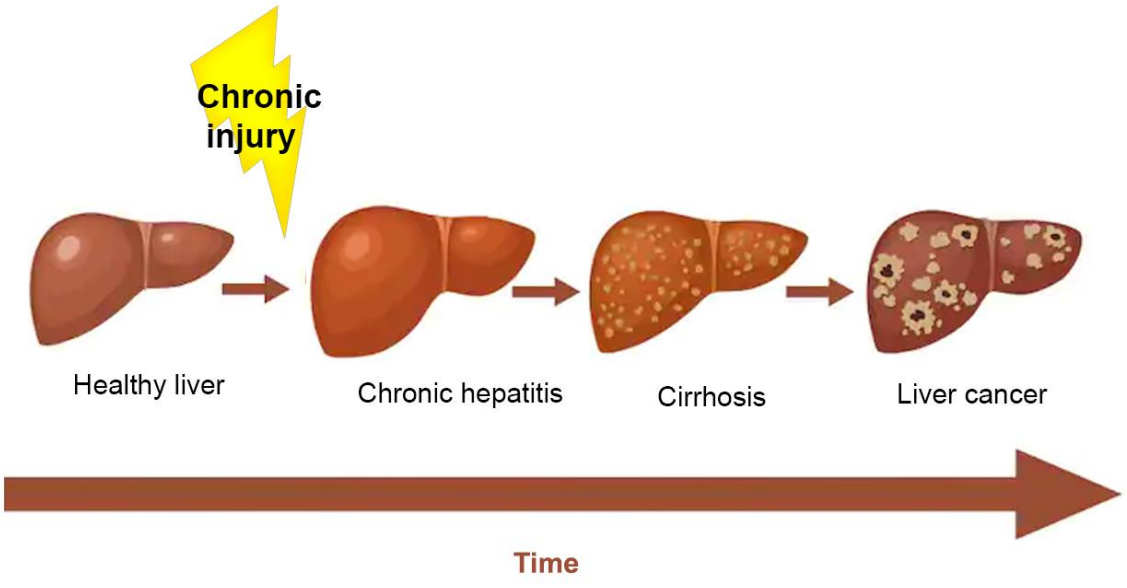


Fig 1. Schematic representation of HCC development

On a genetic level, HCC tumors presents different chromosomal alterations as well as somatic mutations, such as mutations in the *TERT* promoter (60%), *TP53* (30%), *CTNNB1* (30%) and *AXIN1* (10%), none of which can be used in daily medical practice as predictors of therapeutic outcome due to insufficient testing. Moreover, as far as molecular therapies are concerned, alterations involving *VEGFA* and *CCND1/FGF19,* which could be targeted by such therapies, have been reported but are quite rare. These aspects put HCC among the group of solid tumors with few somatic mutations that can be targeted by molecular therapies.[9,10]

Recent progress has allowed the identification of different molecular subtypes of HCC which, although they corelate with clinical features, are not used in clinical practice due to the insufficient proof of their capacity to predict a therapeutic response.[6,11] The molecular subtypes can be divided into two main groups: the proliferation class and the nonproliferation class. The first is seen more frequently in tumors originating on a chronic HBV infection and characterized by an aggressive clinical behavior. This molecular subtype includes poorly differentiated tumors with *TP53* mutations, activation of various oncogenic pathways such as MAPK, MET and mTOR as well as chromosomal instability. Regarding the nonproliferation class, tumors falling in this category present a gene expression signature more similar to that of a normal liver tissue and present more mutations in the *CTNNB1* gene (Fig 2).[12,13]

Depending on the stage, current treatments include surgical resection, liver transplantation, radiofrequency ablation and transarterial chemoembolization. For patients with an advanced disease, systemic therapies are recommended. With most drugs failing in phase 3 trials, Sorafenib remained the only available first line treatment for these patients until the FDA approval of levantinib in 2018. Unfortunately, both drugs show a moderate increase in survival, from 7.9 months for placebo to 10.7 months for sorafenib (SHARP trial), and 13.6 months for levantinib compared to 12.3 months for sorafenib (Levantinib vs Sorafenib non-inferiority phase 3 trial).[14] For second line treatment, the first drug to be approved by the FDA was Regorafenib, which increased survival of patients showing tumor progression under sorafenib from 7.8 to 10.6 months (RESORCE trial).[15]

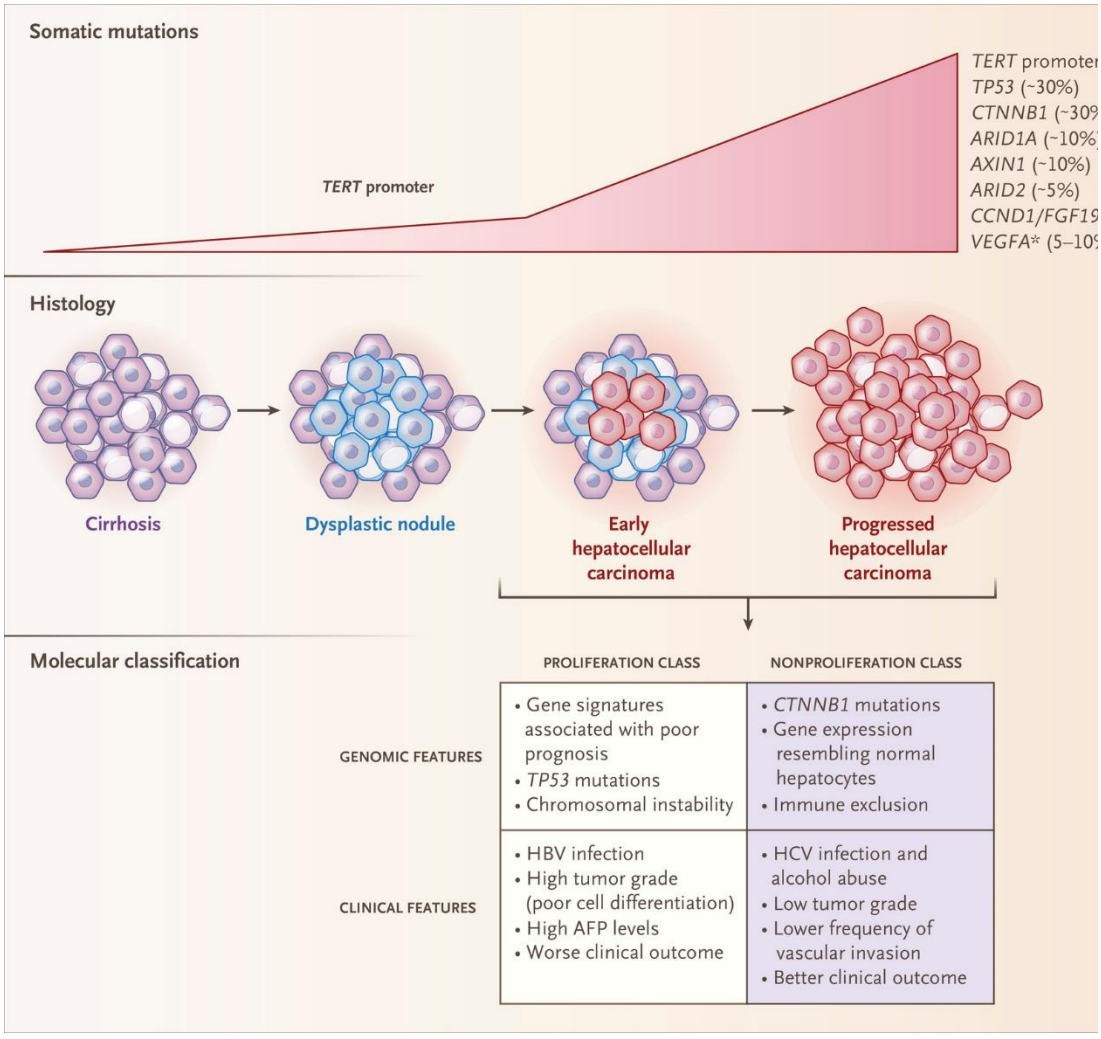


Fig 2. Molecular features of hepatocellular carcinoma[12]

Considering the unsatisfactory results obtained with the current therapeutic procedures (including surgical, locoregional and pharmacological therapies), the unmet clinical need is the identification and correction of the mechanisms leading to HCC occurrence in patients at risk. Although chronic liver disease, especially in the cirrhotic stage is the most common background condition, the explanation why some patients will develop HCC while others will not, despite having common risk factors, remains to be clarified.

The role of mitochondrial metabolism in cancer

The process of malignant transformation describes the transformation of a normal cell into a neoplastic one, which by gaining additional alterations, achieves the limitless potential to replicate, disseminate and metastasize. In probably what is the most influential review in the field of oncology, Hanahan and Weinberg defined six major hallmarks of cancer: increased proliferation, insensitivity to growth suppression signals, resisting apoptosis, unlimited replicative potential, sustained angiogenesis, and tissue invasion and metastasis. Later, in their second review, they add two new hallmarks: reprogramming of metabolism and the capacity to evade immune response (Fig 3).[16–18]

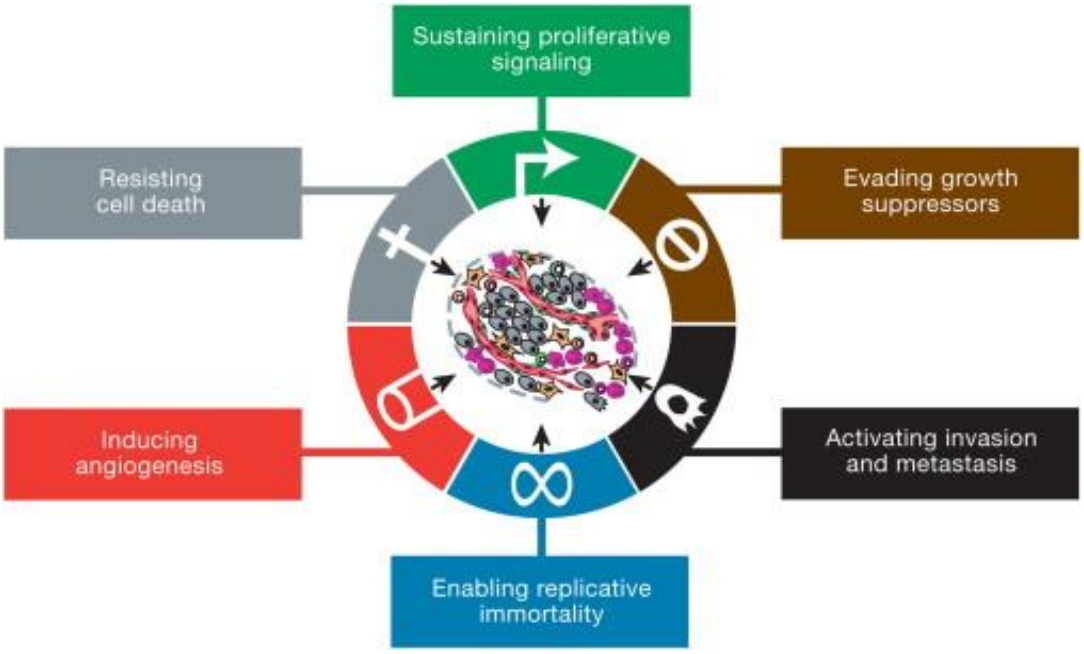


Fig 3. The hallmarks of cancer.[17]

Otto Heinrich Warburg was the first to reveal that cancers go through metabolic reprogramming by observing that they are characterized by an elevated level of glucose uptake and its conversion to lactate through glycolysis even in the presence of oxygen.[19] The phenomenon called “aerobic glycolysis” would be later termed the Warburg effect. This discovery allowed the development of the 2-[^{18}F]fluoro-2-deoxy-D-glucose (^{18}F-FDG) positron emission tomography (PET), an extensively used imaging technique in the clinics for the detection and follow-up of tumors.[20]

For a long period of time cancer cells have been thought to rely only on glycolysis to supply their energetic demands. Thus, the metabolism of cancer cells has been thought to be independent of the mitochondrial status, as they were considered insufficient to satisfy the energetic demands of a cell with such an accelerated proliferation rate. Moreover, the Warburg effect suggested that the increased glycolysis in cancer cells was in fact due to damaged mitochondria. Later, different studies have shown that many tumors still retain a functional mitochondria which allows them to use both glycolysis and oxidative phosphorylation (OXPHOS) in order to adapt to the harsh microenvironment conditions.[21]

Mitochondria have been shown to contribute to tumoral development in different ways. Apart from the metabolic support in terms of ensuring appropriate redox balance and Krebs cycle functioning for anaplerotic biosynthetic reactions, mitochondria through the electron transport chain represent the main cellular source of ROS, which may exert mitogenic activity.[22] By transferring an electron from their redox core to molecular oxygen, respiratory complexes I and III are capable of generating superoxide anions. In normal conditions, these molecules are rapidly converted to H_2O_2 by two superoxide dismutases, namely SOD1 located in the intermembrane space and SOD2 in the mitochondrial matrix.[23]

In a physiological state, ROS are known to play an important role as signaling molecules, influencing cell proliferation and differentiation, but excessive levels can lead to cellular damage resulting in cell death. Nearly all cancers present high rates of ROS and also in parallel, increased levels of antioxidant enzymes in order to maintain the intracellular balance required for its function and survival.[24] It is known that low concentrations of ROS exert a mitogenic activity by promoting cell survival and proliferation, while intermediate levels lead to cell cycle arrest and differentiation. But at high levels, ROS can lead to oxidative damage to different cellular components including DNA with the consequent generation of mutations. This phenomenon is thought to occur at a higher rate in the mtDNA because of its proximity to the source of ROS production and its lack of the complex protective and repair mechanisms that the nuclear DNA exploits.[25,26]

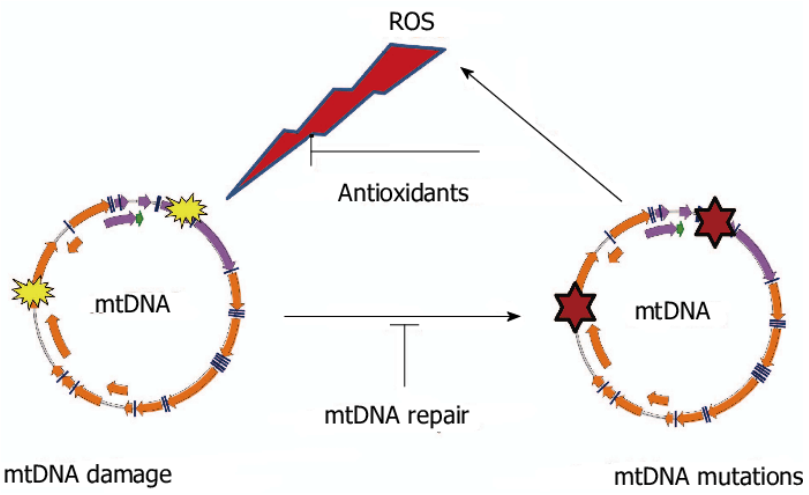


Fig 4. The vicious cycle of ROS induced DNA damage. mitochondrial dysfunction with increased ROS production can induce mtDNA injury with the generation of mutations

Growing tumors present a complex metabolic rewiring which does not resume only to the increased uptake of glucose and increased glycolysis. Cancers have also the capacity to oxidize glutamine in order to produce energy via the Krebs cycle and OXPHOS chain, as well as to process it reductively for the synthesis of fatty acids.[27]

Metabolic reprogramming has also been seen in the case of HCC tumors. These tumors present the characteristic Warburg effect with the increased glucose uptake by upregulating glucose transporters GLUT1 and GLUT2 and an increased glycolytic rate.[28] Another alteration seen is the upregulation of hexokinase 2 isoform which is the first enzyme of the glycolytic pathway, its expression being associated with the pathological stage. Moreover, targeting of this enzyme has shown to increase the response to Sorafenib of HCC cells.[29]

The metabolic plasticity of tumor cells allows them to adapt to conditions such as hypoxia and decreased availability of nutrients. In normal cells this would activate apoptosis resulting in cell death. However, in cancer cells the apoptotic threshold is increased by different mechanisms such as BCL2 overexpression. This increased resistance plays an important role when aiming to treat tumors, since the purpose of chemotherapeutics agents is to induce the death of malignant cells or a permanent growth arrest. Therefore, mitochondria play also a role in the cancer response to therapy, not only in its progression. Such an example can be seen in the case of melanoma cells which present a mutation in the

BRAF gene (*BRAF*^{V600E}). Targeting of this mutation with vemurafenib is associated with a switch to a more oxidative metabolism in order to gain resistance to the treatment. This allows cells to be vulnerable to the inhibition of the electron transport chain using honokiol which can restore the response to vemurafenib.[30] The switch from glycolysis to OXPHOS can be seen also in the case of pancreatic cancer driven by the *KRAS*^{G12D} mutation, as well as in the ablation of MYC/KRAS or MYC/ERBB2 in breast cancers.[31,32]

Indeed, in the case of HCC, mitochondria have also been shown to play an important role in the resistance to sorafenib. For example, a study has shown that the activation of OXPHOS using dichloroacetate can overcome HCC's resistance to sorafenib.[33]

Considering the central role held by mitochondria in metabolic reprogramming of cancer cells as well as in the control of cell death, targeting this organelle in an attempt to cure tumors has gained popularity in the past years with positive results.[22,34] Another example of the role of mitochondria in HCC metabolism can be seen in the case of studies with metformin. Metformin, an inexpensive drug frequently used in the treatment of diabetes, has been shown to inhibit respiratory complex I (CI), reducing tumorigenicity[35–39]. A great body of research shows an association between the use of metformin and reduced cancer incidence in HCC.[40, 41] Metformin has been shown to inhibit HCC cells proliferation and migration[42] as well as the recurrence of metastasis in mouse models [43] Despite these positive results, some findings have failed to show a benefit of the treatment,[44,45] thus further clarifications are needed in order to understand the role of mitochondria in the development and evolution of HCC.

Mitochondrial genetics

Mitochondria are highly specialized cytoplasmatic organelles best known for their energy generating function in the form of ATP. They are semiautonomous organelles as they contain their own circular genome with a different genetic code than the nuclear one as well as a replication pattern independent from the cell cycle. Mitochondrial DNA (mtDNA) is a double stranded circular molecule with a size of approximately 16.6 kb[46]. It comprises a heavy (H) strand and a light (L) strand, names based on the asymmetrical G/C distribution which renders the H strand of a "heavier" molecular weight. The L strand was defined in the original publication as the sense strand containing the coding sequence of most of the mitochondrial genes. Therefore, the RNAs corresponding to these genes are transcribed from the H strand.[47,48]

The mitochondrial genome contains 37 genes of which thirteen are coding for proteins participating in the OXPHOS chain (Fig 5). The rest encode for tRNAs (22 genes) and two rRNA (the 16s for the large subunit of the mitochondrial ribosome and the 12s for the small subunit). As opposed to nuclear genes, mitochondrial genes do not have introns. Apart from the non-coding nucleotides which separate contiguous genes, mtDNA has only one non-coding regulatory region, the displacement loop (D-loop). This region represents the site where the mtDNA replication and translation begin.[49,50]

The inheritance of mtDNA has long been considered to follow exclusively a maternal pattern, as sperm mtDNA is removed by ubiquitination during the formation of the mammalian zygote. Because of the mtDNA is inherited only from the mother, usually in a homoplasmic state, this small molecule offers a valuable opportunity to explore genealogical relationships and track populational migration. Having an known mutational rate which is 5-10 times higher than that of the nuclear genome, mitochondria represent an important tool to distinguish evolutionary patterns of the human population. The lack of recombination of the mtDNA allows all variants to be organized in the shape of a phylogenetic tree, with all branched being traced back to the first ever mtDNA known as the mitochondrial Eve. These branches containing a specific combination of mtDNA variants could be labelled by a distinctive migration pattern on agarose gels upon restriction enzyme digestion. Therefore, these branches were named using alphabetic labels which are known as mtDNA haplogroups.[51,52]

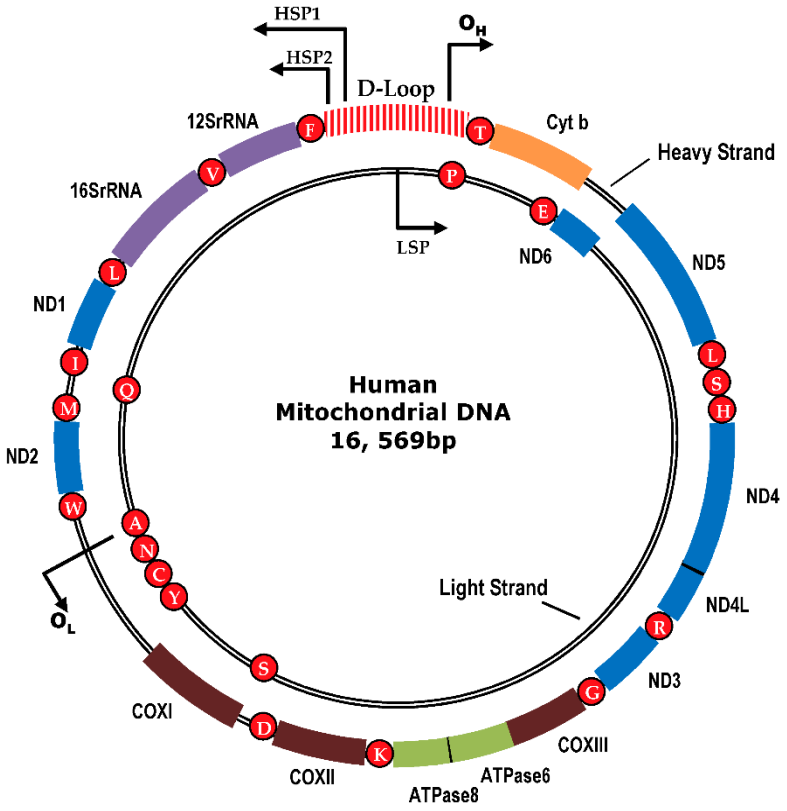


Fig 5. Structure of the human mtDNA. *MT-RNR1* and *MT-RNR2* code for the 12S and 16S rRNAs subunits. *MT-ND1* to 6 genes code for NADH dehydrogenase subunits. *MT-CO1*, *MT-CO2*, *MT-CO3* - cytochrome oxidase subunits 1,2 and 3. *MT-ATP6* and *MT-ATP8* genes codes for mitochondrial ATPase subunits 6 and 8. *MT-CYB* gene codes for cytochrome b. The 22 tRNAs encoding genes are in red. Promoters *HSP1* and *HSP2* transcribe genes on the H-strand and *LSP* on the L-strand.[53]

Unlike the nuclear DNA, mtDNA is found in multiple copies with cells having between 100 to 10000 copies depending on cellular energetical requirements. Homoplasmy defines then state in which all the mtDNA copies are identical. However, due to different factors contributing to an increased risk of mutational events, mtDNA can acquire different mutations which often co-exist together with the wild-type copies. This mixture of different mtDNA molecules which can be in various proportions, is defined as heteroplasmy. Because of the high number of mitochondrial DNA copies in a single cell, some wild type, some mutated, the percentage of any mtDNA variant/mutation can vary over time. Thus, a normal mitochondrial function can be maintained even in the presence of a pathogenic mutation, as long as the ratio of wild type mtDNA is high enough. When this ratio favors the mtDNA copies containing pathogenic mutations, the mitochondrial dysfunction caused by the

alteration can become apparent and cause a visible phenotype. This phenomenon is known as the phenotypic threshold effect and it is characterized by a high variability between different mtDNA mutations as well as between different affected tissues. This threshold effect has been shown to vary between 20 and 98% for different mitochondrial diseases.[54]

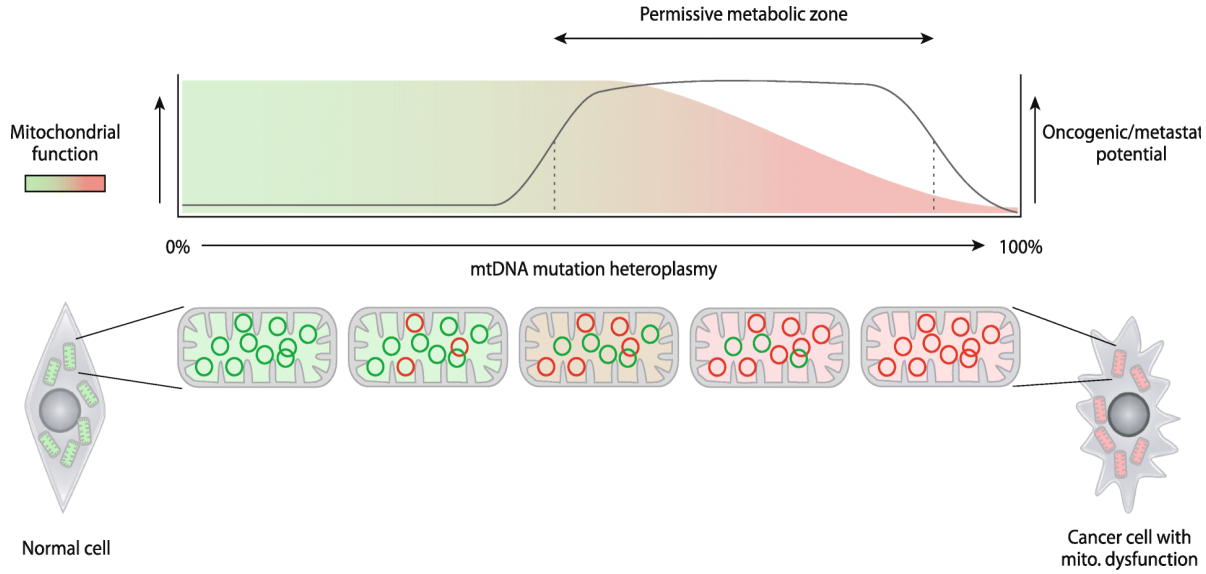


Fig 6. mtDNA polyploidy. Green represents wild type mtDNA while red shows mutated mtDNA. Mutations are required to reach a certain threshold in order to have an impact on the mitochondrial function[55]

Mitochondria generate ATP through the OXPHOS reaction which is carried out in the internal membrane by the electron transport chain (ETC). The ETC comprises five protein complexes: NADH-ubiquinone oxidoreductase - complex I, succinate-ubiquinone oxidoreductase - complex II, ubiquinone-cytochrome c oxidoreductase - complex III, cytochrome c oxidase - complex IV, and ATP synthase - complex V (Fig 6). During evolution, the majority of the genes coding for mitochondrial proteins, including those of the ETC have moved in the nuclear DNA (Fig 7).

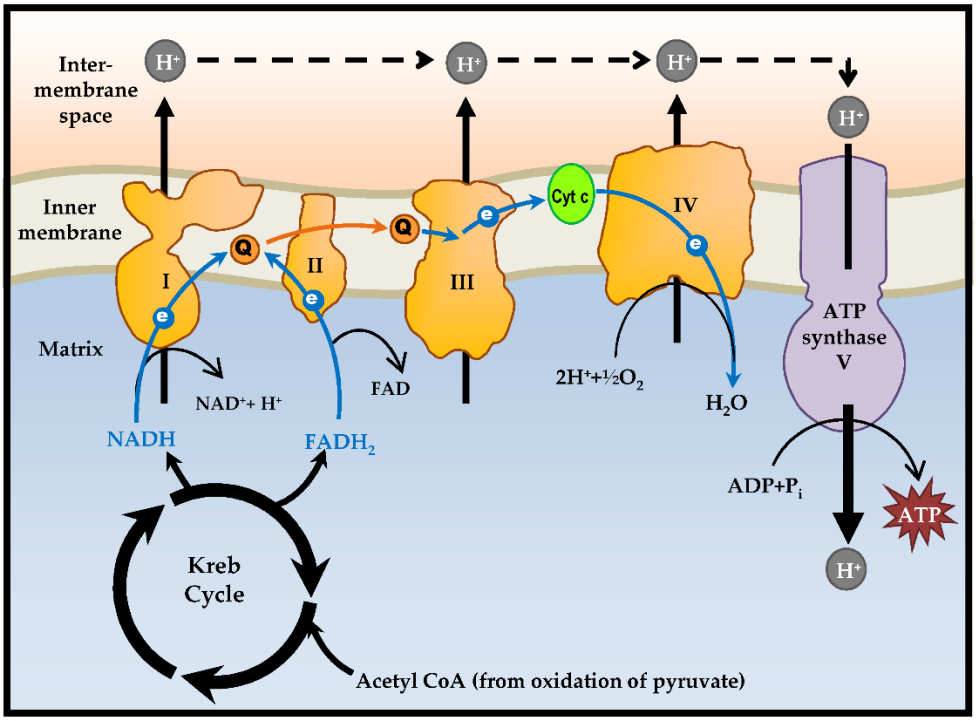


Fig 7. The structure of the electron transport chain.

Apart from ATP production, mitochondria carry out other functions such as maintaining the Ca^{2+} homeostasis, fatty acid beta-oxidation and redox signaling through the generation of reactive oxygen species (ROS). They also participate in certain steps of the heme synthesis though the biogenesis of iron-sulphur clusters as well as in pyrimidine synthesis.[56,57]

Mitochondrial mutations in cancer

Mitochondrial mutations have been described to occur in nearly 60% of all cancer types, but their functional role in tumorigenesis remains unclear. The lack of highly efficient repair mechanisms and the vicinity to the site of reactive oxygen generation renders mitochondrial DNA susceptible to a higher rate of accumulating mutations than the nuclear genome.[58]

The majority of mtDNA mutations found are somatic, but in some cases germline variants which in the normal tissue have a low heteroplasmy can become positively selected in the tumoral tissue.[59]

The most reported variants in cancer studies are located in the D-loop region of the mtDNA genome. This region is known to be highly polymorphic, and many authors fail to consider a proper stratification of the discovered variants, as the simple presence of a nucleotide change is not sufficient for its association with a cancer type. Thus, such studies need to be read with caution. The second most frequent mutations seen are those in genes coding for CI subunits, but this could be mainly due to the fact that most of the mtDNA genome codes for CI subunits. To date, no preferential mutational hotspot site in the mtDNA has been seen in cancer cells.[60]

In cancers, the most frequently seen variants occurring in protein-coding genes are missense. This accumulation of variants with a functional impact has led to the idea that they might offer tumors a selective advantage. Contrarily, highly pathogenic mutations are contra-selected from tumors supporting the current accepted knowledge that functional mitochondria is needed in order to ensure the progression of the tumors. Unfortunately, there is a low availability of functional studies which characterizing completely the impact of mtDNA mutations which limits the current understanding of this topic.[61] The association between mtDNA mutations and cancer development is a difficult task also due to the presence of heteroplasmy, which adds another layer of complexity. The current accepted view is that mtDNA mutations are not capable of inducing malignant transformation but can play an important role as modifiers of the tumorigenic potential. For example, mutations leading to a decreased efficiency of OXPHOS could be beneficial when tumors go through hypoxia but may impair the metastatic capacity of the tumoral cell bearing it.[62,63]

Another issue contributing to the complex roles of mtDNA mutations is the fact that the same mutation can have opposite effects depending on the context in which it occurs.

For example, the mtDNA mutation present in the 143B osteosarcoma cell line has been shown to increase glycolytic metabolism and promote tumor progression when found in a heteroplasmy state, while increasing apoptotic rate and decreasing tumoral growth when found in homoplasmy.[64,65]

Mitochondria play important roles in the metabolism of glutamine and lipids. Glutamine is used in the anaplerotic reactions of the TCA cycle and for the synthesis of nucleotides while the reprogramming of lipid metabolism is required for the synthesis of the lipids required to construct the cellular membranes. Because these processes are influenced by the functional status of the OXPHOS chain, mtDNA mutations can play an important role in the capacity of cancer cells to adapt to metabolic stress.[66,67]

In the case of HCC, there is a limited number of studies tackling the role of mtDNA mutations in hepatocarcinogenesis. Firstly, many reports limit themselves to sequence only the D-loop regulatory region.[68] This may lead to the finding of polymorphisms rather than to true pathogenic variants. Next, due to the lack of a proper classification and stratification of the variants, often studies consider all events found in the tumoral tissue and not in the healthy tissue to be potentially pathogenic. This issue is also present in other studies investigating the role of mtDNA mutations in different cancers as well.[69]

An interesting finding is that of a common deletion affecting CI which is purified from the HCC tumors, suggesting its probable unfavorable impact on tumor progression.[70] Other findings report a lower number of mtDNA copies which also seems to corelate with the clinical features.[71] Overall, the scarce data available, combined with the lack of a careful selection of the variants contributes to an unclear picture of the role of mtDNA mutations in HCC.

The Cancer Genome Atlas (TCGA): a valuable resource for ‘omic studies

The Cancer Genome Atlas (TCGA) is a joint effort between the National Cancer Institute and the National Human Genome Research Institute which began in 2006. By joining multiple research centers from diverse disciplines, the consortium managed to collect and analyze more than 20.000 primary cancers and their matched normal samples from 33 cancer types. Containing a wide array of large scale ‘omic techniques such as genomic, transcriptomic, epigenomic and proteomic data, TCGA is one of the most comprehensive and complete cancer datasets available to date.

The data from the TCGA projects can be divided into two types: Open Access Tier and Controlled Access Tier:

Open Access data tier comprises information which does not pose a risk for the research patient re-identificaton. Therefore, for this particular category the access does not require an approved user certification and are freely accessible through the TCGA Data Portal.

Controlled Access data tier comprises data which are unique to an individual such as Individual germline variant data (SNP .cel files); Primary sequence data (.bam files); Clinical free text fields and Exon Array files. In order to gain access to such data, a researcher must receive approval by the dbGaP Authorized Access.

Aim

The importance of mitochondria in carcinogenic processes is well known in general, but insufficient work is available in the literature which explores the role of mitochondrial metabolism in the instance of HCC development.

Mitochondrial mutations have been known to play a role in the process of carcinogenesis with many works reporting the occurrence of mtDNA mutations in tumors. To date, few studies correctly prioritize the mitochondrial DNA variants discovered. Thus, an extensive analysis of the mitochondrial status in a large dataset of patients may be an opportunity to discover novel mechanisms in the pathogenesis of HCC.

The main objective of this study was to analyze the occurrence and role of mtDNA mutations in HCC tumors and to explore their value as predictive markers for HCC prognosis.

To this purpose we first identified and defined the mitochondrial DNA mutational landscape in the hepatocellular carcinoma cohort from The Cancer Genome Atlas consortium.

By prioritizing and stratifying the variants found based on their potential pathogenic impact, we aimed to analyze the potential impact of such mutations on the evolution of patients with HCC.

Methodology

Bioinformatics analysis of mtDNA mutations

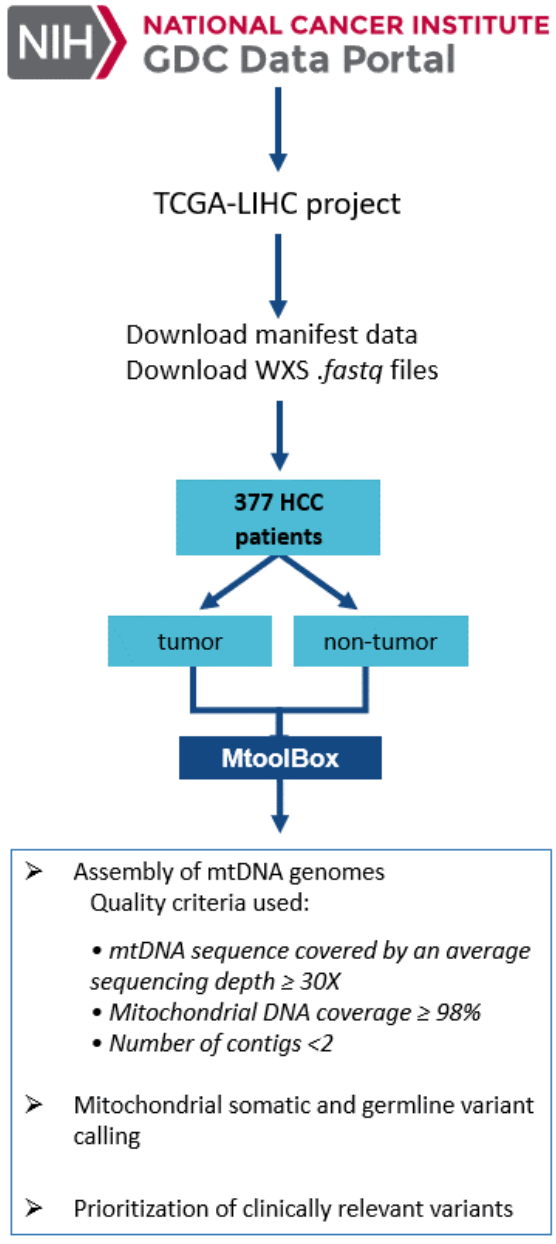


Fig 8. MToolBox analysis workflow

To achieve our goal, we downloaded whole exome sequencing (WXS) data from the TCGA-LIHC project for both tumoral and paired non-tumoral tissues (blood or surrounding liver tissue) after submitting a project request to the NIH.

The *.fastq* files deriving from 377 HCC patients (paired tumor and non-tumor tissues) were processed to reconstruct the mitochondrial genomes using the MToolBox automated pipeline.[72] The process includes the extraction and assembly of mitochondrial DNA fragments from the off-target reads and also discarding the sequences originating from the nuclear sequences derived from mitochondrial DNA, namely the Nuclear mt Sequences (NumtS).

In order to achieve a high level of accuracy of our analysis and avoid false positive results the quality criteria set for this study was as follows:

- an average sequencing depth of the mtDNA sequence ≥ 30X
- Mitochondrial DNA coverage ≥ 98%
- Number of reconstructed mtDNA contigs less than 2

After the variant calling of the discovered events, all variants are given a functional annotation and are prioritized in order

to select those of potential clinical relevance according to the tiers reported in Table 1 and defined, by combining the Disease score and the allele frequency for any possible variants that could occur in any site of the human mitochondrial genome. The disease score for variants occurring in coding for protein genes is estimated acccording to the algorithm described in Santorsola M et al, and is based on different predictors: MutPred, HumDiv and HumVar-trained PolyPhen-2 models, SNPs&GO, PhD-SNP and PANTHER.[73–75] The disease score for tRNA variants is estimated according to the algorithm reported in Di Roma MA et al and then improved in HmtVar. For tRNA variants the MutPred etc predictors is not considered. This classification criteria can be found in Table 1.

	Non-synonymous variants		tRNA variants	
Tier	**Disease Score**	**Allele Frequency**	**Disease Score**	**Allele Frequency**
Polymorphic	DS < 0.43	AF > 0.003264	DS < 0.35	AF > 0.005020
Likely Polymorphic	DS < 0.43	AF ≤ 0.003264	DS < 0.35	AF ≤ 0.005020
Likely Pathogenic	DS ≥ 0.43	AF > 0.003264	DS ≥ 0.35	AF > 0.005020
Pathogenic	DS ≥ 0.43	AF ≤ 0.003264	DS ≥ 0.35	AF ≤ 0.005020

Table.1. HmtVar pathogenicity predictions

Pairwise comparison between blood/normal solid tissue and tumor was performed in order to identify the potentially germline and tumor-specific somatic mtDNA variants. The schematic analysis workflow of this study is presented in Fig 8.

To achieve a correct estimation of the tumoral mitochondrial mutational burden, heteroplasmic fractions of the variants discovered was adjusted using the estimated tumor purity value from TCGAbiolinks R/Bioconductor package. This value combines four parameters: expression profiles from 141 immune genes and 141 stromal genes, somatic copy-number variantion data (not available for the LIHC cohort) and 44 non-methylated immune-specific CpG sites.[76]

Complementary analysis involving other somatic mutations was carried out using the Mutation Annotation Format (MAF) files available in the Genomic Data Commons Data Portal (CGD)[77]. Explorative data visualization and mining was done using the Orange3

software package from Anaconda[78]. Graphical representations and analyses were constructed in R/Bioconductor using the appropriate packages such as TCGAbiolinks[79,80,81] and the CRAN TCGA2STAT[82] and in GraphPad Prism 8.

Statistical analysis using appropriate tests was done by using IBM SPSS Statistics software version 20. For continuous variables Student's T-test was used while for categorical variables we used Chi-Square test of Fisher's exact test. For survival analysis, the Kaplan Meyer curve analysis was employed. For multivariate analysis the Cox Proportional Hazards Regression analysis was used. The results were considered significant if the p-value was less than 0.05.

Immunohistochemistry staining

Paraffin embedded tissues of patients diagnosed with HCC were obtained from the archives of the Pathology department of Sant'Orsola-Malpighi Hospital, Bologna, Italy. Serial cuts of 3um were performed from each paraffin block. Sections were deparaffinized, rehydrated, and retrieved using a Tris-EDTA pH 8 solution (20 min at 98*C). Primary antibodies were purchased from Abcam:

For complex I - NDUFS4 mouse diluted 1/1000 (Catalog no. AB55540)

For complex III – Core 2 (UQCRC2) mouse diluted 1/1500 (Catalog no. AB14745)

For complex IV – COX I mouse diluted 1/1000 (Catalog no. AB14705)

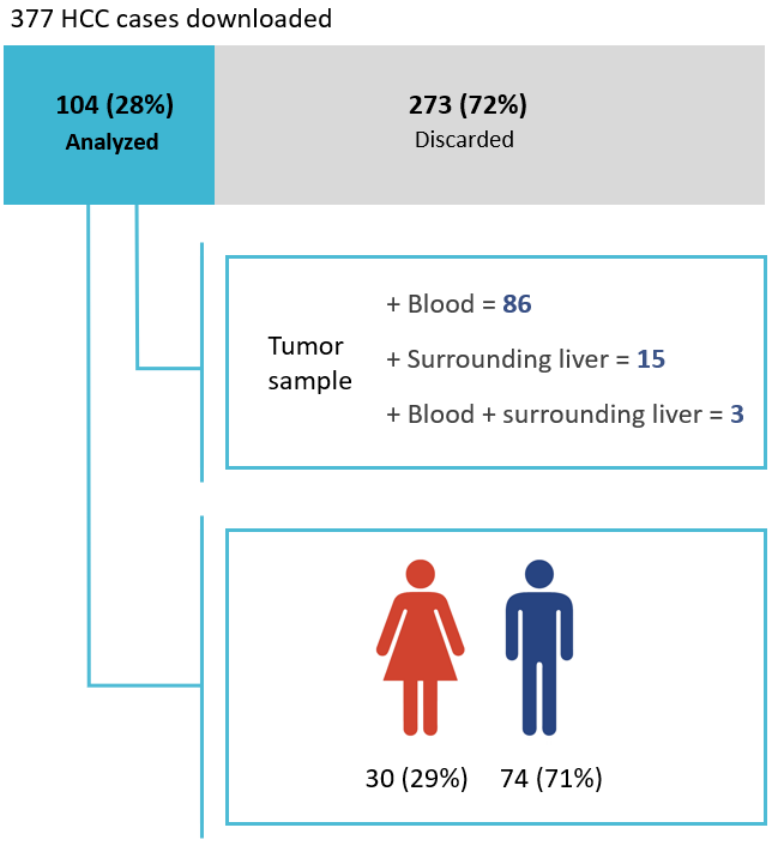
For complex V - ATPase5a mouse diluted 1/2000 (Catalog no. AB14748)

Incubation for the primary antibody was 1h at room temperature. The reaction was developed with a DAB-Peroxidase Substrate Solution according to the manufacturer's instructions. Sections were counterstained with hematoxylin, dehydrated, and mounted.

Results

Evaluation of the TCGA LIHC cohort revealed 104 cases eligible for mtDNA mutation analysis

The TCGA-LIHC cohort comprises a total of 377 HCC cases, all of which were run through the MToolBox pipeline. A number of 104 patients fulfilled our criteria and were further analyzed. These included patients for which data were available for both tumor samples and for their paired non-tumoral control.



The majority of the cases were males representing 71% of total. Mean age for diagnosis for males was 58.5±11.8 years while for females it was 58.68±15.3 years.

Disease stages as defined by the American Joint Committee on Cancer (AJCC) cancer stage manual was predominantly represented by the early stages with almost 70% of the patients having either a stage I or II diagnosis (Fig 9).

Almost 80% of the patients had a documented risk for developing HCC, the most frequent one being alcohol consumption which was reported by 31 patients. All risk factors are presented in Table 2.

The mean minimum reading depth for tumoral samples was 30.5x while the maximum was 429.5x. The non-tumoral samples were represented by either a blood derived normal, solid normal tissue or both. Mean reading depths of these samples spun between a minimum of 30x and a maximum of 642.2x. Detailed quality information regarding the samples can be found in Annex 1.

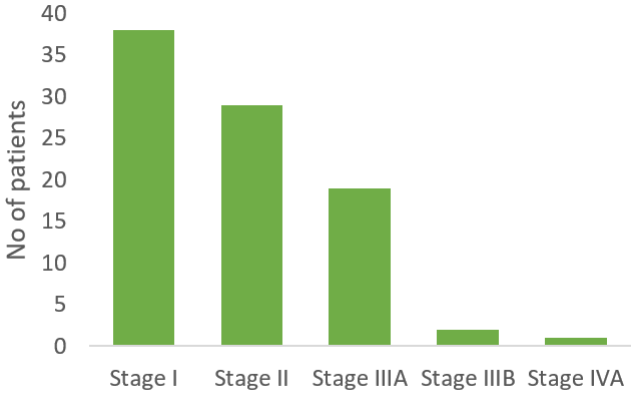


Fig 9. Number of patients for each HCC stage. Stages are according to the AJCC classification of cancer stages

	No of patients	Percent
Alcohol consumption	23	22
Alcohol consumption|Hepatitis B	6	5.8
Alcohol consumption|Hepatitis C	2	1.9
Hemochromatosis	3	2.9
Hepatitis B	22	21.2
Hepatitis B|Hepatitis C	1	1.0
Hepatitis C	8	7.7
Hepatitis C|Hemochromatosis	1	1.0
Non-Alcoholic Fatty Liver Disease	6	5.8
No History of Primary Risk Factors	24	23.1

Table 2. Risk factors in the TCGA HCC cohort

Non-tumor specific mutations in HCC

A total of 1216 mtDNA germline variants recognized by the Revised Cambridge Reference Sequence (rCRS), the Reconstructed Sapiens Reference Sequence (RSRS) and macro-haplogroup consensus sequence (MHCS)[83] were identified. Among these, a total of 821 variants defined the patients haplogroups and were not considered in the subsequent analysis steps. A general view of the distribution of the germline variants based on their heteroplasmy can be found in Fig 10.

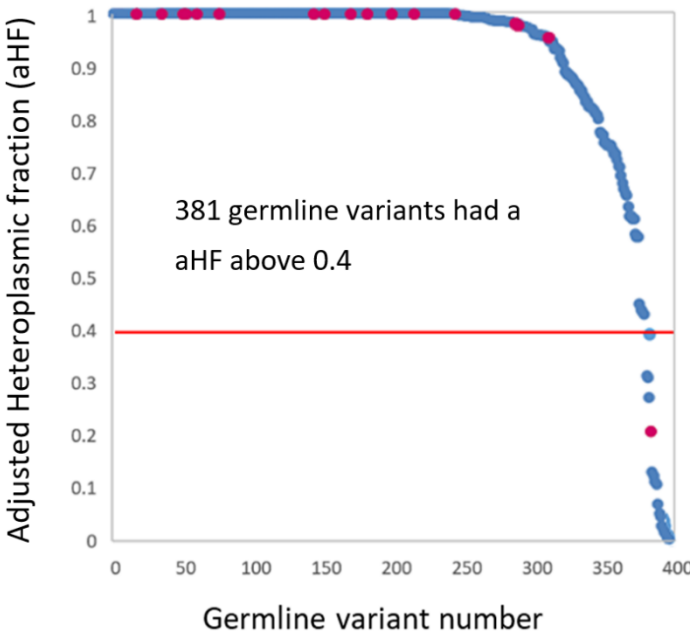


Fig 10. Distribution of the 395 germline variants found in the TCGA-HCC cohort. Blue dots: polymorphic/likely polymorphic variants, Purple dots: pathogenic/likely pathogenic variants. Red line: fixed heteroplasmy threshold.

After fixing a heteroplasmic fraction (HF) threshold of 0.4 to the remaining 395 variants, 381 variants were filtered, among which 132 (35%) were synonymous, 59 (16%) non-synonymous and 189 (50%) were variants belonging to the non-protein coding class. Five variants were not previously reported in the HmtVar database. Summary of the characteristics of the variants found can be seen in Table 3.

After applying the HmtVar pathogenicity criteria previously described, variants were divided into four classes: pathogenic, likely pathogenic, likely polymorphic and polymorphic. A total of 13 non-synonymous and three tRNA pathogenic/likely-pathogenic mutations were found in 15 (14%) patients (Fig 11).

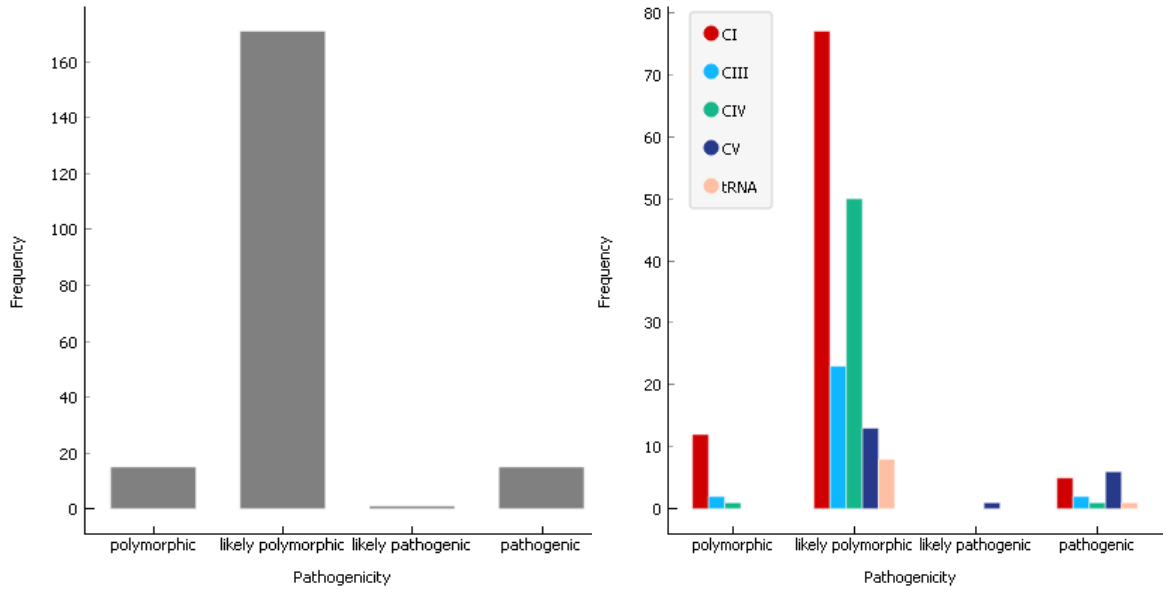


Fig 11. Distribution of germline variants in the LIHC TCGA cohort. The majority of discovered events are classified as benign variants (likely polymorphic and polymorphic).

Patient ID	Variant	AF	HF non-Tu	HF Tumor	Locus	Codon Position	AA Change	AA Var	DS	Pathogenicity
LIHC_16	8505G	0.000000	1	1	MT-ATP8	2	Y47C	0.0167	0.55	pathogenic
LIHC_19	12535T	0.001017	1	1	MT-ND5	1	H67Y	0.0141	0.72	pathogenic
LIHC_24	8516C	0.000218	1	1	MT-ATP8	1	W51R	0.0073	0.89	pathogenic
LIHC_37	14180C	0.003246	1	1	MT-ND6	2	Y165C	0.135	0.78	pathogenic
LIHC_51	15617A	0.001163	1	1	MT-CYB	1	V291I	0.0059	0.76	pathogenic
LIHC_52	15383C	0.000460	1	1	MT-CYB	1	S213P	0.0054	0.77	pathogenic
LIHC_65	8393T	0.003342	0.94	1	MT-ATP8	1	P10S	0.124	0.67	likely_pathogenic
LIHC_76	5979A	0.000654	1	1	MT-CO1	1	A26T	0.0132	0.73	pathogenic
LIHC_87	9163A	0.001308	1	0.98	MT-ATP6	1	V213I	0.0022	0.81	pathogenic
LIHC_88	5095C	0.000460	1	1	MT-ND2	2	I209T	0.0122	0.65	pathogenic
LIHC_89	12965C	0.000048	1	0.98	MT-ND5	2	L210P	0.0022	0.72	pathogenic
LIHC_99	9163A	0.001308	1	0.95	MT-ATP6	1	V213I	0.0022	0.81	pathogenic
LIHC_100	8944G	0.000097	1	1	MT-ATP6	1	M140V	0.0023	0.57	pathogenic
LIHC_58	15924G	0.033061	1	1	MT-TT				0.4	likely_pathogenic
LIHC_65	15927A	0.007581	0.97	1	MT-TT				0.35	likely_pathogenic
LIHC_67	5814C	0.003028	1	1	MT-TC				0.75	pathogenic

Table 3. Germline mutations found in HCC patients. HF: MToolBox output heteroplasmic fraction; AF: allele frequency; AA Var: aminoacid variability; DS: disease score. *Nt Var, AA Var and AF are estimated on the basis of the human mitochondrial genomes annotated in GenBank and then available through the HmtDB database and hence in HmtVar.*

Tumor specific mutations in HCC

Tumor specific mitochondrial variants were defined as new events reported only in the tumoral MToolBox output and not present in the non-tumoral output.

We detected a total of 302 mtDNA variants exclusively in 89 (86%) tumor samples. Fifteen patients did not present any mitochondrial variants in their tumoral tissue.

A general overview of all the somatic mtDNA variants present in the TCGA-HCC cohort is illustrated in Fig 12.

Variants were grouped into protein coding and non-protein coding variants. Protein coding variants were either synonymous when the aminoacid remained the same, non-synonymous (missense) when the nucleotidic change determined also an aminoacid change and nonsense if the nucleotide alteration induced a premature stop codon.

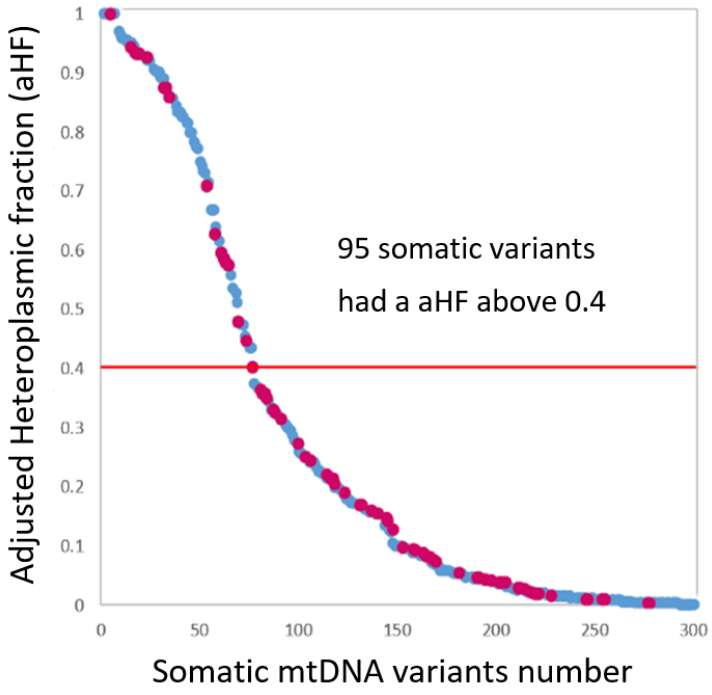


Fig 12. Distribution of all somatic variants found in the TCGA-HCC cohort. Blue dots: polymorphic/likely polymorphic variants, purple dots: pathogenic/likely pathogenic variants. Red line: fixed heteroplasmy threshold.

The TCGA-LIHC cohort presented a total of 33 (11.7%) synonymous variants found in 19 patients, and 70 (24.7%) missense variants in 45 patients, while the vast majority of the somatic events were present in the non-coding part - 180 (63.6%) and reported in 72 cases. From this category, the variants occurring in the D-loop or rRNAs were excluded from the analysis where pathogenicity predictions were required, as for these variants the the algorithm to estimate the disease score has not been designed.

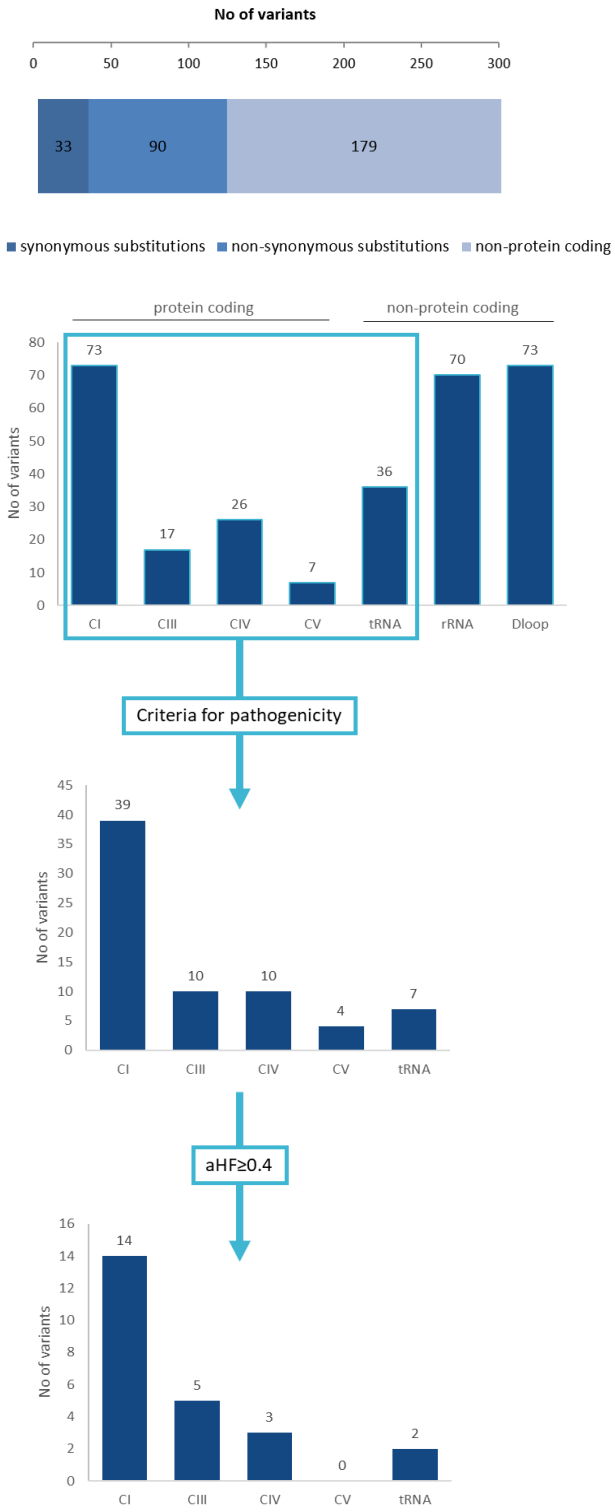


Fig 13. Filtering and selection of somatic mtDNA variants found in the TCGA-LIHC tumors

After applying the pathogenicity scores and filtering, a total of 70 pathogenic mutations are left with the main majority being located in genes coding for CI proteins. Although classified as pathogenic or likely pathogenic, the simple presence of a mitochondrial mutation is not sufficient to induce a phenotype. Thus, a further filtering was applied namely an adjusted heteroplasmic fraction of 0.4 (Fig.13). The remaining 24 mutations were distributed in 22 (21%) patients. Their characteristics are resumed in Table 4.

One patient presented a mutation at position 6579A in MT-CO1 which leads to the gain of a premature stop codon at position 226 out of 513. This leads to the formation of a truncated protein. For stop-gain mutations, MToolBox does not calculate a disease score thus it does not estimate the pathogenicity of this particular variant. In order to gain insight in the possible impact of this mutation, we explored the phyloP and PhastCons scores which measure the evolutionary conservation of a site and the probability of negative selection, respectively. The resulted phyloP positive value of 9.4 means that this region is well conserved and the PhastCons value of 1 predicts this mutation as deleterious.

Mitochondrial genomes are prone to accumulate high levels of variants in tumoral cells. These mutations may serve to promote tumor cell proliferation and allow tumors to adapt to the microenvironment.[84] We analyzed the distribution patterns of the somatic variants found. We observed that the major proportion observed in HCC tumors were missense variants. This suggests that tumoral cells have an increased tolerance for the accumulation of these variants. (Fig 14).

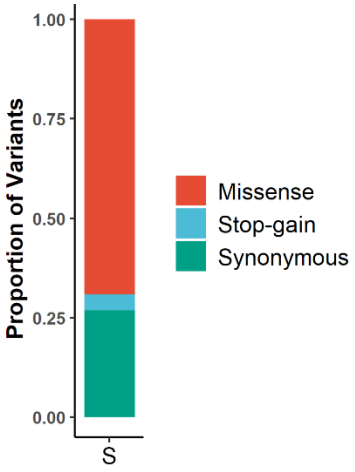


Fig 14. Proportion of variants in HCC tumors

Upon further analysis we noticed that tumors show an increased number of variants classified as pathogenic (Fig 15).

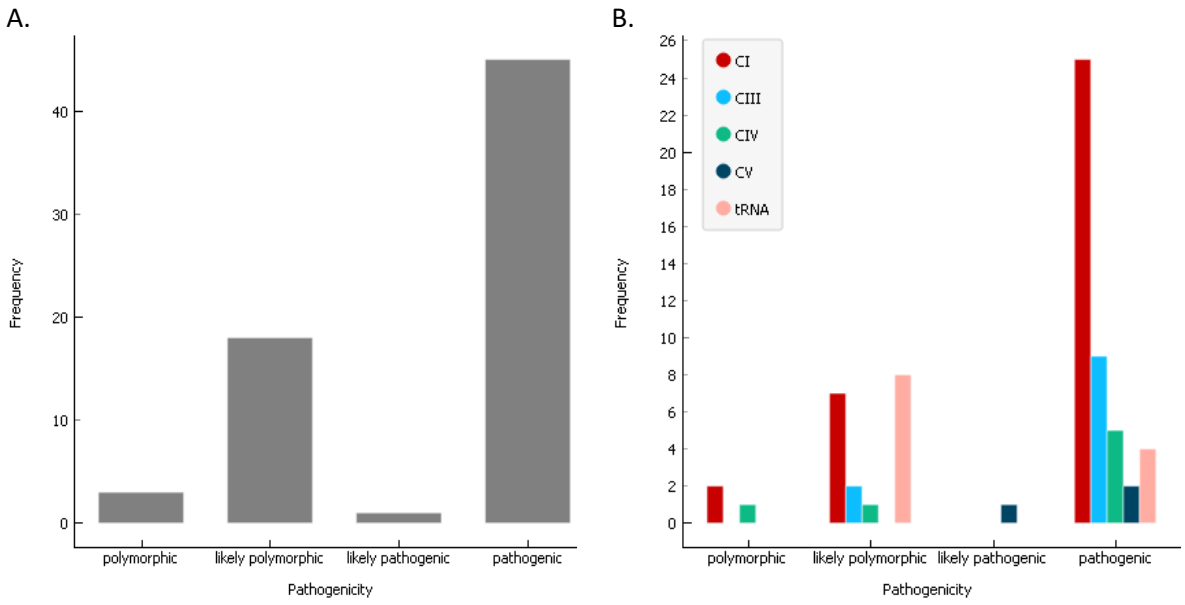


Fig 15. Distribution of somatic variants in the LIHC TCGA cohort. Mitochondrial complexes I-V are abbreviated as CI-V

The majority of the somatic pathogenic mutations falls into genes coding for CI (57%) and CIII (20%) subunits.

Patient ID	HmtVar ID	Variant Allele	HF	aHF	Location	Gene	Nt Var	AF	AA Change Type	AA Change	Aa Var	DS	Pathogenicity
LIHC_67	2723	3946A	0.40	0.47	CI	MT-ND1	0.0004	0.0000	missense	E214K	0.0030	0.91	pathogenic
LIHC_40	24286	12769A	0.58	0.65	CI	MT-ND5	0.0001	0.0000	missense	E145K	0.0053	0.91	pathogenic
LIHC_92	2666	3922A	0.48	0.58	CI	MT-ND1	0.0000	0.0000	missense	E206K	0.0013	0.89	pathogenic
LIHC_44	5220	4950C	0.36	0.53	CI	MT-ND2	0.0001	0.0000	missense	S161P	0.0018	0.83	pathogenic
LIHC_29	3176	4139T	0.87	0.92	CI	MT-ND1	0.0000	0.0000	missense	P278L	0.0014	0.82	pathogenic
LIHC_104	23905	12613A	0.93	1.00	CI	MT-ND5	0.0012	0.0003	missense	A93T	0.0067	0.82	pathogenic
LIHC_21	20213	11087C	0.59	1.00	CI	MT-ND4	0.0098	0.0021	missense	F110L	0.1348	0.81	pathogenic
LIHC_90	24915	13042A	0.57	0.85	CI	MT-ND5	0.0001	0.0000	missense	A236T	0.0022	0.78	pathogenic
LIHC_63	21667	11711A	0.99	1.00	CI	MT-ND4	0.0009	0.0000	missense	A318T	0.0000	0.78	pathogenic
LIHC_62	20990	11420A	0.92	1.00	CI	MT-ND4	0.0000	0.0000	missense	V221I	0.0000	0.75	pathogenic
LIHC_48	23897	12610A	0.36	0.45	CI	MT-ND5	0.0002	0.0000	missense	V92M	0.0022	0.69	pathogenic
LIHC_77	23442	12419T	0.92	0.99	CI	MT-ND5	0.0000	0.0000	missense	K28M	0.0157	0.69	pathogenic
LIHC_21	27085	13976C	0.25	0.48	CI	MT-ND5	0.0001	0.0000	missense	N547T	0.0059	0.63	pathogenic
LIHC_97	25485	13289A	0.32	0.41	CI	MT-ND5	0.0007	0.0000	missense	G318D	0.0082	0.88	pathogenic
LIHC_70	29546	15005A	0.70	0.82	CIII	MT-CYB	0.0002	0.0000	missense	A87T	0.0000	0.87	pathogenic
LIHC_18	29175	14846A	0.87	0.91	CIII	MT-CYB	0.0000	0.0000	missense	G34S	0.0009	0.87	pathogenic
LIHC_60	30387	15357A	0.58	0.67	CIII	MT-CYB	0.0000	0.0000	missense	G204E	0.0009	0.85	pathogenic
LIHC_28	31262	15699A	0.44	0.65	CIII	MT-CYB	0.0000	0.0000	missense	R318H	0.0009	0.82	pathogenic

Patient ID	HmtVar ID	Variant Allele	HF	aHF	Location	Gene	Nt Var	AF	AA Change Type	AA Change	Aa Var	DS	Pathogenicity
LIHC_104	30692	15489C	0.92	1.00	CIII	MT-CYB	0.0000	0.0000	missense	D248A	0.0006	0.7	pathogenic
LIHC_5	10748	7210C	0.85	1.00	CIV	MT-CO1	0.0000	0.0000	missense	M436T	0.0026	0.83	pathogenic
LIHC_34	16596	9591A	0.94	1.00	CIV	MT-CO3	0.0033	0.0007	missense	V129I	0.0297	0.55	pathogenic
LIHC_8	9251	6579A	0.82	1.00	CIV	MT-CO1	0.0000	0.0000	Stop-gain	G226X	0.0021		
LIHC_44	1065	3243G	0.36	0.52	tRNA	MT-TL1	0.0004	0.0001				0.8	pathogenic
LIHC_42	6640	5549A	0.62	0.66	tRNA	MT-TW	0.0001	0.0000				0.35	pathogenic

Table 4. Pathogenic somatic mutations with an aHF≥0.4 found in the TCGA HCC patients.

HF: MToolBox output heteroplasmic fraction; aHF: tumor purity adjusted heteroplasmic fraction; Nt Var: Nucleotide variability; AF: allele frequency; AA Var: aminoacid variability; DS: disease score. *Nt Var, AA Var and AF are estimated on the basis of the human mitochondrial genomes annotated in GenBank and then available through the HmtDB database and hence in HmtVar.*

Adjusting for the microenvironment component of the HCC tumor increases mtDNA variant detection

The structure of solid tumors comprises a mass of heterogenous cell types which can be divided into two distinct compartments: the parenchyma and the microenvironment. The parenchyma is represented by the tumoral cells while the microenvironment includes a variety of different types of neoplastic induced cells such fibroblasts, blood vessels and immune cells.[85] The proportion between these two parts can vary between different types of cancers with some tumors displaying a more evident desmoplastic reaction, with a well-represented stromal compartment, while in others this component will grow to a smaller extent. This variation, also seen in the case of tumors belonging to the same cancer type, represents a constant challenge in research especially when using high throughput techniques such as next generation sequencing as it can highly underestimate the true mutational load.

In our analyzed cohort the minimal tumor purity was 36% while the maximum was 95% with an average of 77%. In Fig XX an overall increase in the value of heteroplasmy can

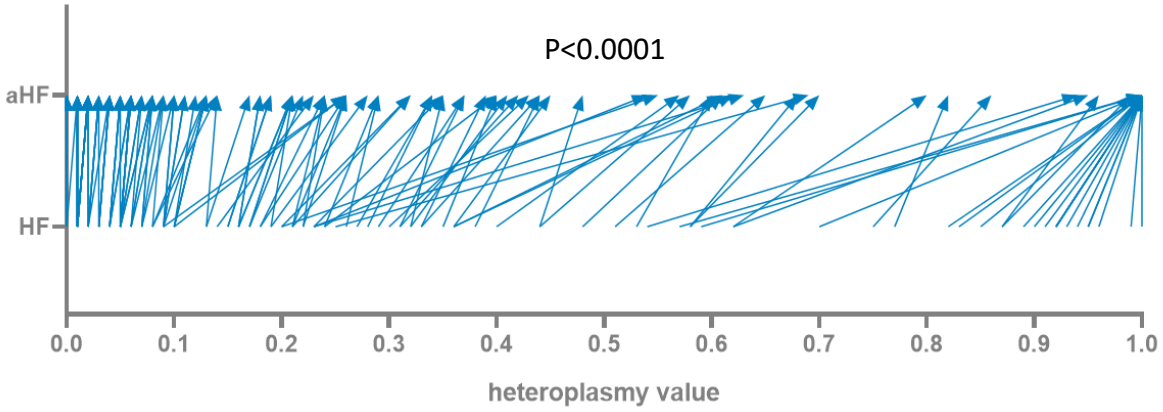
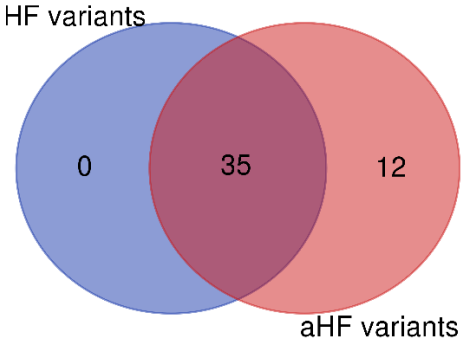


Fig 16. Differences between MToolBox Heteroplasmic fraction (HF) output and adjusted heteroplasmic fraction (aHF) calculated in respect to the tumor purity parameter from TCGAbiolinks.

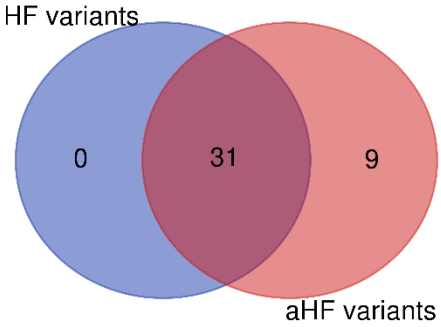
be observed after taking into account the tumor purity correction. In order to see whether this adjustment influences the HF in a significant way, we conducted a two tailed Paired Samples Wilcoxon Test which yielded a highly significant increase between the HF values prior to the tumor purity adjustment and those after ($p<0.0001$; $rs=0.99$).

A. heteroplasmy threshold =0.4



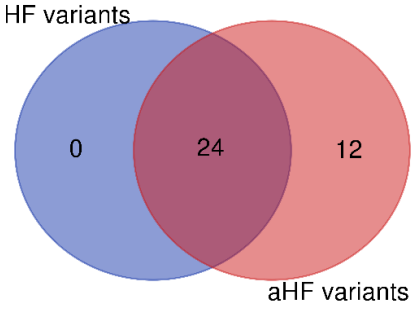
% of variants potentially lost 25.5%

B. heteroplasmy threshold =0.5



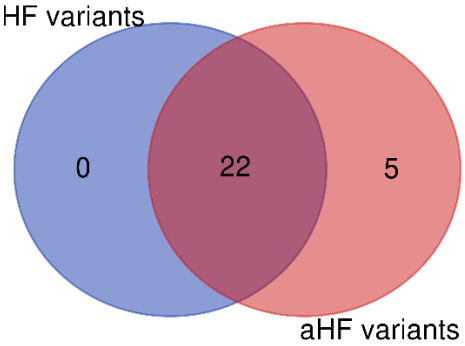
% of variants potentially lost 22.5%

C. heteroplasmy threshold =0.6



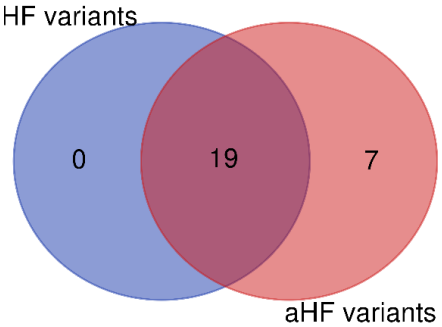
% of variants potentially lost 33.3%

D. heteroplasmy threshold =0.7



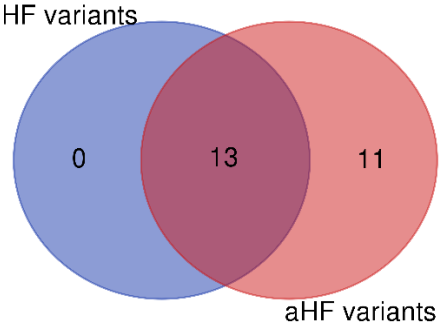
% of variants potentially lost 18.5%

E. heteroplasmy threshold =0.8



% of variants potentially lost 26.9%

F. heteroplasmy threshold =0.9



% of variants potentially lost 45.8%

Fig 17. Venn diagrams representing mitochondrial variants with a major HF underestimation. Different thresholds of 0.4 (A), 0.5 (B), 0.6 (C), 0.7 (D), 0.8 (E), 0.9 (F) were set.

Germline mtDNA variants show heteroplasmic change under positive or negative selective pressure in HCC tumors

Mitochondrial variants are known to be subjected to pressure selection. Mutations favoring tumoral adaptation to the harsh tumoral microenvironment will preferentially expand while those showing a disadvantage would diminish.[84] We asked whether any of the germline variants discovered showed a heteroplasmic shift, be it negative or positive.

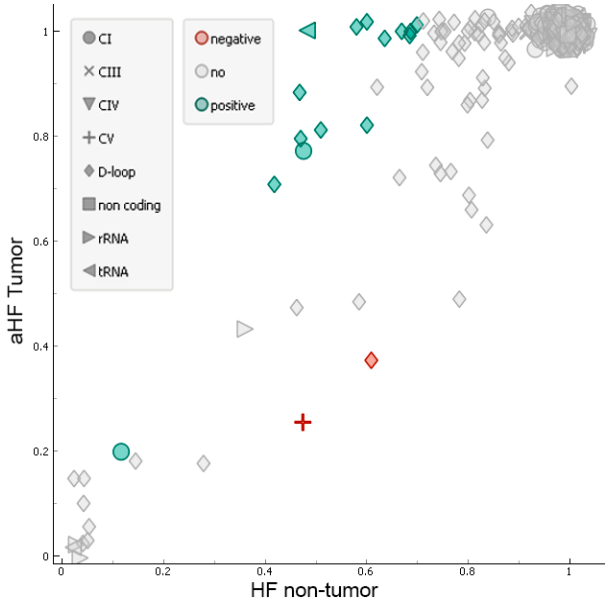


Fig 18. Germline variants which present a positive (green) or negative (red) selection. For the non-tumoral sample the HF from the MToolBox output was used while for the tumoral sample we used the adjusted HF.

For this analysis we plotted the tumoral HF of all germline variants remaining after the exclusion of haplogroup defining events against the original ones found in the non-tumoral tissue (Fig 18). We considered variants showing a shift of 40% increase or decrease in their heteroplasmic fraction. The majority of variants (95.7%) showed a low variation while 13 variants were seen to present a positive selection and two a negative one. Most of these were located in the D-loop region.

Four variants were located in protein coding genes and in a tRNA gene, respectively.

Two synonymous variants located in genes coding for complex I proteins show an increase in the tumoral sample. The A3606G variant was previously reported in a study in type 2 diabetes[86]. It does not appear in the ClinVar database.

Variant Allele	%increase/ decrease	Location	Gene	Nt Var	AF	AA change	DS	Pathogenicity
A3606G	↑ 54	CI	MT-ND1	0.062	0.013	syn L100L		likely polymorphic
A10598G	↑ 58	CI	MT-ND4L	0.004	<0.001	syn M43M		likely polymorphic
T10454C	↑ 117	tRNA	MT-TR	0.014	0.003		0.2	likely polymorphic
G9025A	↓ 46	CV	MT-ATP6	0.003	<0.001	mis G167S	0.89	pathogenic

Table 5. Germline variants showing a tumoral heteroplasmic shift.
AA: aminoacid; syn: synonymous; mis: missense; Nt Var: Nucleotide variability; AF: allele frequency; DS: disease score. *Nt Var and AF are estimated on the basis of the human mitochondrial genomes annotated in GenBank and then available through the HmtDB database and hence in HmtVar.*

The 10454C variant although scored as polymorphic (https://www.hmtvar.uniba.it/varCard/18707), was described as a mitochondrial tRNA mutation associated with non-syndromic deafness. Patients presenting this mutation showed a higher penetrance of the disease and an earlier age of onset of hearing loss[87]. It has been also described in Leigh syndrome[88] and lung cancer[89]. In this the last study the authors also classify it as a "neutral polymorphism".

The 9025A variant located in the mitochondrially encoded ATP synthase 6 is the only protein coding variant to show a decrease in the tumoral sample and is predicted to be pathogenic in HmtVar although no data regarding pathogenicity are available through ClinVar, OMIM and MitoMap.

mtDNA mutations correlate with immune infiltration in HCC tumors

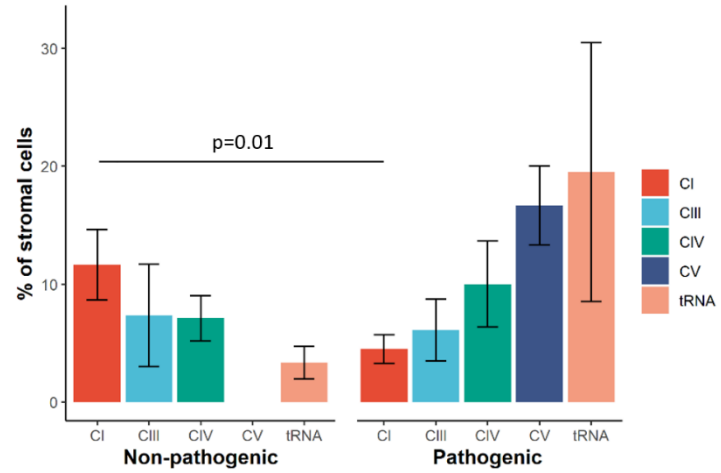
The role of the tumoral microenvironment in cancer progression cannot be overstated. Growing evidence suggests that stromal cells and innate as well as adaptive immune cells can influence the proliferation and drug response of tumoral cells, with major impact on the outcomes of the patients.[90,91]

Escaping immune control of tumoral growth is one of the hallmarks of cancer.[17] The intratumoral level of immune infiltration has been described recently as a factor with prognostic value in different cancers.[92,93] Exhaustion of T cells is the process in which T cells become dysfunctional and therefore allow tumoral immune escape.[94] Elevated ROS levels in the tumoral microenvironment have been associated with lymphocyte inactivation and tumor induced immunosuppression[95]. Because mtDNA mutations could lead to the generation of higher levels of ROS[96,97], we sought to evaluate whether the presence of pathogenic mtDNA mutations could corelate with the immune component of the tumoral samples. For this we made use of the paraclinical immunohistochemical data available on the GDC portal which contains for each tumor specimen, the percentage of immune cell infiltration (monocyte, lymphocyte and neutrophils) as well as the percentage of stromal component.[80]

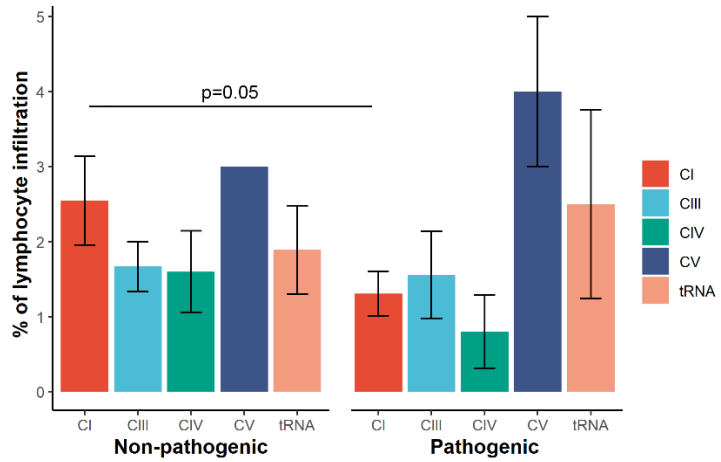
In the analysed TCGA HCC cohort, we first focused our attention on mitochondrial variants located in genes coding for subunits of CI and CIII, the main generators of ROS from the ETC. We observed that the HCC tumors harboring pathogenic variants had a significantly lower proportion of the stromal component (p-value for Student's t-test = 0.01) as well as a lower level of lymphocytic infiltration (p=0.05). When analyzing each complex separately, only mutations located in CI genes reached statistical significance for both the stromal component as well as for lymphocyte infiltration (p-value=0.05 for both).

Regarding the impact of variants situated in genes coding for other components of complexes or for tRNA genes, the p-value remained not significant although a positive trend could be observed in favor of the pathogenic variants (Fig 19).

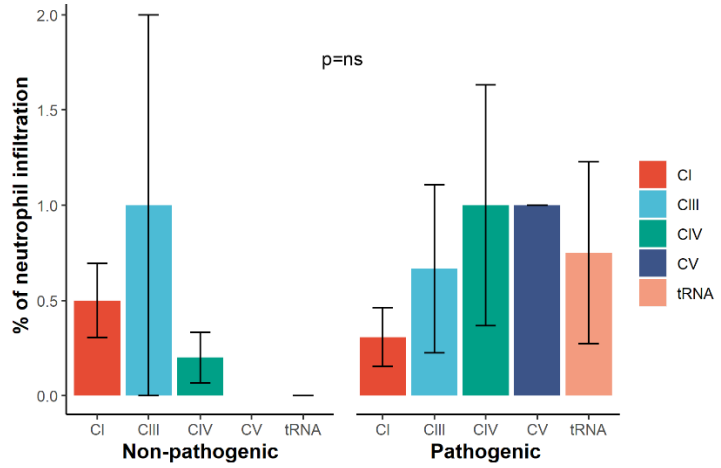
A.


% of stromal cells
30
20
10
0
p=0.01
CI
CIII
CIV
CV
tRNA
Non-pathogenic
Pathogenic

B.


% of lymphocyte infiltration
5
4
3
2
1
0
p=0.05
CI
CIII
CIV
CV
tRNA
Non-pathogenic
Pathogenic

C.


% of neutrophil infiltration
2.0
1.5
1.0
0.5
0.0
p=ns
CI
CIII
CIV
CV
tRNA
Non-pathogenic
Pathogenic

D.

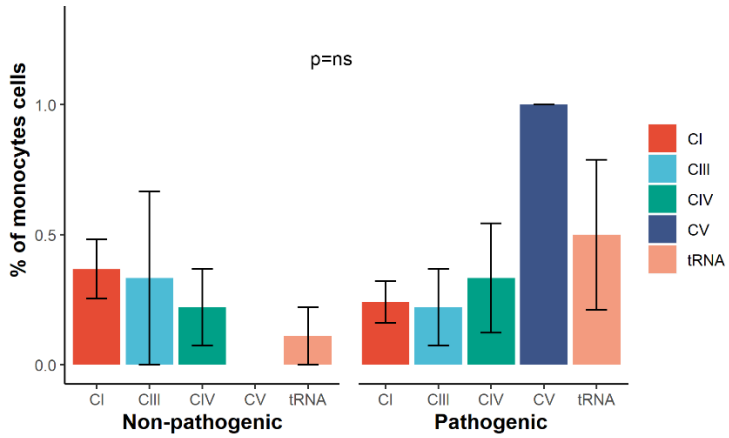


Fig 19. Somatic complex I mutations correlate with the level of intratumoral stromal component (A), lymphocyte infiltration (B), neutrophils (C) and monocytes (D).
Mitochondrial complexes I-V are abbreviated as CI-V

Of note, for mutations located in CV genes we could not conduct any analysis as the non-pathogenic group comprised only one tumor sample.

Pathogenic mtDNA mutations predict overall survival of TCGA LIHC patients

Because of the known pro-tumorigenic effects of some mtDNA mutations, we sought to analyze the possible relationship of pathogenic mtDNA mutations on different clinical features of HCC samples.

First we analyzed whether the somatic mutations presented a different distribution between the different stages of the disease (Fig 20).

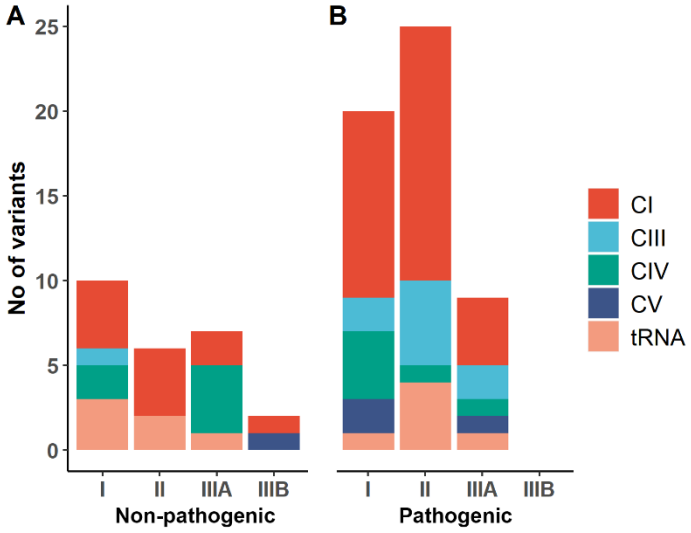


Fig 20. Distribution of non-pathogenic (**A**) and pathogenic (**B**) somatic mtDNA mutations between patients having different stages of HCC.
Mitochondrial complexes I-V are abbreviated as CI-V

Chi-square test showed no significant difference in the distribution of pathogenic mtDNA mutations between the stages of HCC. The analysis was conducted only for stages I, II and IIIA, as the only two patients having stage IIIB of the disease showed only non-pathogenic mtDNA variants. Because tumoral resection, hence access to tumoral tissue for the analysis, is recommended for early stages of the disease, samples from advanced stages are often limited, as seen also in this cohort. Thus, a comprehensive analysis of the mutational burden between diferent stages cannot be fully carried out.

We next sought to see if the presence of pathogenic mutations was influenced by any of the backgound risk factors present in the HCC patients. Again, the Chi-Square test remained insignificant, with the mutations showing a uniform distribution between patients (Fig 21).

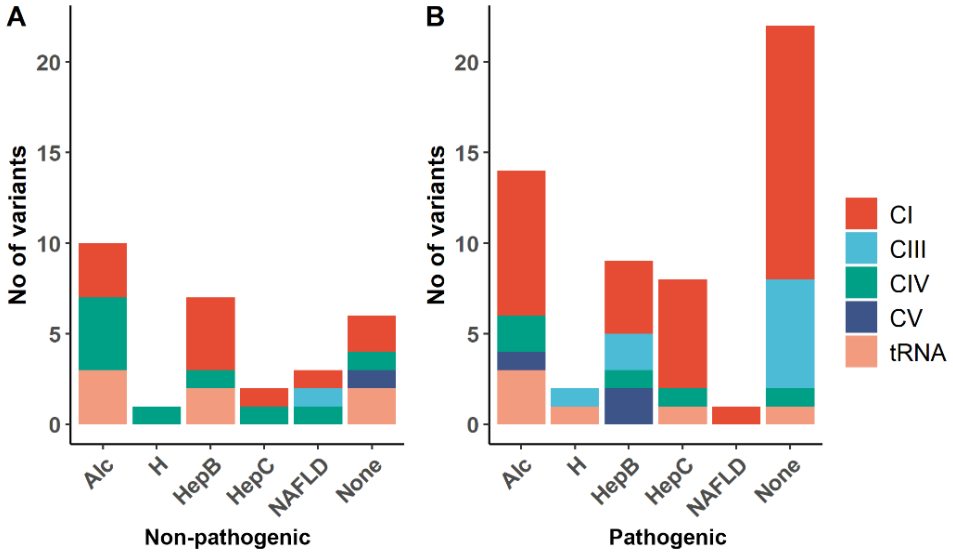


Fig 21. Distribution of non-pathogenic (**A**) and pathogenic (**B**) mtDNA mutations between patients with HCC developed on different backgound conditions. Alc: Alcohol consumption; H: Hemochromatosis; HepB: Hepatitis B; HepC: Hepatitis C; NAFLD: Non-Alcoholic Fatty Liver Disease; None: Absence of a background condition. Mitochondrial complexes I-V are abbreviated as CI-V

We then proceeded to see whether the presence of a pathogenic mtDNA mutation in tumors would change the survival outcome of the TCGA-HCC patients. The required information was available for 102 out of 104 patients. Survival analysis was performed between the group of patients carrying a pathogenic mtDNA mutations and those carrying non-pathogenic variants.

Given the known threshold effect presented by mtDNAmutations, in order to avoid selecting a random threshold to analyze our data, we used different adjusted heretoplasmy thresholds in the range of the values reported in literature. Thus, the following Kaplan-Meyer survival analyses were conducted using thresholds values between 0.4 and 0.9 aHF (presented in Fig 22). For all survival curves, patients harboring non-pathogenic mitochondrial mutations presented statistically significant higher survival rates when compared to their counterparts with pathogenic mutations (p values < 0.05 for all curves).

Because of the small number of patients available for the analysis, the median survival times could not be calculated for the non-pathogenic patients groups, therefore in order to evaluate the average survival of the two groups we calculated the estimated mean

survival times for each group. Since time is a type of variable with a non-normal distribution, this limitation of the analysis could represent a source of bias. The results are resumed in Table 6. It can be noted that no matter the heteroplasmy threshold set for analysis, the patients harboring non-pathogenic variants presented a significantly longer survival time.

	Mean Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
aHF≥0.4				
non-pathogenic	2084.241	220.169	1652.710	2515.772
pathogenic	1362.300	257.771	857.069	1867.532
aHF≥0.5				
non-pathogenic	2047.198	241.950	1572.976	2521.419
pathogenic	1358.512	286.210	797.540	1919.484
aHF≥0.6				
non-pathogenic	2199.714	208.574	1790.909	2608.519
pathogenic	1110.908	251.429	618.107	1603.709
aHF≥0.7				
non-pathogenic	2162.167	239.933	1691.898	2632.435
pathogenic	1231.833	320.467	603.717	1859.950
aHF≥0.8				
non-pathogenic	2162.167	239.933	1691.898	2632.435
pathogenic	1195.875	323.836	561.156	1830.594
aHF≥0.9				
non-pathogenic	2109.600	282.102	1556.679	2662.521
pathogenic	1519.500	341.250	850.650	2188.350

Table 6. Estimated mean survival of the TCGA-HCC cohort based on different aHF thresholds

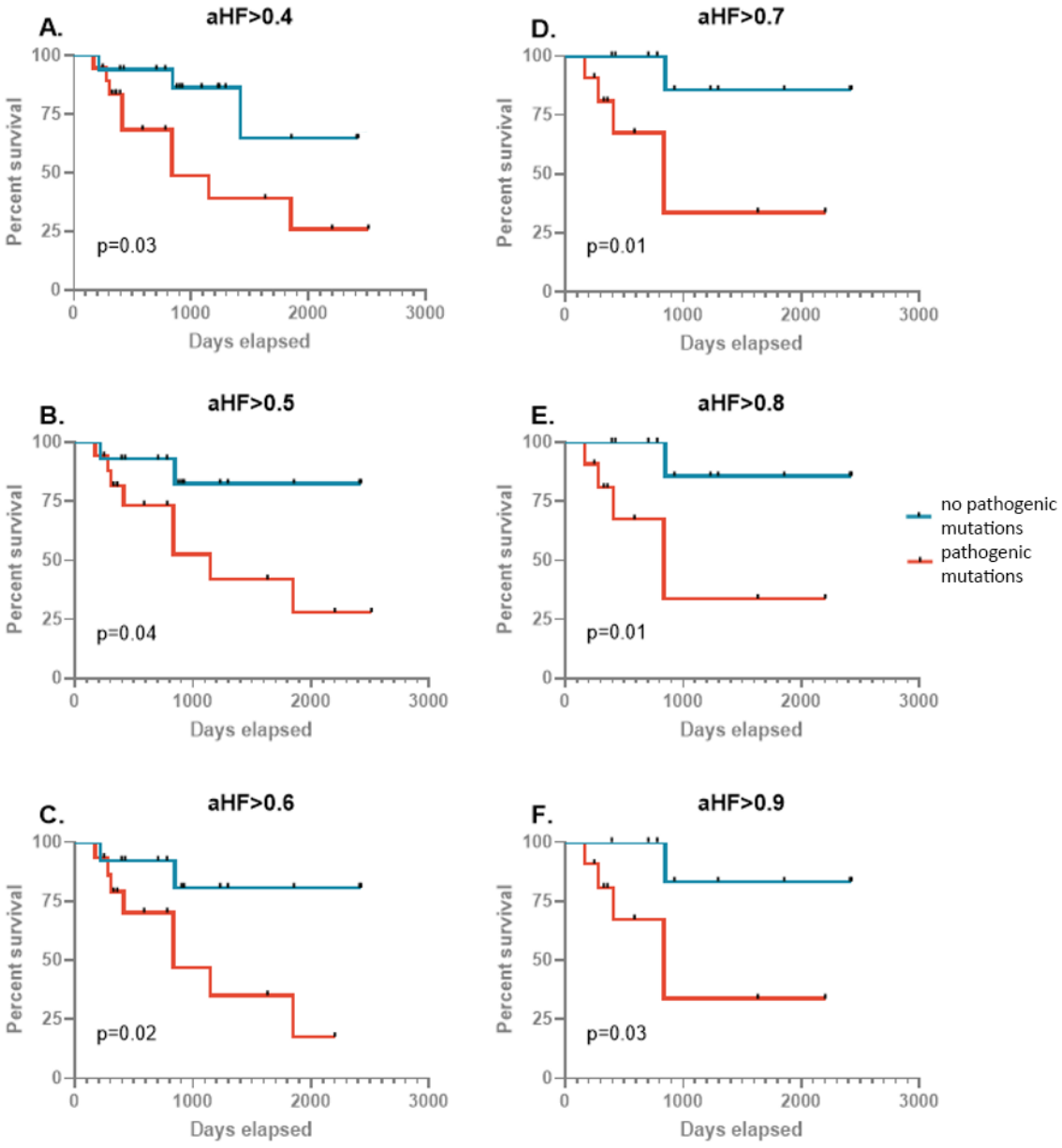


Fig 22. Survival analysis of TCGA-HCC patients with different aHF thresholds of 0.4 (A), 0.5 (B), 0.6 (C), 0.7 (D), 0.8 (E), 0.9 (F). Patients carrying somatic pathogenic mtDNA mutations had a poorer survival.

In order to evaluate whether survival was influenced by other factors and their effect size, we performed a multivariate analysis using Cox regression model. To identify covariates which could potentially affect patient prognosis we considered for the analysis the following variables: age at diagnosis, gender, AFP levels, stage and microvascular invasion. In order to select the significant variables to introduce in our model we first performed separate univariate analysis. Gender attained statistical significance for aHF thresholds of ≥0.5 ($p<0.01$) , ≥0.6 ($p<0.02$), ≥0.7 ($p<0.03$) and ≥0.8 ($p<0.03$) while all other variables remained unsignificant for every heteroplasmy fraction threshold.

Results of the multivariate analysis can be found in Table 7.

				Hazard Ratio	
		Hazard	**p**	**(HR)**	**HR 95.0% CI**
aHF≥0.4 overall p=0.023	mtDNA mutations	1.008	0.066	2.739	0.936-8.016
	Gender	-1.108	0.040	0.330	0.115-0.951
aHF≥0.5 overall p=0.002	mtDNA mutations	1.136	0.043	3.113	1.035-9.366
	Gender	-1.424	0.011	0.241	0.081-0.719
aHF≥0.6 overall p=0.003	mtDNA mutations	1.348	0.023	3.850	1.204-12.308
	Gender	-1.329	0.024	0.265	0.083-0.841
aHF≥0.7 overall p=0.005	mtDNA mutations	1.922	0.021	6.831	1.341-34.800
	Gender	-2.159	0.008	0.115	0.024-0.566
aHF≥0.8 overall p=0.006	mtDNA mutations	1.910	0.019	6.754	1.361-33.510
	Gender	-2.097	0.010	0.123	0.025-0.603
aHF≥0.9 overall p=0.285	mtDNA mutations	1.068	0.288	2.911	0.406-20.883
	Gender	-1.339	0.145	0.262	0.043-1.588

Table 7. Multivariate Cox regression analysis of overall survival of HCC patients included in the study

We can see that carrying pathogenic mtDNA mutations significantly increases the death hazard ratio for all the aHF thresholds between 0.5 and 0.8. Harbouring somatic mtDNA mutations was associated with a risk of death up to 6.8 times higher for a aHF≥0.7. On the other hard, male sex was associated with a lower death hazard ratio in this group of patients.

It is known that prognosis of HCC patients is influenced by the stage of the disease. We can observe this trend also in the HCC cases studies here albeit not statistically significant (Fig 23).

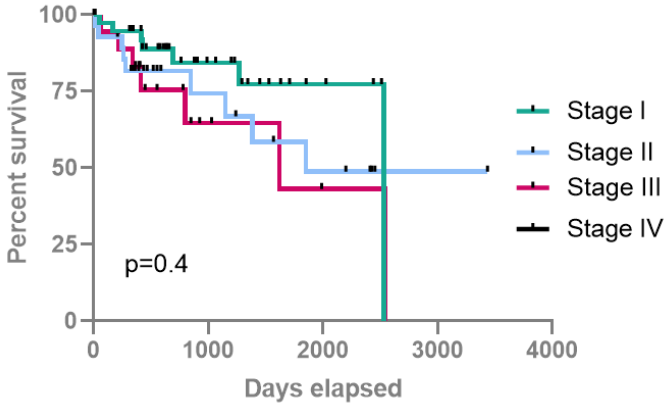
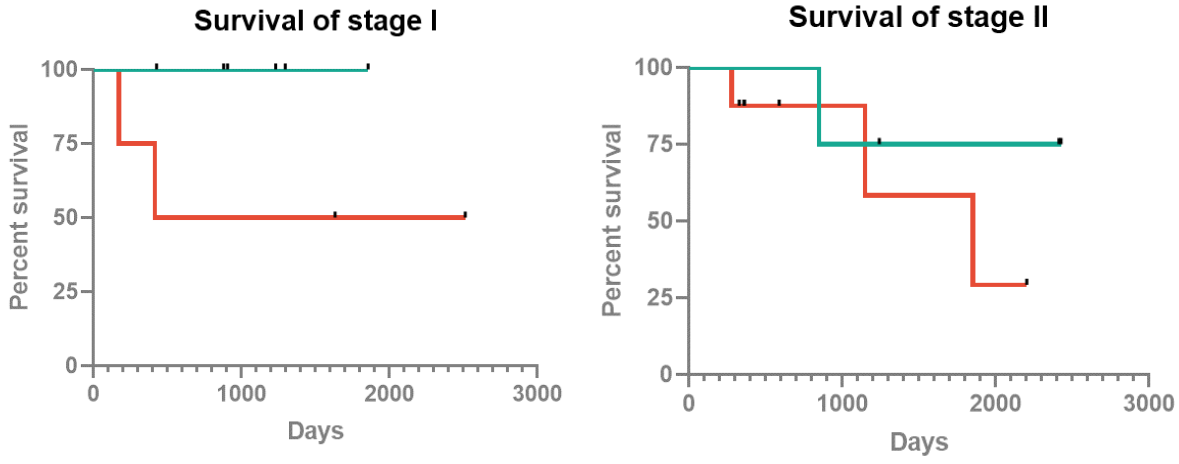


Fig 23. Survival of the TCGA HCC patients based on the stage of the disease

We therefore sought to analyse whether the presence of pathogenic mtDNA mutations was also capable of influencing the prognosis of the different stages of the disease.

The analysis was possible only for mtDNA variants with a aHF>0.4, as filtering further would have left insufficient patients to construct a curve (Fig 24)



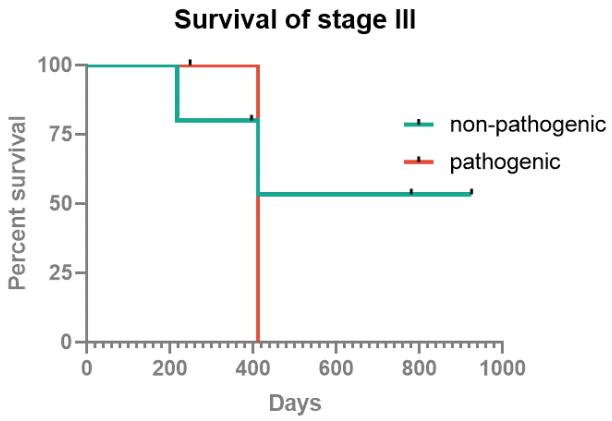


Fig 24. Survival of patients with different HCC stages based on the burden of pathogenic mtDNA mutations

The significance was not calculated due to the small number of patients in each subgroup, a trend can be observed further confirming the prognostic value of pathogenic mtDNA mutations in HCC patients.

TP53 status predict overall survival independent of mitochondrial DNA mutations burden

The TP53 gene is the most frequently mutated gene in tumors. Out of all 104 HCC patients, 24 (23%) presented mutations in the TP53 gene (Fig 25). The missense c.747G>T which changes the Arginine in position 249 to a Serine was seen in 4 patients. This mutation is frequently seen as a consequence of exposure to aflatoxin B1. Exposure information is unfortunately lacking from the TCGA clinical files.

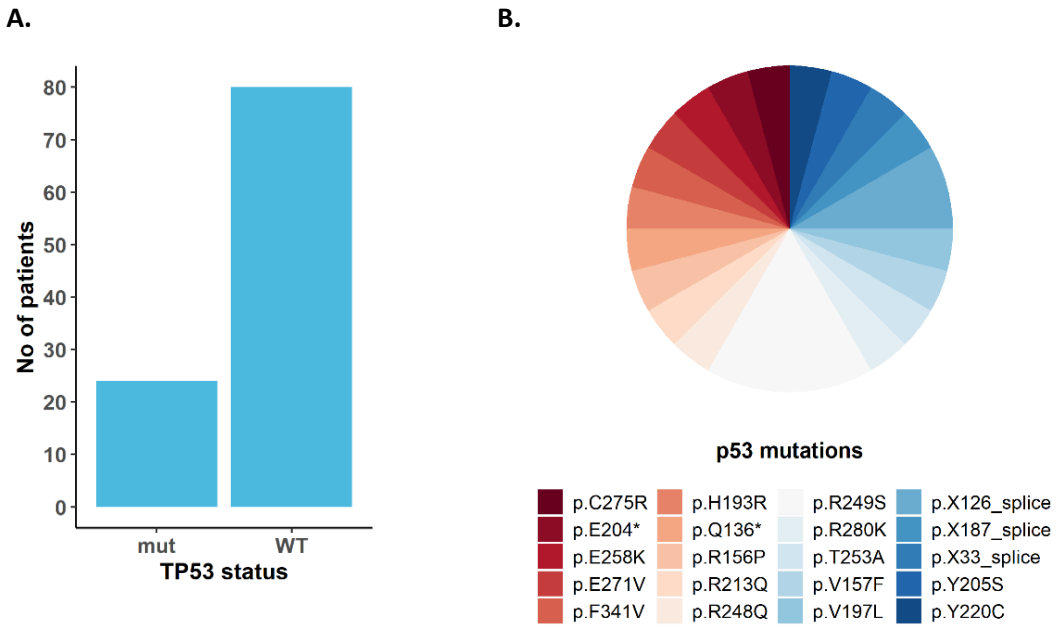


Fig 25. A. Number of patients carrying TP53 mutations. B. TP53 mutations present in our analyzed patients

To gather information about these mutations, including pathogenicity predictions, we queried the IARC TP53 database (http://p53.iarc.fr/) which contains information regarding TP53 mutations associated with human cancers that have been either curated from genomic databases or published in peer reviewed papers (Table 8)

The majority were in the DNA binding domain of the protein and had an impact on the transactivational capacity of reported genes under the control of a p53-response element. These mutations have also shown dominant-negative effects as well as loss of growth-suppression activities.

HGVSc	HGVSp	Effect	Hotspot	Domain function	SIFT Class	Polyphen2	Transactivation Class	Loss of growth Suppression
c.823T>C	p.C275R	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.610G>T	p.E204*	nonsense	no	DNA binding				notDNE_LOF
c.772G>A	p.E258K	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.812A>T	p.E271V	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.1021T>G	p.F341V	missense	no	Tetramerisation	D	B	partially functional	notDNE_LOF
c.578A>G	p.H193R	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.406C>T	p.Q136*	nonsense	yes	DNA binding				notDNE_notLOF
c.467G>C	p.R156P	missense	no	DNA binding	T	D	non-functional	DNE_LOF
c.638G>A	p.R213Q	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.743G>A	p.R248Q	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.747G>T	p.R249S	missense	yes	DNA binding	D	D	non-functional	DNE_LOF
c.839G>A	p.R280K	missense	yes	DNA binding	D	D	non-functional	notDNE_LOF
c.757A>G	p.T253A	missense	no	DNA binding	D	D	non-functional	DNE_LOF
c.469G>T	p.V157F	missense	yes	DNA binding	D	D	non-functional	unclass.
c.589G>T	p.V197L	missense	yes	DNA binding	D	D	partially functional	unclass.
c.376-1G>A	p.X126_splice	splice	no					
c.560-1G>A	p.X187_splice	splice	no					
c.97-1G>A	p.X33_splice	splice	no					

HGVSc	HGVSp	Effect	Hotspot	Domain function	SIFT Class	Polyphen2	Transactivation Class	Loss of growth Suppression
c.614A>C	p.Y205S	missense	yes	DNA binding	D	P	non-functional	DNE_LOF
c.659A>G	p.Y220C	missense	yes	DNA binding	D	D	non-functional	DNE_LOF

Table 8. Characteristics of the TP53 mutations present in the HCC cohort.
HGVS: Human Genome Variation Society; SIFT Class D: Deleterious, T: Tolerated; Polyphen2 D: probably damaging, P: Possibly damaging, B: Benign; DNE: dominant negative effect; LOF: loss of function

Apart from the classical roles played “as the guardian of the genome”, p53 has been shown to be involved also in the regulation of metabolism and mtDNA proofreading. Thus, we investigated whether the mutational status of TP53 correlated with the presence and number of mtDNA mutations.

For this analysis, we included all somatic variants discovered with an aHF higher than 0.1 to exclude false positive variants possibly due to sequencing errors. We observed a slight increase in the overall number of mtDNA variants as well as of the mtDNA pathogenic variants (Fig 26) but without attaining statistical significance (p-value of Fisher’s exact test = ns).

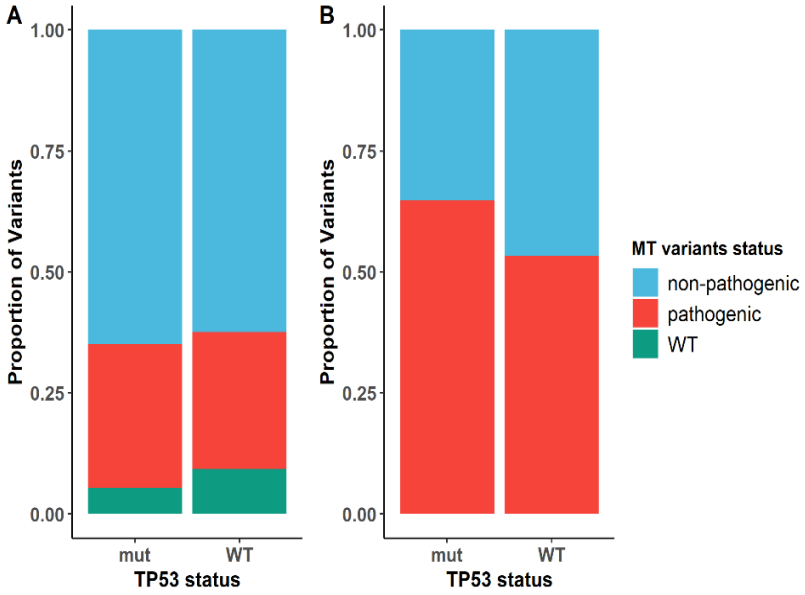


Fig 26. Proportion of mitochondrial variants in HCC patients with different TP53 status. **A.** All discovered somatic variants with an aHF>0.1 (WT[green] are patients without any mtDNA somatic events). **B.** Only variants found in coding regions and tRNA genes with an aHF>0.1.

Given the known influence of *TP53* mutations survival of cancer patients, we sought to analyze whether in our cohort of HCC patients *TP53* would have an impact on the overall survival regardless of the burden of mtDNA mutations. The Kaplan-Meyer curve shows that patients with a mutated *TP53* had a significant poorer prognosis p=0.007 (Fig 27).

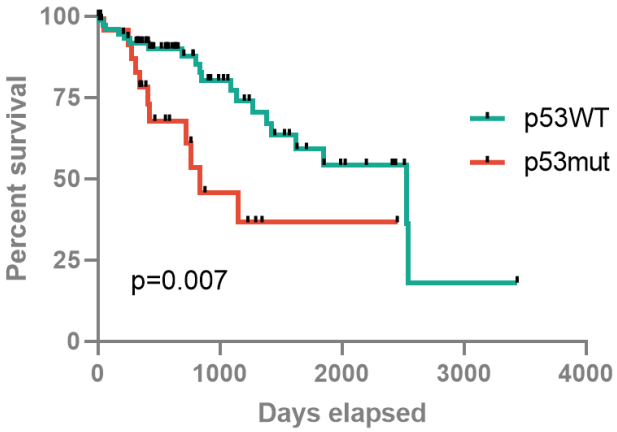


Fig 27. Survival of HCC patients with wild-type or mutated TP53

Sources of mtDNA damage in the LIHC-TCGA cohort

Mutations in the mitochondrial genome are known to occur at a much higher rate than in the nuclear genome. One of the currently accepted explanations for this observation is that the close proximity of the mtDNA to the OXPHOS chain renders it more vulnerable to oxidative damage induced by ROS which in tumors are known to reach higher levels. On the other hand, these mutations can be attributed also to polymerase gamma (POLG) replication errors, as the mitochondrial do not benefit from the complex DNA repair mechanisms of the nuclear DNA.

Mitochondria through their activity of oxidative phosphorylation represent the main cellular generators of ROS, which can ultimately induce mtDNA injury. The currently accepted oxidative damage signature is characterized by the high rate of G:C to T:A transversions via the generation of 8-hydroxy-2'-deoxyguanosine (8-oxo-dG). Moreover, ROS may induce damage to the dGTP pool with the generation of 8-OH-dGTP. This leads to a transversion from A:T to C:G in the mtDNA genome.

The second source for the occurrence of mtDNA mutation is the erroneous DNA replication due to polymerase errors, especially in the homopolymeric regions. Indeed, mutations in the *POLG* gene have been associated with increased levels of mtDNA mutations.[98,97] Moreover, mutations in this gene have been shown to occur in 63% of breast tumors leading to mtDNA depletion as well as increased tumorigenicity.[99] This mutational pattern is characterized by C>T and A>G substitutions.[98]

We sought to analyze whether the somatic mutations found in our patients follow such DNA damage models and found that 53% of all events were represented by the G>A transition. A major part of these changes was seen in genes encoding for CI subunits. The second most frequent base change seen was T>C representing 20%, followed by C>T in 15%.

The A>C and C>A changes represented only 5.5% of all the mtDNA alterations.

The distribution of the type of alterations found is presented in Fig 28.

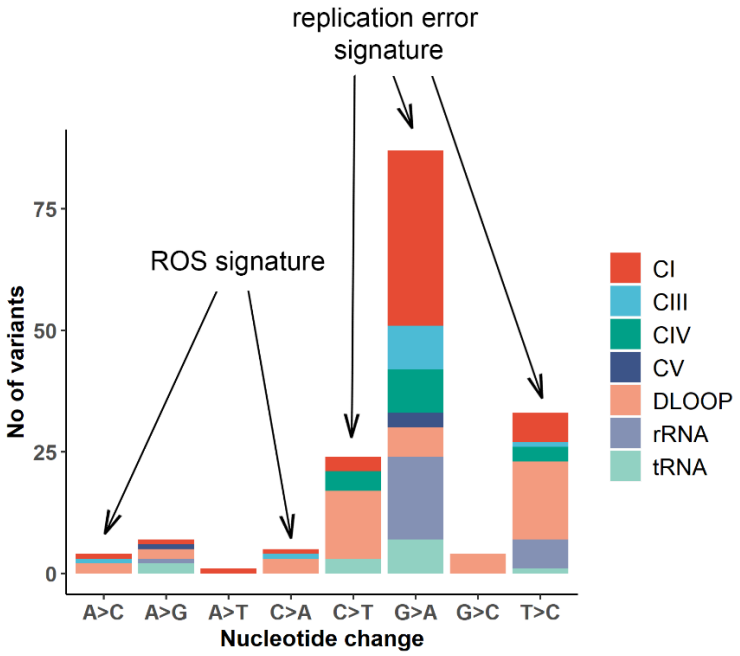


Fig 28. Number of somatic transitions and transversions present in our HCC cohort.

To further investigate the possible causes which might explain the accumulation of such mutations in these patients, we controlled for the presence of mutations in genes responsible for ROS detoxification and for mtDNA replication machinery. We proceeded to extract the somatic mutations of our genes of interest from the *.maf* files.

We searched for enzymes involved in ROS-detoxification and homeostasis on Gene Ontology (GO) knowledgebase (http://www.geneontology.org/) and on Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/)[100] we selected the following list of 36 genes:

Gene	Ensembl ID	Gene	Ensembl ID
GLRX1	ENSG00000173221	TXNDC2	ENSG00000168454
GLRX2	ENSG00000023572	TXNDC3/NME8	ENSG00000086288
GLRX3	ENSG00000108010	TXNDC5	ENSG00000239264
GLRX5	ENSG00000182512	TXNDC6/NME9	ENSG00000181322
GRXCR1	ENSG00000215203	TXNDC8	ENSG00000204193
NOS1	ENSG00000089250	TXNDC9	ENSG00000115514

NOS2	ENSG00000007171	TXNDC11	ENSG00000153066
NOS3	ENSG00000164867	TXNDC12	ENSG00000117862
PRDX1	ENSG00000117450	TXNDC15	ENSG00000113621
PRDX2	ENSG00000167815	TXNDC16	ENSG00000087301
PRDX3	ENSG00000165672	TXNDC17	ENSG00000129235
PRDX5	ENSG00000126432	TXNL1	ENSG00000091164
PRDX6	ENSG00000117592	TXNL4A	ENSG00000141759
SOD-1	ENSG00000142168	TXNL4B	ENSG00000140830
SOD-2	ENSG00000112096	TXNRD1	ENSG00000198431
SOD-3	ENSG00000109610	TXNRD2	ENSG00000184470
TXN	ENSG00000136810	TXNRD3	ENSG00000197763
TXN2	ENSG00000100348	TXNIP	ENSG00000265972

We found 13 single nucleotide variations distributed in 11 patients. Eight variants were missense mutations with 6 having a pathogenic SIFT/Polyphen prediction (Table 9 A). Given the small number of these patients no differences in the distribution of mtDNA variants was seen (Fig 29). Moreover, upon closer analysis we saw that only two of these patients had an alteration typical of ROS induced DNA damage. This leads us to conclude that, at least in this cohort of HCC patients, mutations in genes coding for enzymes responsible for ROS detoxification are not responsible for the mtDNA mutations found.

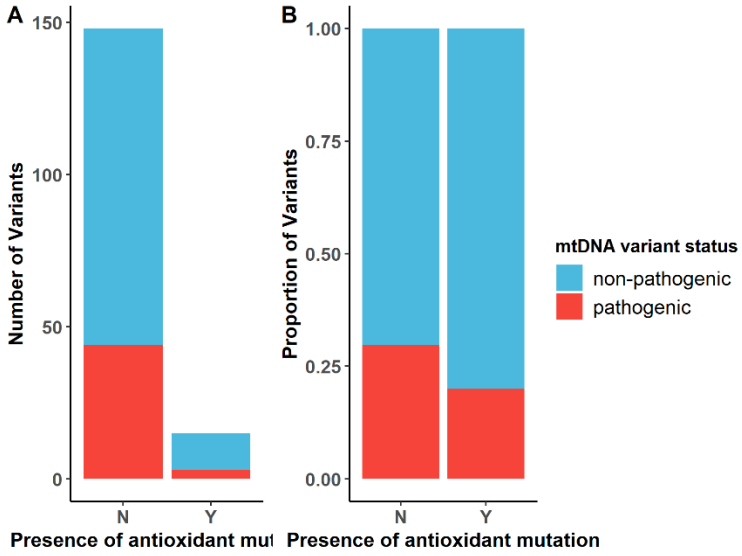


Fig 29. Distribution of mtDNA mutations between patients with and without mutations in antioxidant genes.
N: no; Y: yes

Next, we searched for mutations in the *POLG* gene as well as in other genes encoding for the mitochondrial DNA replication machinery.

Gene	Ensembl ID
POLG2	ENSG00000256525
Twinkle	ENSG00000107815
RNASEH1	ENSG00000171865
DNA2	ENSG00000138346
MGME1	ENSG00000125871
TFAM	ENSG00000108064
FEN1	ENSG00000168496
LIG3	ENSG00000005156
mtSSB	ENSG00000106028

We found that out of all the 377 LIHC patients on TCGA, only 14 patients presented variations in the selected genes and out of these, only three patients were present in our cohort. These patients did present also mtDNA variants characterized by a C>T change. While mutations in genes encoding components of the mtDNA replication machinery might explain the occurrence of mtDNA mutations in these patients, it does not offer an explanation for the variants seen in the other patients. Thus, we cannot attribute the number of mtDNA mutations in the TCGA-LIHC cohort to mutations.

Discovered mutations are summarized in Table 9 B.

A.

Gene	dbSNP	HGVSc	HGVSp	Consequence	SIFT	PolyPhen	IMPACT	COSMIC
GRXCR1	novel	c.619T>A	p.Y207N	missense	Deleterious (0.02)	D (0.999)	MODERATE	
NME8	novel	c.444T>C	p.C148C	synonymous			LOW	
NME8		c.157A>T	p.K53*	stop_gained			HIGH	COSM4923173
NOS1		c.1254G>A	p.S418S	synonymous			LOW	COSM1359118
NOS1		c.2444G>T	p.G815V	missense	Deleterious (0)	D (0.999)	MODERATE	COSM4925301
NOS2	rs759255233	c.1824G>A	p.S608S	synonymous			LOW	
NOS2	novel	c.2013C>A	p.S671R	missense	Tolerated (0.18)	P (0.58)	MODERATE	
NOS2		c.2824G>A	p.G942S	missense	Deleterious (0)	D (0.974)	MODERATE	COSM4937890
NOS3	novel	c.720C>A	p.D240E	missense	Deleterious (0.03)	D (0.98)	MODERATE	
NOS3	rs145811781	c.1296G>T	p.E432D	missense	Tolerated (0.08)	D (0.99)	MODERATE	
PRDX6	novel	c.521G>A	p.R174K	missense	Tolerated (1)	B (0)	MODERATE	
TXN2	novel	c.388G>C	p.V130L	missense	Deleterious (0)	P (0.859)	MODERATE	
TXNDC2	novel	c.279A>C	p.S93S	synonymous			LOW	

B.

FEN1	rs573495657	c.1016G>A	p.R339H	Missense	Deleterious s(0)	D (0.993)	MODERATE	COSM1355266
LIG3	novel	c.2011C>T	p.R671W	Missense	Deleterious (0)	D (1)	MODERATE	
LIG3	novel	c.2743G>T	p.V915L	Missense	Tolerated (0.54)	B (0.002)	MODERATE	

Gene	dbSNP	HGVSc	HGVSp	Consequence	SIFT	PolyPhen	IMPACT	COSMIC
LIG3	rs761685440	c.1427C>T	p.S476L	Missense	Deleterious (0.01)	B (0.068)	MODERATE	COSM3819302 COSM3819303
POLG2	novel	c.438G>T	p.R146S	Missense	Deleterious (0.01)	D (0.92)	MODERATE	
POLG2	novel	c.438G>T	p.R146S	Missense	Deleterious (0.01)	D (0.92)	MODERATE	

Table 9. Summary of variants located in ROS detoxifying genes (A) and in genes encoding for components of the mitochondrial replisome found in HCC patients.

Polyphen D: probably damaging, P: Possibly damaging, B: Benign

HCC patients present also mutations in OXPHOS genes encoded by the nuclear genome

Studies tackling the subject of mutations involving the respiratory chain often focus only on those found in the mitochondrial genome and leave behind the many OXPHOS genes coded by the nuclear genome. This might be justified by the fact that given the particular vulnerability of the mtDNA to the acquisition of mutations, a tumor is more likely develop OXPHOS impairment due to mutations in the mtDNA rather than in the nuclear one. Nevertheless, such an approach might miss out some valuable information. Therefore, to have a complete view over the mitochondrial mutational status, we sought to see whether any of the patients presented mutations in mitochondrial proteins encoded by the nuclear DNA. For this purpose, we selected genes included in the KEGG OXPHOS pathway dataset. All genes were listed also in the MitoCarta 2.0, an online resource comprising genes with strong evidence of mitochondrial localization.[101]

We found a total of 34 single nucleotide variations in the tumoral samples of 30 patients. Out of the 21 located in coding regions, five were mutations predicted to be pathogenic by both SIFT/Polyphen. Four patients presented mutations in both mitochondrial as well as nuclear encoded OXPHOS genes Table 10.

Location	Gene	Consequence	dbSNP	HGVSc	HGVSp	SIFT	PolyPhen	IMPACT
CI	**NDUFS1**	missense	novel	c.1697T>G	p.I566S	**deleterious(0)**	**D (0.925)**	MODERATE
CI	**NDUFB5**	missense	novel	c.413C>T	p.A138V	**deleterious(0.01)**	**P (0.493)**	MODERATE
CV	**ATP6AP1**	missense		c.229T>A	p.Y77N	**deleterious(0)**	**D (0.993)**	MODERATE
CV	**ATP6V1B2**	missense		c.563A>C	p.K188T	**deleterious(0)**	**D (0.996)**	MODERATE
CV	**ATP6V1B1**	missense	novel	c.883A>C	p.S295R	**deleterious(0)**	**D (0.978)**	MODERATE

Table 10. Variants found in nuclear encoded mitochondrial genes. Polyphen D: probably damaging, P: Possibly damaging, B: Benign

Livers with HCC present a heterogenous staining for mitochondrial complexes

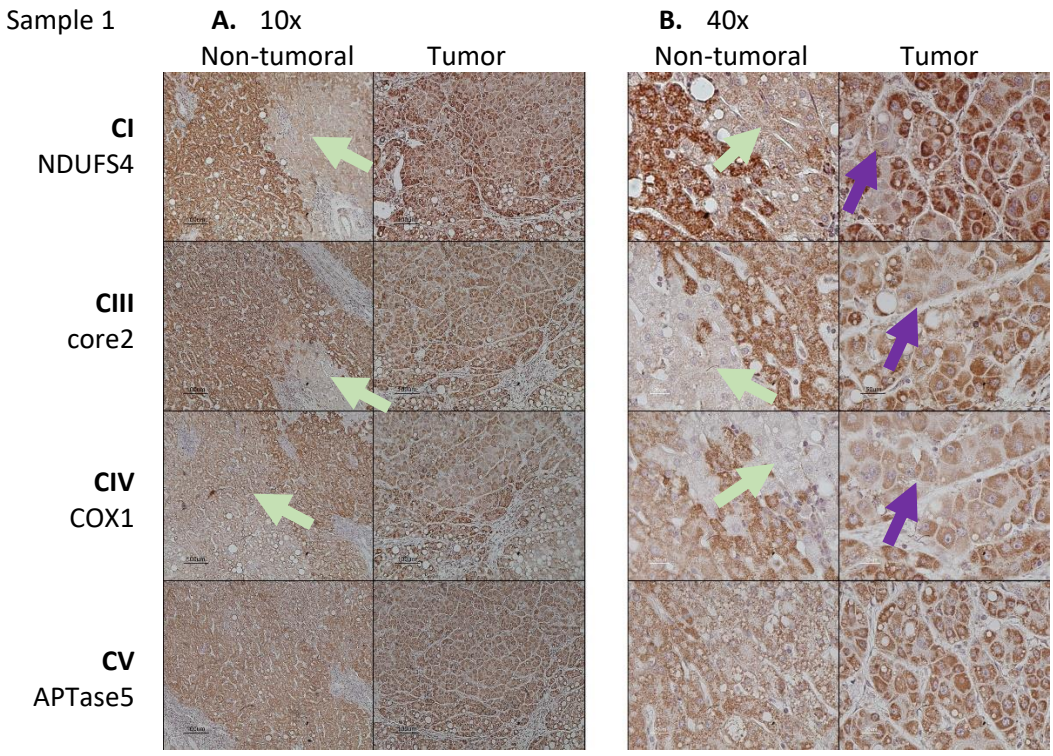
Mutations affecting genes coding for OXPHOS chain complexes can cause the disassembly of the respective complex resulting in the disruption of the respiratory chain.[102,103] Given the considerable number of mtDNA mutations found previously, we sought to see whether by staining HCC tumors and surrounding non-tumoral tissues for mitochondrial complexes, we could find a possible hint of such defects. We noticed that both non-tumoral as well as tumoral tissues presented a heterogenous pattern of the staining. However, there was a difference between the two: in the non-tumoral tissue there was a well delimited border between the intensely stained zones and the weakly/negatively stained areas, whereas the tumors were exibiting a more disorganized pattern, with a mixture of cells showing an intense staining with those with a lower staining intensity.

Another observation that can be made is that the non-tumoral tissue seems to display a gradient in staining intensity, with a lower intensity towards the central part of the tumoral mass. This phenomenon was seen more intensely in the case of sample no 4 and 5, where at the line of interface with the stromal component, the tumor presented a much highly intense signal than in the rest of the mass. This could be explained by the fact that as tumors grow, the center becomes hypoxic forcing cells to rely more on glycolysis as the oxygen demand by the OXPHOS chain is not fulfilled.

Next we sought to confirm the presence of mtDNA mutations in the nodules and areas showing a different staining pattern but unfortunately, the small size of the nodules and limited available tissue combined with the long storage of the samples did not allow for a DNA extraction of a sufficient quality and quantity for the amplification of the mitochondrial genome. Thus, we can only suppose that these patterns could be due to the presence of mtDNA mutations.

Fig 30 presents the pattern of immunohistochemical staining of mitochondrial complexes at 10x (A) and 40x (B) magnification. The non-tumoral surrounding tissue is on the left of each panel while on the right there is the HCC tumoral tissue

Fig 30. Staining pattern of mitochondrial proteins in HCC tumors



Light green arrows show well delimited areas of weak/negative staining in the non-tumoral tissue.
Purple arrows indicate cells with a weak staining dispersed in the tumoral mass

Sample 2

A. 10x

Non-tumoral Tumor

B. 40x

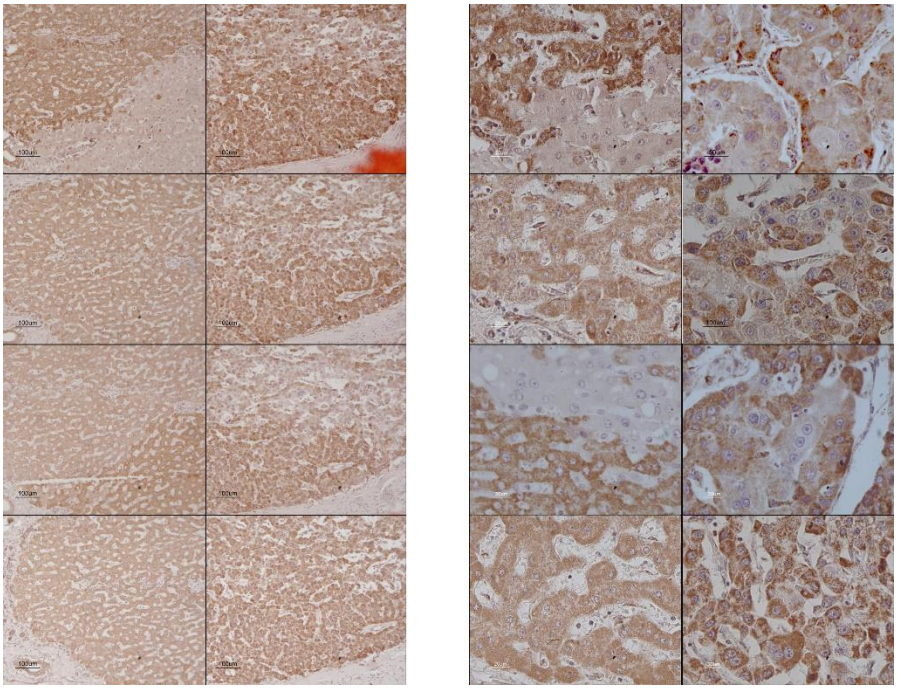
Non-tumoral Tumor

CI
NDUFS4

CIII
core2

CIV
COX1

CV
APTase5



Sample 3

A. 10x

Non-tumoral Tumor

B. 40x

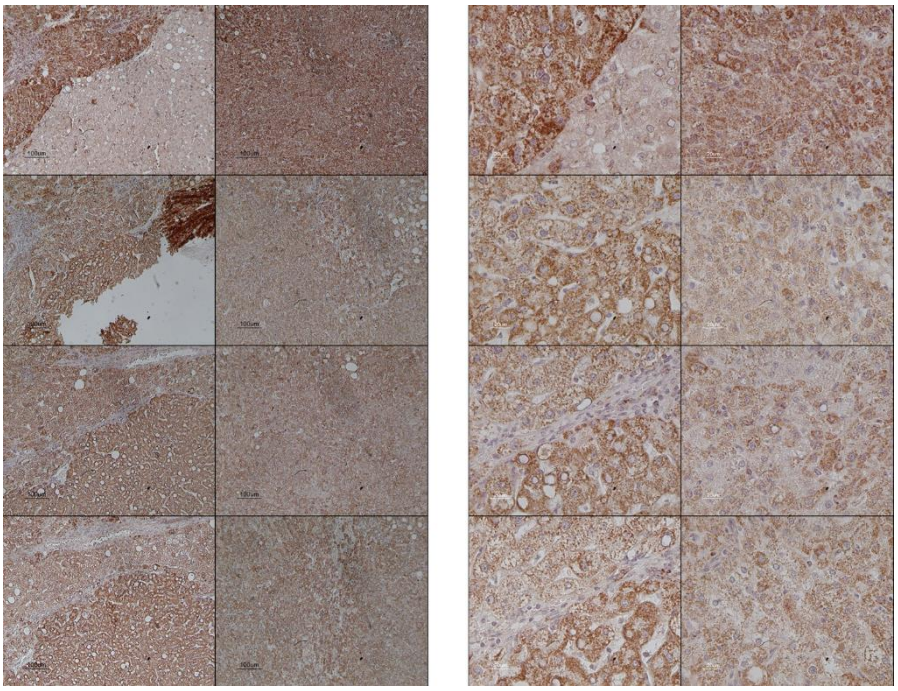
Non-tumoral Tumor

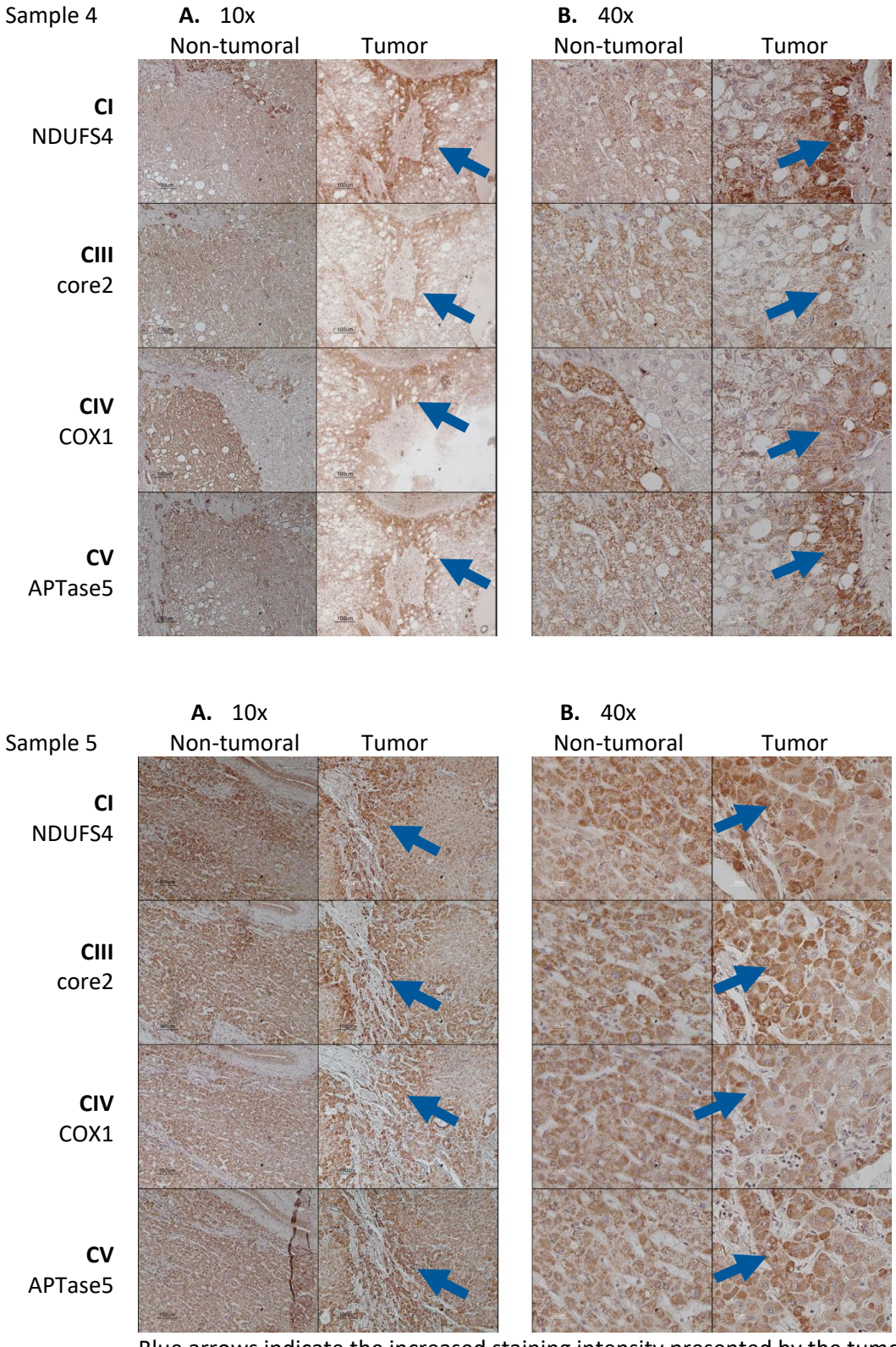
CI
NDUFS4

CIII
core2

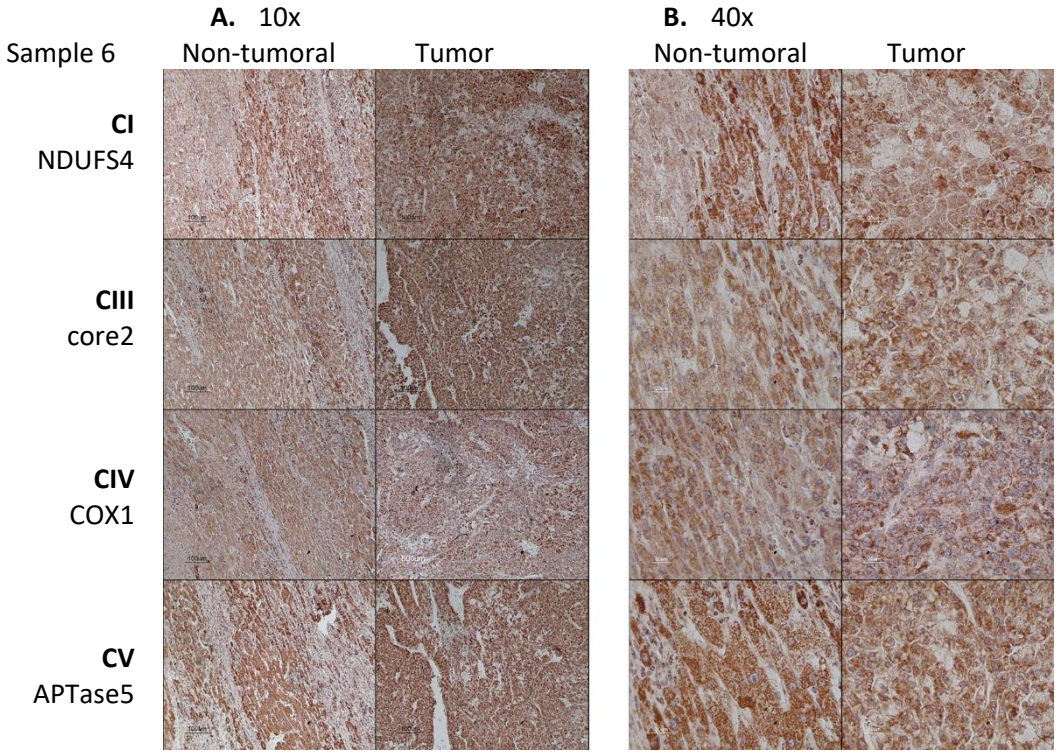
CIV
COX1

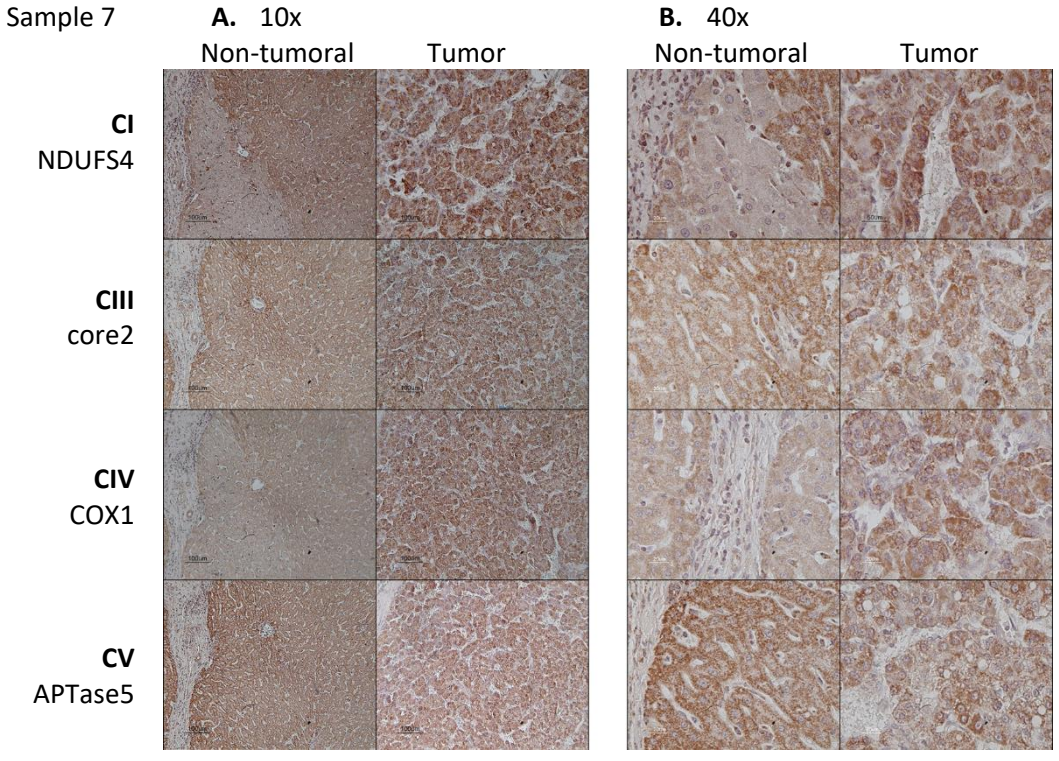
CV
APTase5

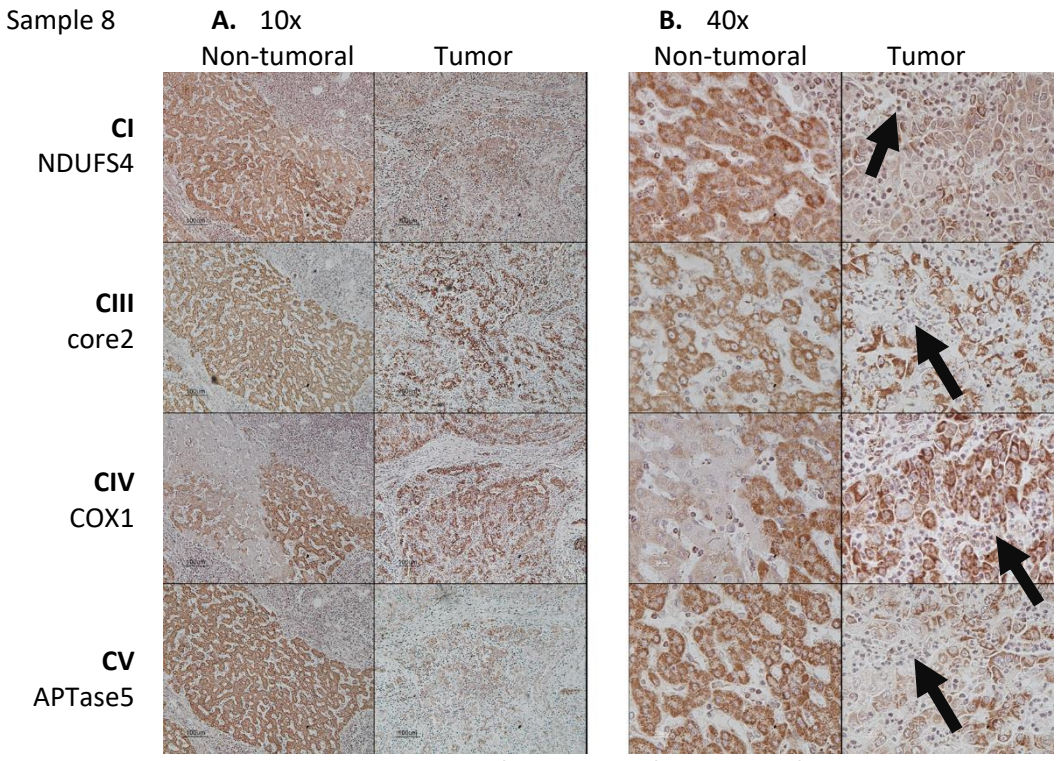




Blue arrows indicate the increased staining intensity presented by the tumoral tissue in the zones of contact with the stroma


A. 10x
B. 40x
Sample 6
Non-tumoral
Tumor
Non-tumoral
Tumor
CI
NDUFS4
CIII
core2
CIV
COX1
CV
APTase5


Sample 7
A. 10x
B. 40x
Non-tumoral
Tumor
Non-tumoral
Tumor
CI
NDUFS4
CIII
core2
CIV
COX1
CV
APTase5



Black arrows indicate a significant level of immune infiltrate in the tumoral tissue

Discussion

The study of mtDNA mutations and their functional impact in different diseases including cancer continues to remain a challenge despite technological advancements in the field of genetics and molecular biology.

In this work, we extracted mtDNA variants from WXS data of the TCGA LIHC cohort using MToolBox pipeline. We report a high number of mtDNA variants found in the analyzed HCC patients, among which 21% of the harbored tumor variants were deemed as pathogenic.

Discussing the impact of mtDNA mutations in cancer is difficult to summarize and estimating their impact on the clinical outcome of cancer patients is even more challenging due to several issues present in literature. First, many papers define a mtDNA mutation as any variant found in the tumor sample and not in the paired non-tumoral tissue. But this simple occurrence of a mtDNA variant in a tumor does not necessarily imply its pathogenicity. This aspect seems to not be fully understood by some authors. Also, some papers analyze only tumoral tissues, thus making it impossible to conclude whether the variants found are somatic or germline. Other approaches include deeming variants as "novel" if they were not found in the MITOMAP database, which although extensive, does not index all variants reported in healthy individuals. In the absence of a further evaluation and a proper pathogenicity scoring, these variants may be wrongly regarded as pathogenic by the authors. [104] Moreover, some reports have shown that in some papers the variational patterns which include different haplogroups sites being present in the same sample would be more likely explained by technical errors rather than real events.[105] A further problem is the focus of many works on variants present in the variable regions of the mtDNA, namely the D-loop. This site is the major regulatory region of the mtDNA and it contains many sites of functional relevance. The polymorphic sites are located inside the HVSI and HVSII regions which represent the markers used to assign the haplogroups before the era of the seuqencing of the entire mitochondrial genome. Thus, many variants reported here are also present in the normal healthy population. Another layer of complexity when dealing with mtDNA mutations is added by the presence of heteroplasmy. Thus, the simple relationship between the mere presence of a mutation causing a defective respiratory chain and the occurrence of a specific phenotypical sign has remained problematic as many authors fail to consider this aspect.

In this work we considered all the aforementioned limitations and provided a thorough analysis of the variants discovered. First, we included in the analysis only the patients for which good quality sequences were extracted from both tumoral and non-tumoral sample. This allows for a correct differentiation between germline variants and somatic ones. Next, we discarded samples which have shown different haplogroups as such a situation would be impossible to find in a normal setting. Moreover, the ambiguous assignment of more haplogroups does not guarantee good assembling and mitochondrial DNA coverage. Moreover, the non-synonumous mutations found were considered pathogenic based on the HmtVar disease score which uses the pathogenicity prediction given by six tools associated to the allele frequency observed in healthy subjects.[106,107,83] All of this, together with discarding of the variants with a low heteroplasmy ensure a minimal level of false positive variants.

The mitochondrial genome of HCC tumors shows an increased tendency of accumulating mutations, but the functional impact of such mutations is yet to be fully understood. Moreover, given the polyploid nature of the mtDNA, such mutations are often found in a heteroplasmic state. This allows cells having mtDNA mutations which offer a metabolic advantage to prevail, while those bearing unfavorable variants will subside. The result is seen as a shift in heteroplasmy and in our cohort was analyzed in the case of germline mutations. Although many variants presented a different value of HF between the non-tumoral and the tumoral tissue, only in the case of few mutations was this variation present to a greater extent. The three variants which shown a positive selection were considered likely polymorphic. A closer analysis in literature has found that these variants have been reported in different diseases such as diabetes, deafness or cancer, but given the lack of a proof for a functional role, a simple co-occurrence of a mtDNA variant with a disease is not enough. Even though the selection of a mtDNA variant in a tumor might lead to the instinctive thought that it could somehow favor tumoral development, mathematical models have shown that a variant can arrive to higher heteroplasmy levels, even homoplasmy, by chance.[108] Regarding negative selection one mutation in CV presented a decrease in HF in the tumoral sample when compared to the non-tumoral tissue. The reason behind the counterselection of such a variant might rely in the function of the complex itself. While mutations occurring in other mitochondrial complexes might cause a mild OXPHOS impairment, a mutation in CV, the place of ATP synthesis, could lead to a severe energy imbalance which would be detrimental for the cancer cell. But what about the other germline pathogenic mtDNA mutations which do not seem to present any variation in HF? In the HCC patients analyzed we observed six such mutations. There is interesting *in vivo* data

showing that some mutations occurring in CV might actually increase the tumor fitness favoring its growth as well as causing resistance to apoptosis. [109–111] This might also be the case of the mutations in CV reported in this study.

Some of the somatic mtDNA mutations found in the studied patients have been reported previously in literature in mitochondrial diseases. For example, the 3946A in the MT-ND1 gene and the 3243G mutation from the MT-TL1 gene were found present in patients diagnosed with MELAS, the latter accounting for up to 80% of diagnosed cases.[112] The 3243G mutation in the MT-TL1 gene could impact the synthesis of mitochondrial complexes subunits causing mitochondrial dysfunction and probably a deficient tumoral growth.

Another mutation found, the 13042A in the MT-ND5 gene, has been previously reported in patients with LHON-like optic neuropathy.[113]

The 14846A mutation in the MT-CYB gene previously reported in patients with exercise intolerance has also been found in one HCC tumor.[114]

Apart from the OXPHOS mutations found in the mtDNA genome, we also report few patients carrying mutations in the OXPHOS coding genes from the nuclear genome. We found five pathogenic mutation in CI and CV genes. Mutations in the OXPHOS genes encoded by the nucleus are often overlooked in cancer studies. Depending on their site of occurrence, mutations in the OXPHOS chain can have a different impact. While CI and CIII mutations might increase ROS levels and offer a tumoral advantage, mutations in CV can result in a decreased ATP supply which would represent a disadvantage for the cancer cells. Regarding the impact of mutations in nuclear OXPHOS genes, a study on families with a genetically established OXPHOS deficiency has shown that patients owning their disease to a mutation occurring in a nuclear gene coding for OXPHOS proteins have an earlier age of onset and present a more severe clinical course.[115] In the case of cancer patients, larger studies are required to be able to analyze the influence of such mutations on cancer progression.

One of the most studied proteins in cancer, p53 encoded by the TP53 gene, carries out numerous cellular roles including cell cycle regulation, DNA damage repair and cell death. TP53 is the most mutated gene in cancer showing a frequency of over 30% on the COSMIC database. These mutations are mainly missense mutations located in highly conserved functional parts of the protein, unlike the deletions of frameshifts usually reported in other tumor suppressor genes. This, together with the fact that until now there

are no reports of inactivation of p53 by hypermethylation hints toward the idea that mutant variants of p53 could play a role in tumorigenesis.[116],[117]. Even though in HCC *TP53* is the third most frequently mutated gene after *TTN* and *CTNNB*, this value remains rather low (20-30% vs 60-80% as in colon, lung or skin cancer for example). A recent study in cell reports comes forward to provide insight into this paradoxically retained wild-type p53 by showing its role in increasing glycolysis and suppressing pyruvate-driven oxidative phosphorylation by upregulating PUMA.[118] This finding comes in contradiction with the commonly accepted idea that wild-type p53 supports oxidative metabolism. Control of glucose metabolism is one of the many non-canonical roles of p53. It has been shown in many papers that p53 inhibits glycolysis by targeting multiple steps, even leading to an attenuation of the Warburg effect. First, it has a suppressive effect on the main cellular glucose transporters (GLUT1 GLUT3 and GLUT4) and represses the promoter of the insulin receptor (*INSR*) to limit the entry of glucose inside the cells. Moreover, by repressing the transporter of lactate, monocarboxylic acid transporter 1 (MCT1), it causes an build-up of lactate inside the cell which will consequently limit the glycolytic rate.[119] Another mechanism deployed to inhibit glycolysis is by inducing the TP53-induced glycolysis and apoptosis regulator (*TIGAR*). TIGAR dephosphorylates fructose-2,6-bisphosphate ($F2,6P_2$) into fructose-6-phosphate (F6P) which is later converted to fructose-1,6-bisphosphate ($F1,6P_2$) by phosphofructokinase-1 (PFK1). This conversion is a rate limiting reaction in glycolysis. Furthermore, p53 plays a role in the assembly of the CIV subunit, by activating the transcription of the cytochrome c oxidase assembly protein (SCO2).[120,121]

Another recently discovered role of p53 was that of proofreading and maintaining the integrity of the mtDNA genome, a role long thought to be held only by *POLG1*. This *POLG1* independent repair mechanism has been shown in pol-g proofreading deficient mice prone to the accumulation of mtDNA mutations [122] Because of this role we expected to find an increase in the number of mutations found in tumors harboring *TP53* mutations. While the mutated samples did show a small tendency towards an accumulation of mtDNA variants, it did not reach a statistically significant difference. This could be due to the small number of patients having a mutated *TP53* being compared with a much higher number of patients with the wild type. Another reason could be the fact that *TP53* mutations, with the exception of aflatoxinB1 induced HCCs, can occur as a late event in the tumorigenesis of HCC[123], thus the time required for a consequent build-up of mtDNA mutations is not reached.

The hotspot mutation R249S seen in 4 patients is associated with Aflatoxin B1 exposure which is a risk factor for HCC. The exposure data available on the CGD portal is scarce thus a direct association with this issue was not possible, but a study identified the Aflatoxin B1 Mutational signature (COSMIC signature 24 - https://cancer.sanger.ac.uk/cosmic/signatures_v2) is these patients[124]. Aflatoxin B1 exposure has been described to cause a significant level of mitochondrial dysfunction with induction of ROS generation and activation of apoptosis[125,126] but associations with mtDNA mutations have yet to be investigated.

Increased production of ROS could also be a possible contributor to the accumulation of mtDNA in tumors. Experiments on cell culture models show that a high concentration of ROS is needed in order to exert an oxidative damage to the mtDNA most probably due to the protective role of the TFAM histone.[127] The ROS induced mutational theory in cancer has been widely debated with some reports showing *in vivo* that high oxidative stress even in the presence of KOs of detoxifying genes does not induce a higher mutational rate.[58,128]. Since oxidative damage does not seem to offer a satisfactory explanation for the majority of the mutations seen, the next culprit would be POLG replication errors. This pattern is characterized by C>T and A>G substitutions. [129] Mitochondrial genome replication process is hypothesized to have a bidirectional initiation. When reading mitochondrial variants, we use the Revised Cambridge Reference Sequence which uses the light (L) strand of the mtDNA. This strand contains the sense sequence for most of the mitochondrial genes. If a mutation, as for example a C>T change, is to occur on the L stand of the mtDNA it will be read as it is. But if the same change occurred on the H stand, this would appear as a G>A change in our gene and therefore in our report. This is precisely the situation we observe in this study, with more than half of all variants showing this mutational pattern. Since POLG is the enzyme replicating the mtDNA, we examined whether the patients from this study presented any somatic mutations in this gene and in others forming the mitochondrial DNA replication machinery and found only three patients. This led us to conclude that in the case of HCC, the mtDNA mutations are not caused by a dysfunctional replication due to mutations in the genes for replication machinery.

Nevertheless, mutations in the replicative genes are not needed for the occurrence and accumulation of replication-related mtDNA mutations. These could be due to the increase mitochondrial biogenesis, thus the increase of POLG activity as it needs to keep the pace of the high tumoral proliferation rates. In such a situation even a wild-type replicaton machinery would allow the occurrence and accumulation of mtDNA mutations.[130] Some

authors suggest that C>T POLG replication errors occur more frequently on the H strand[131] thus increasing the number of G>A events on the L strand.

Another possible situation is that during mtDNA replication the displaced H strand is single stranded rendering it vulnerable to cytosine deamination.[49,132] This creates a C>T substitution on the H strand which will appear as a G>A change on our reference L strand. Moreover, in this situation the H strand is also prone to adenine deamination leading to a T>C change.[130] Indeed approximately 35% of our discovered variants are C>T and T>C alterations. In addition, there is also evidence for a different efficacy of the repair mechanisms between the H and the L strand.[133–135]

Another known source of DNA variation is represented by technical error introduced during the prepararion steps of NGS. It has been shown that the C>T/G>A change has been shown to have the highest error rate during the steps of the technique, event most likely due to the spontaneous deamination of a methylated cytosine to uracil[136]. Nevertheless, in the absence of a confirmation with a secondary technique such as Sanger sequencing, the contribution of NGS technical errors cannot be excluded.

Different pharmaceutical drugs as well as environmental toxins begin to be discovered as mitochondrial intoxicants.[137] The most well-known mutation induced by environmental toxins is the aflatoxinB induced mutation in the *TP53* gene. The liver is the main organ responsible for detoxification of the whole body. Frequent aggressions caused by chronic alcohol consumption, drug abuse as well as environmental factors can lead to liver damage and increased levels of ROS which could lead to a favorable environment for the acquisition of different mutations. Moreover, Chronic HBV and HCV infections have been found to disrupt the normal mitochondrial function. In our study we fail to find a difference in the distribution of the mtDNA mutations between the patients showing different backgrounds of HCC risks. Nevertheless, given the high frequency of comorbidities found in these patients that require the administration of different drugs which are mostly metabolized by the liver, one could also suppose that mtDNA mutations might arise as a consequence of the subsequent drug induced liver injury.

It has been widely accepted that the tumoral microenvironment plays an important role in tumoral development. One of its components, the tumoral infiltrating immune cells have been shown to be of high relevance for prognosis in cancer as it can predict clinical outcome of patients and their response to therapy. In HCC, there a growing body of

research in the past years as the usage of immunotherapy with immune checkpoint inhibitors expands also for these tumors. Here we find that the burden of somatic pathogenic mitochondrial DNA mutations located in genes coding for CI and CIII subunits influences the percentage of intratumoral lymphocytic infiltration.

We also see that patients with pathogenic mutations have a lower survival rate that those with non-pathogenic variants. Since immunohistochemical analysis is limited to the total lymphocyte infiltration, we could not asses the different subsets of these immune cells. Tumoral infiltrating lymphocytes have been shown to have bidirectional role in the development of HCC with a opposite impact of prognosis. For example, tumors showing a high density of the CD3+ antigen have a better outcome as evaluated by a longer disease free survival and overall survival[138]. The opposite situation in which tumors having a low density of CD3+ show a poor prognosis is also true.[139,140] Since CD3 is a common antigen expressed on all T-lymphocytes, these results are in accordance with our findings. Further detailed research is required to investigate this aspect and characterize the subsets of lymphocytes. Missense pathogenic mutations in CI and CIII are thought to cause an increase in ROS production which could raise the intratumoral level of oxidative stress [95]. Such a milieu has been shown to inactivate and reduce the survival of T cells ultimately leading to a diminished antitumoral activity. With T cell-based immunotherapy on the rise, understanding and counteracting the mechanisms which lead to induced T cell suppression and hyporesponsiveness is critical to ensure therapeutic success.

An interesting observation is that of the infiltration pattern presented by tumors harboring mutation in CV subunits. Even though we could not perform any statistical analysis because only one tumor presented a benign CV variant, we can see that this tumor had an insignificant level of stromal component while the others bearing pathogenic mutations presented up to 20% of stromal component. This could be due to the fact that mutations in ATP synthase subunits would cause an energetic deficit[141] which the tumor might compensate by attracting stromal cells such as cancer associated fibroblasts (CAFs) or tumor associated macrophages (TAMs) to aid their growth. This idea is supported also by the fact that we observed a negative selection of a CV germline pathogenic mutation. Another interesting observation following the same idea is that of mutations located in genes coding for transporter RNAs. Although failing to achieve statistical significance, tumors presenting pathogenic mutations in tRNA genes seem to have a higher stromal component as well as a higher level of immune infiltrate. This could be explained by the fact that tRNA mutations may cause an impairment in mitochondrial protein synthesis

consequently leading to a decrease in proteins constructing the OXPHOS chain and thus a decline in ATP production required for cellular function.[142,89] This, again, would render cancer cells dependent on the microenvironment for their proliferation. For example, CAFs have been shown to transfer to tumoral cells various molecules such as non-coding RNAs, different metabolites amino acids and whole proteins through exosomes in various types of cancers,[143–145] as well as in HCC.[146,147] Moreover, metabolites originating from such vesicles have been shown to regulate mitochondrial OXPHOS and glycolysis in recipient tumoral cells.[148] Furthermore, in breast cancer exosomes deriving from stromal cells have been shown to transfer mtDNA to tumoral cells, supporting OXPHOS and mediating the exit from chemotherapy induces dormancy.[149] Exosomes deriving from tumor-associated macrophages have also been shown to transfer proteins to cancer cells and contribute to tumor aggressiveness and metastasis.[145]

As far as the immune infiltration represented by monocytes and neutrophils, statistical analysis revealed no difference in infiltration rates between the tumors harboring pathogenic mutations and those without, although we can observe an overall general tendency for tumors with pathogenic variants to show a higher amount of these two components.

Neutrophils have been known to play an important role in chronic inflammatory diseases as well as in cancer. Recent studies have shown that they play vital functions in the tumoral milieu although the precise nature in different cancers is still under discussion[150]. In general, tumor associated neutrophils are considered to have a pro-tumorigenic effect facilitating tumoral progression and to even aid in establishing the premetastatic niche.[151] In liver cancer, neutrophils have been associated with a significant decrease in survival rates.[152]

Similarly to neutrophils, high infiltration rates of monocytes into the tumor milieu are associated with poor clinical outcome in different cancers such as breast, lung and ovarian cancers,[153–155] as well as in HCC.[156,157] TAMs which can derive from monocytes, have been shown to modulate the tumor microenvironment and decrease therapeutic success.[158,159] For example, tumors with CI deficiency have been shown to be capable of inducing angiogenesis by recruiting TAMs.[160]

The impact of mtDNA mutations on clinical outcome and survival of cancer patients has been under debate by many authors.[161–164] In this study survival analysis showed that in the TCGA-HCC cohort, mtDNA mutations can predict overall survival, as patients harboring

pathogenic mutations presented a lower survival rate than those without such variants. It is known that some mtDNA mutations induce mitochondrial dysfunction with consequent increase of ROS levels,[165] and HCC tumors are characterized by high levels of endogenous ROS.[166] This persistent oxidative stress could lead to DNA damage, activation of oncogenes ultimately causing cancer development. Moreover, ROS contribute to angiogenesis, activation of epithelial to mesenchymal transition and metastasis, all features found in HCC.[167] Even though the mutational signature found in the cohort of HCC samples analyzed in this study, was not of ROS origin, the increased generation of ROS as a consequence of mtDNA mutations is likely to contribute to the aggressiveness by facilitating tumoral proliferation and invasion, thus explaining the poor prognosis of these patients. In spite of such a positive correlation found in this study, because of the relatively small sample size, larger studies are required to strengthen the value mtDNA mutations as a prognostic factor in HCC.

The remarkable ability of the liver to regenerate after large hepatectomies in which more than half of the liver is removed has had researchers believe in the existence of a stem cell compartment in the liver.[168] A stem cell has two main features: the capacity to self-renewal and expand clonally. This has been shown by the fact that transplanted hepatocytes in a diseased liver have the capacity of expanding clonally. [169] Similar to the tracing of the origins and migration of human populations, cells can be traced back using mtDNA variants. Tracking differentiated cells to their stem cell progenitor has been shown has been shown in the case of colonic crypt cells which show the accumulation of mtDNA mutations particularly in CIV that originate from the stem cells laying in the base of the crypt.[170] On the same note, one study reports the presence of cytochrome c oxidase deficient patches in normal livers[171] similar to those we found in our study. The authors go further to report the presence of clonal mtDNA mutations in these nodules, further confirming the idea of that these cells originate from a common progenitor cell. Thus, one could think that a similar situation could be found also in our samples. Considering all the aforementioned arguments, the fact that we see a mixture of cells with low and high intensity signals for different mitochondrial complexes sustains the idea of a polyclonal origin of HCC tumors and is also in line with the well-known heterogeneity of this type of cancer.[172,173]

Conclusion

Hepatocellular carcinoma remains an aggressive tumor with a poor prognosis despite the medical advancements. Until present the treatment options available showing modest results combined with the advanced stage at diagnosis allows these tumors to remain a difficult and often incurable clinical challenge.

This study has analyzed the occurrence of mtDNA mutations in the LIHC-TCGA cohort and found pathogenic mutations in 21% of patients. In such cases, it may be hypothesized that mtDNA mutations could cause mitochondrial dysfunction and contribute to hepatocarcinogenesis, by increasing the production of ROS and oxidative stress, thus fueling the tumor machinery.

The search for a better therapy for liver cancer has proven to be a difficult challenge. The high heterogeneity of the tumor combined with the lack of established makers for therapeutic response have led to a great failure of many clinical trials. Thus, gaining further insight into the mechanisms which contribute to HCC development and exploring them in the context of therapy response represents a must in order to develop efficient therapeutic strategies. Here we found a correlation between mutations present in genes coding for CI subunits and the amount of stromal component and intratumoral infiltration of immune cells, factors which are known to support tumoral growth, invasion and metastasis. Moreover, we found a significant impact of mtDNA mutations on the overall survival, as the patients bearing tumors with pathogenic mutations showed a poorer survival.

Currently there is a limited understanding of the potential impact of mitochondrial dysfunction in tumoral aggressiveness and clinical outcome of HCC. In this setting, ongoing research is needed in order to completely understand the impact of mitochondrial metabolism and their targeting in the development and evolution of HCC.

Acknowledgements

Doing a PhD pretty much feels like a journey through a maze. Many moments in which you are either wandering around in circles or stuck, or even fighting off the occasional surprise Cthulhu which is popping out of nowhere. Luckily for me, I had my mentors helping and bringing me on the right path every time. My eternal gratitude goes to Fabio, Giuseppe and Ivana, who have been by my side at my best and most importantly at my worst. For your time, your patience, your scientific and moral support and for showing me the light at the end of the tunnel.

I am grateful to Giulia, Monica and Laura A., for their countless and most valuable pieces of advice and for their stimulating discussions.

I would like to thank my fellow maze runners, Nikkitha, Licia, Maria, Greta, Lorena, Sara and Silvia, for their cooperation, their friendship and all the fun and dramas we had in the past two years. And of course, for all the Nutella mascarpone moments…

A very special gratitude goes out to Laura G. and all the CRBa staff, especially to Catia, Francesca and Michele, who were the ones to initiate me on this road, and to Milena, for the constant encouragement and eternal source of cheerfulness.

I would also like to thank all my dear friends from home, for their enthusiasm and for making my visits home so much more joyful.

To Ruggero, for all the support and for giving my mind and heart a happy distraction outside of the lab.

To my family, for their constant care, whether it was in the form of a good word in a moment of need, or a heavy package full of cookies. And a special recognition goes to my mother, for her stubborn mitochondria, which gave me the power to carry on.

Last but not least, I would like to thank Alexandra for removing the barriers in the way of science and all my reviewers for their precious time.

I wish you all to live long and prosper.

Bibliography

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7-34. doi:10.3322/caac.21551

2. Terzi E, Salvatore V, Negrini G, Piscaglia F. Ongoing challenges in the diagnosis of hepatocellular carcinoma. *Expert Rev Gastroenterol Hepatol*. 2016;10(4):451-463. doi:10.1586/17474124.2016.1124758

3. Giannini EG, Farinati F, Ciccarese F, et al. Prognosis of untreated hepatocellular carcinoma. *Hepatology*. 2015;61(1):184-190. doi:10.1002/hep.27443

4. Parkin DM, Pisani P, Ferlay J. Global cancer statistics. *CA Cancer J Clin*. 49(1):33-64, 1. http://www.ncbi.nlm.nih.gov/pubmed/10200776. Accessed July 26, 2016.

5. McGlynn KA, London WT. The global epidemiology of hepatocellular carcinoma: present and future. *Clin Liver Dis*. 2011;15(2):223-243, vii-x. doi:10.1016/j.cld.2011.03.006

6. Llovet JM, Zucman-Rossi J, Pikarsky E, et al. Hepatocellular carcinoma. *Nat Rev Dis Prim*. 2016;2(1):16018. doi:10.1038/nrdp.2016.18

7. Sukowati CH, El-Khobar KE, Ie SI, Anfuso B, Muljono DH, Tiribelli C. Significance of hepatitis virus infection in the oncogenic initiation of hepatocellular carcinoma. *World J Gastroenterol*. 2016;22(4):1497. doi:10.3748/wjg.v22.i4.1497

8. Park B-J, Lee Y-J, Lee H-R. Chronic liver inflammation: clinical implications beyond alcoholic liver disease. *World J Gastroenterol*. 2014;20(9):2168-2175. doi:10.3748/wjg.v20.i9.2168

9. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017;23(6):703-713. doi:10.1038/nm.4333

10. Schulze K, Nault J, hepatology AV-J of, 2016 undefined. Genetic profiling of hepatocellular carcinoma using next-generation sequencing. *Elsevier*. https://www.sciencedirect.com/science/article/pii/S0168827816302495. Accessed February 12, 2020.

11. Goossens N, Sun X, Hoshida Y. Molecular classification of hepatocellular carcinoma: potential therapeutic implications. *Hepatic Oncol*. 2015;2(4):371-379. doi:10.2217/hep.15.26

12. Villanueva A. Hepatocellular Carcinoma. Longo DL, ed. *N Engl J Med*. 2019;380(15):1450-1462. doi:10.1056/NEJMra1713263

13. Shimada S, Mogushi K, Akiyama Y, et al. Comprehensive molecular and immunological

characterization of hepatocellular carcinoma. *EBioMedicine*. 2019;40:457-470. doi:10.1016/j.ebiom.2018.12.058

14. Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet (London, England)*. 2018;391(10126):1163-1173. doi:10.1016/S0140-6736(18)30207-1

15. Bruix J, Qin S, Merle P, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet (London, England)*. 2017;389(10064):56-66. doi:10.1016/S0140-6736(16)32453-9

16. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100(1):57-70. doi:10.1016/s0092-8674(00)81683-9

17. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell*. 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013

18. Fouad YA, Aanei C. Revisiting the hallmarks of cancer. *Am J Cancer Res*. 2017;7(5):1016-1036. http://www.ncbi.nlm.nih.gov/pubmed/28560055. Accessed October 25, 2019.

19. Warburg O, Wind F, Negelein E. THE METABOLISM OF TUMORS IN THE BODY. *J Gen Physiol*. 1927;8(6):519-530. doi:10.1085/jgp.8.6.519

20. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science (80-)*. 2009;324(5930):1029-1033. doi:10.1126/science.1160809

21. Anderson RG, Ghiraldeli LP, Pardee TS. Mitochondria in cancer metabolism, an organelle whose time has come? *Biochim Biophys Acta - Rev Cancer*. 2018;1870(1):96-102. doi:10.1016/J.BBCAN.2018.05.005

22. Porporato PE, Filigheddu N, Pedro JMB-S, Kroemer G, Galluzzi L. Mitochondrial metabolism and cancer. *Cell Res*. 2018;28(3):265-280. doi:10.1038/cr.2017.155

23. Cazzaniga M, Bonanni B. Relationship Between Metabolic Reprogramming and Mitochondrial Activity in Cancer Cells. Understanding The Anticancer Effect of Metformin and Its Clinical Implications. *Anticancer Res*. 2015;35(11):5789-5796. http://www.ncbi.nlm.nih.gov/pubmed/26503999. Accessed October 30, 2019.

24. Zhao R, Jiang S, Zhang L, Yu Z. Mitochondrial electron transport chain, ROS generation and uncoupling (Review). *Int J Mol Med*. 2019;44(1):3-15. doi:10.3892/ijmm.2019.4188

25. Kauppila JHK, Stewart JB. Mitochondrial DNA: Radically free of free-radical driven

mutations. *Biochim Biophys Acta - Bioenerg*. 2015;1847(11):1354-1361. doi:10.1016/J.BBABIO.2015.06.001

26. Chatterjee A, Mambo E, Sidransky D. Mitochondrial DNA mutations in human cancer. *Oncogene*. 2006;25(34):4663-4674. doi:10.1038/sj.onc.1209604

27. Cannino G, Ciscato F, Masgras I, Sánchez-Martín C, Rasola A. Metabolic Plasticity of Tumor Cell Mitochondria. *Front Oncol*. 2018;8:333. doi:10.3389/fonc.2018.00333

28. De Matteis S, Ragusa A, Marisi G, et al. Aberrant Metabolism in Hepatocellular Carcinoma Provides Diagnostic and Therapeutic Opportunities. *Oxid Med Cell Longev*. 2018;2018:1-13. doi:10.1155/2018/7512159

29. DeWaal D, Nogueira V, Terry AR, et al. Hexokinase-2 depletion inhibits glycolysis and induces oxidative phosphorylation in hepatocellular carcinoma and sensitizes to metformin. *Nat Commun*. 2018;9(1):446. doi:10.1038/s41467-017-02733-4

30. Trotta AP, Gelles JD, Serasinghe MN, Loi P, Arbiser JL, Chipuk JE. Disruption of mitochondrial electron transport chain function potentiates the pro-apoptotic effects of MAPK inhibition. *J Biol Chem*. 2017;292(28):11727-11739. doi:10.1074/jbc.M117.786442

31. Havas KM, Milchevskaya V, Radic K, et al. Metabolic shifts in residual breast cancer drive tumor recurrence. *J Clin Invest*. 2017;127(6):2091-2105. doi:10.1172/JCI89914

32. Viale A, Pettazzoni P, Lyssiotis CA, et al. Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. *Nature*. 2014;514(7524):628-632. doi:10.1038/nature13611

33. Shen Y-C, Ou D-L, Hsu C, et al. Activating oxidative phosphorylation by a pyruvate dehydrogenase kinase inhibitor overcomes sorafenib resistance of hepatocellular carcinoma. *Br J Cancer*. 2013;108(1):72-81. doi:10.1038/bjc.2012.559

34. Zong W-X, Rabinowitz JD, White E. Mitochondria and Cancer. *Mol Cell*. 2016;61(5):667-676. doi:10.1016/j.molcel.2016.02.011

35. Cavalli LR, Varella-Garcia M, Liang BC. Diminished tumorigenic phenotype after depletion of mitochondrial DNA. *Cell Growth Differ*. 1997;8(11):1189-1198. http://www.ncbi.nlm.nih.gov/pubmed/9372242. Accessed October 25, 2019.

36. Pustylnikov S, Costabile F, Beghi S, Facciabene A. Targeting mitochondria in cancer: current concepts and immunotherapy approaches. *Transl Res*. 2018;202:35-51. doi:10.1016/j.trsl.2018.07.013

37. Rozanov D, Cheltsov A, Nilsen A, et al. Targeting mitochondria in cancer therapy could provide a basis for the selective anti-cancer activity. Ahmad A, ed. *PLoS One*. 2019;14(3):e0205623. doi:10.1371/journal.pone.0205623

38. Fulda S, Galluzzi L, Kroemer G. Targeting mitochondria for cancer therapy. *Nat Rev Drug Discov*. 2010;9(6):447-464. doi:10.1038/nrd3137

39. Wheaton WW, Weinberg SE, Hamanaka RB, et al. Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife*. 2014;3:e02242. doi:10.7554/eLife.02242

40. Donadon V, Balbi M, Mas MD, Casarin P, Zanette G. Metformin and reduced risk of hepatocellular carcinoma in diabetic patients with chronic liver disease. *Liver Int*. 2010;30(5):750-758. doi:10.1111/j.1478-3231.2010.02223.x

41. Chen H-P, Shieh J-J, Chang C-C, et al. Metformin decreases hepatocellular carcinoma risk in a dose-dependent manner: population-based and in vitro studies. *Gut*. 2013;62(4):606-615. doi:10.1136/GUTJNL-2011-301708

42. Hu L, Zeng Z, Xia Q, et al. Metformin attenuates hepatoma cell proliferation by decreasing glycolytic flux through the HIF-1α/PFKFB3/PFK1 pathway. *Life Sci*. 2019;239:116966. doi:10.1016/J.LFS.2019.116966

43. You A, Cao M, Guo Z, et al. Metformin sensitizes sorafenib to inhibit postoperative recurrence and metastasis of hepatocellular carcinoma in orthotopic mouse models. *J Hematol Oncol*. 2016;9(1):20. doi:10.1186/s13045-016-0253-6

44. Casadei Gardini A, Marisi G, Scarpi E, et al. Effects of metformin on clinical outcome in diabetic patients with advanced HCC receiving sorafenib. *Expert Opin Pharmacother*. 2015;16(18):2719-2725. doi:10.1517/14656566.2015.1102887

45. Casadei Gardini A, Faloppi L, De Matteis S, et al. Metformin and insulin impact on clinical outcome in patients with advanced hepatocellular carcinoma receiving sorafenib: Validation study and biological rationale. *Eur J Cancer*. 2017;86:106-114. doi:10.1016/j.ejca.2017.09.003

46. Homo sapiens mitochondrion, complete genome; https://www.ncbi.nlm.nih.gov/nuccore/251831106.

47. Revised Cambridge Reference Sequence (rCRS) of the Human Mitochondrial DNA ; https://www.mitomap.org/foswiki/bin/view/MITOMAP/HumanMitoStrands.

48. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet*. 1999;23(2):147-147. doi:10.1038/13779

49. Holt IJ, Reyes A. Human mitochondrial DNA replication. *Cold Spring Harb Perspect Biol*. 2012;4(12). doi:10.1101/cshperspect.a012971

50. Chinnery PF, Hudson G. Mitochondrial genetics. *Br Med Bull*. 2013;106(1):135-159. doi:10.1093/bmb/ldt017

51. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes. *Investig Genet*. 2015;6:3. doi:10.1186/s13323-015-0022-2

52. Torroni A, Schurr TG, Cabell MF, et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet*. 1993;53(3):563-590. http://www.ncbi.nlm.nih.gov/pubmed/7688932. Accessed October 25, 2019.

53. Mohamed Yusoff A. Role of mitochondrial DNA mutations in brain tumors: A mini-review. *J Cancer Res Ther*. 2015;11(3):535. doi:10.4103/0973-1482.161925

54. Rossignol R, Faustin B, Rocher C, Malgat M, Mazat J-P, Letellier T. Mitochondrial threshold effects. *Biochem J*. 2003;370(Pt 3):751-762. doi:10.1042/BJ20021594

55. Https://bmcbiol.biomedcentral.com/articles/10.1186/s12915-019-0668-y/figures/4. No Title.

56. Wang L. Mitochondrial purine and pyrimidine metabolism and beyond. *Nucleosides, Nucleotides and Nucleic Acids*. 2016;35(10-12):578-594. doi:10.1080/15257770.2015.1125001

57. Friedman JR, Nunnari J. Mitochondrial form and function. *Nature*. 2014;505(7483):335-343. doi:10.1038/nature12985

58. Kauppila JHK, Bonekamp NA, Mourier A, et al. Base-excision repair deficiency alone or combined with increased oxidative stress does not increase mtDNA point mutations in mice. *Nucleic Acids Res*. 2018;46(13):6642-6669. doi:10.1093/nar/gky456

59. Liu Q, Lin D, Li M, Gu Z, Zhao Y. Evidence of Neutral Evolution of Mitochondrial DNA in Human Hepatocellular Carcinoma. Enard D, ed. *Genome Biol Evol*. 2019;11(10):2909-2916. doi:10.1093/gbe/evz214

60. Hertweck KL, Dasgupta S. The Landscape of mtDNA Modifications in Cancer: A Tale of Two Cities. *Front Oncol*. 2017;7:262. doi:10.3389/fonc.2017.00262

61. Ju YS, Alexandrov LB, Gerstung M, et al. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife*. 2014;3. doi:10.7554/eLife.02935

62. Chatterjee A, Mambo E, Sidransky D. Mitochondrial DNA mutations in human cancer. *Oncogene*. 2006;25(34):4663-4674. doi:10.1038/sj.onc.1209604

63. Errichiello E, Venesio T. Mitochondrial DNA Variations in Tumors: Drivers or Passengers? In: *Mitochondrial DNA - New Insights*. InTech; 2018. doi:10.5772/intechopen.75188

64. Park JS, Sharma LK, Li H, et al. A heteroplasmic, not homoplasmic, mitochondrial DNA mutation promotes tumorigenesis via alteration in reactive oxygen species generation and apoptosis. *Hum Mol Genet*. 2009;18(9):1578-1589.

doi:10.1093/hmg/ddp069

65. Gasparre G, Kurelac I, Capristo M, et al. A Mutation Threshold Distinguishes the Antitumorigenic Effects of the Mitochondrial Gene MTND1, an Oncojanus Function. *Cancer Res*. 2011;71(19):6220-6229. doi:10.1158/0008-5472.CAN-11-1042

66. Andréasson C, Ott M, Büttner S. Mitochondria orchestrate proteostatic and metabolic stress responses. *EMBO Rep*. 2019;20(10). doi:10.15252/embr.201947865

67. Picard M, McEwen BS, Epel ES, Sandi C. An energetic view of stress: Focus on mitochondria. *Front Neuroendocrinol*. 2018;49:72-85. doi:10.1016/J.YFRNE.2018.01.001

68. Wheelhouse NM, Lai PBS, Wigmore SJ, Ross JA, Harrison DJ. Mitochondrial D-loop mutations and deletion profiles of cancerous and noncancerous liver tissue in hepatitis B virus-infected liver. *Br J Cancer*. 2005;92(7):1268-1272. doi:10.1038/sj.bjc.6602496

69. Yin P-H, Wu C-C, Lin J-C, Chi C-W, Wei Y-H, Lee H-C. Somatic mutations of mitochondrial genome in hepatocellular carcinoma. *Mitochondrion*. 2010;10(2):174-182. doi:10.1016/j.mito.2009.12.147

70. Yamamoto H, Tanaka M, Katayama M, Obayashi T, Nimura Y, Ozawa T. Significant existence of deleted mitochondrial DNA in cirrhotic liver surrounding hepatic tumor. *Biochem Biophys Res Commun*. 1992;182(2):913-920. doi:10.1016/0006-291x(92)91819-c

71. Yin PH, Lee HC, Chau GY, et al. Alteration of the copy number and deletion of mitochondrial DNA in human hepatocellular carcinoma. *Br J Cancer*. 2004;90(12):2390-2396. doi:10.1038/sj.bjc.6601838

72. Calabrese C, Simone D, Diroma MA, et al. MToolBox: a highly automated pipeline for heteroplasmy annotation and prioritization analysis of human mitochondrial variants in high-throughput sequencing. *Bioinformatics*. 2014;30(21):3115-3117. doi:10.1093/bioinformatics/btu483

73. Li B, Krishnan VG, Mort ME, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics*. 2009;25(21):2744-2750. doi:10.1093/bioinformatics/btp528

74. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249. doi:10.1038/nmeth0410-248

75. Calabrese R, Capriotti E, Fariselli P, Martelli PL, Casadio R. Functional annotations improve the predictive score of human disease-related mutations in proteins. *Hum Mutat*. 2009;30(8):1237-1244. doi:10.1002/humu.21047

76. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nat Commun*. 2015;6(1):8971. doi:10.1038/ncomms9971

77. https://portal.gdc.cancer.gov/.

78. Anaconda Software Distribution. Computer software. Vers. 2-2.4.0. Anaconda, Nov. 2016. Web. <https://anaconda.com>.

79. Mounir M, Lucchetta M, Silva TC, et al. New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx. Wang E, ed. *PLOS Comput Biol*. 2019;15(3):e1006701. doi:10.1371/journal.pcbi.1006701

80. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44(8):e71-e71. doi:10.1093/nar/gkv1507

81. Silva TC, Colaprico A, Olsen C, et al. TCGA Workflow: Analyze cancer genomics and epigenomics data using Bioconductor packages. *F1000Research*. 2016;5:1542. doi:10.12688/f1000research.8923.2

82. Wan Y-W, Allen GI, Liu Z. TCGA2STAT: simple TCGA data access for integrated statistical analysis in R. *Bioinformatics*. 2016;32(6):952-954. doi:10.1093/bioinformatics/btv677

83. Santorsola M, Calabrese C, Girolimetti G, Diroma MA, Gasparre G, Attimonelli M. A multi-parametric workflow for the prioritization of mitochondrial DNA variants of clinical interest. *Hum Genet*. 2016;135(1):121-136. doi:10.1007/s00439-015-1615-9

84. Grandhi S, Bosworth C, Maddox W, et al. Heteroplasmic shifts in tumor mitochondrial genomes reveal tissue-specific signals of relaxed and positive selection. *Hum Mol Genet*. 2017;26(15):2912-2922. doi:10.1093/hmg/ddx172

85. Connolly JL, Schnitt SJ, Wang HH, Longtine JA, Dvorak A, Dvorak HF. Tumor Structure and Tumor Stroma Generation. 2003. https://www.ncbi.nlm.nih.gov/books/NBK13447/. Accessed October 13, 2019.

86. Liu S-M, Zhou X, Zheng F, et al. Novel mutations found in mitochondrial diabetes in Chinese Han population. *Diabetes Res Clin Pract*. 2007;76(3):425-435. doi:10.1016/j.diabres.2006.09.032

87. Jing Z, Yan–chun J, Min–Xin G. Mitochondrial tRNA mutations associated with deafness. *J Otol*. 2012;7(1):36-44. doi:10.1016/S1672-2930(12)50009-7

88. Ma Y-Y, Wu T-F, Liu Y-P, et al. Genetic and biochemical findings in Chinese children with Leigh syndrome. *J Clin Neurosci*. 2013;20(11):1591-1594. doi:10.1016/j.jocn.2013.03.034

89. He ZF, Zheng LC, Xie DY, Yu SS, Zhao J. Mutational analysis of mitochondrial tRNA

genes in patients with lung cancer. *Balkan J Med Genet*. 2016;19(2):45-50. doi:10.1515/bjmg-2016-0035

90. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res*. 2019;79(18):4557-4566. doi:10.1158/0008-5472.CAN-18-3962

91. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. *J Cell Sci*. 2010;123(Pt 24):4195-4200. doi:10.1242/jcs.023820

92. Tormoen GW, Crittenden MR, Gough MJ. Role of the immunosuppressive microenvironment in immunotherapy. *Adv Radiat Oncol*. 2018;3(4):520. doi:10.1016/J.ADRO.2018.08.018

93. Kobold S. Innate and adaptive immunity combined for cancer treatment. *Proc Natl Acad Sci U S A*. 2019;116(4):1087-1088. doi:10.1073/pnas.1820166116

94. Wang J-C, Xu Y, Huang Z-M, Lu X-J. T cell exhaustion in cancer: Mechanisms and clinical implications. *J Cell Biochem*. 2018;119(6):4279-4286. doi:10.1002/jcb.26645

95. Chen X, Song M, Zhang B, Zhang Y. Reactive Oxygen Species Regulate T Cell Immune Response in the Tumor Microenvironment. *Oxid Med Cell Longev*. 2016;2016:1580967. doi:10.1155/2016/1580967

96. Ghelli A, Tropeano C V., Calvaruso MA, et al. The cytochrome b p.278Y>C mutation causative of a multisystem disorder enhances superoxide production and alters supramolecular interactions of respiratory chain complexes. *Hum Mol Genet*. 2013;22(11):2141-2151. doi:10.1093/hmg/ddt067

97. Young MJ, Copeland WC. Human mitochondrial DNA replication machinery and disease. *Curr Opin Genet Dev*. 2016;38:52-62. doi:10.1016/j.gde.2016.03.005

98. Copeland WC, Ponamarev M V, Nguyen D, Kunkel TA, Longley MJ. Mutations in DNA polymerase gamma cause error prone DNA synthesis in human mitochondrial disorders. *Acta Biochim Pol*. 2003;50(1):155-167. doi:035001155

99. Singh KK, Ayyasamy V, Owens KM, Koul MS, Vujcic M. Mutations in mitochondrial DNA polymerase-γ promote breast tumorigenesis. *J Hum Genet*. 2009;54(9):516-524. doi:10.1038/jhg.2009.71

100. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*. 2017;45(D1):D353-D361. doi:10.1093/nar/gkw1092

101. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res*. 2016;44(D1):D1251-D1257. doi:10.1093/nar/gkv1003

102. Marco-Brualla J, Al-Wasaby S, Soler R, et al. Mutations in the ND2 Subunit of Mitochondrial Complex I Are Sufficient to Confer Increased Tumorigenic and Metastatic Potential to Cancer Cells. *Cancers (Basel)*. 2019;11(7):1027. doi:10.3390/cancers11071027

103. Porcelli AM, Ghelli A, Ceccarelli C, et al. The genetic and metabolic signature of oncocytic transformation implicates HIF1α destabilization. *Hum Mol Genet*. 2010;19(6):1019-1032. doi:10.1093/hmg/ddp566

104. Bandelt H-J, Salas A, Taylor RW, Yao Y-G. Exaggerated status of "novel" and "pathogenic" mtDNA sequence variants due to inadequate database searches. *Hum Mutat*. 2009;30(2):191-196. doi:10.1002/humu.20846

105. Bandelt H-J, Salas A. Contamination and sample mix-up can best explain some patterns of mtDNA instabilities in buccal cells and oral squamous cell carcinoma. *BMC Cancer*. 2009;9(1):113. doi:10.1186/1471-2407-9-113

106. Preste R, Vitale O, Clima R, Gasparre G, Attimonelli M. HmtVar: a new resource for human mitochondrial variations and pathogenicity data. *Nucleic Acids Res*. 2019;47(D1):D1202-D1210. doi:10.1093/nar/gky1024

107. Diroma MA, Lubisco P, Attimonelli M. A comprehensive collection of annotations to interpret sequence variation in human mitochondrial transfer RNAs. *BMC Bioinformatics*. 2016;17(S12):338. doi:10.1186/s12859-016-1193-4

108. Chinnery PF, Samuels DC, Elson J, Turnbull DM. Accumulation of mitochondrial DNA mutations in ageing, cancer, and mitochondrial disease: is there a common mechanism? *Lancet (London, England)*. 2002;360(9342):1323-1325. doi:10.1016/S0140-6736(02)11310-9

109. Arnold RS, Sun CQ, Richards JC, et al. Mitochondrial DNA mutation stimulates prostate cancer growth in bone stromal environment. *Prostate*. 2009;69(1):1-11. doi:10.1002/pros.20854

110. Petros JA, Baumann AK, Ruiz-Pesini E, et al. mtDNA mutations increase tumorigenicity in prostate cancer. *Proc Natl Acad Sci*. 2005;102(3):719-724. doi:10.1073/pnas.0408894102

111. Shidara Y, Yamagata K, Kanamori T, et al. Positive Contribution of Pathogenic Mutations in the Mitochondrial Genome to the Promotion of Cancer by Prevention from Apoptosis. *Cancer Res*. 2005;65(5):1655-1663. doi:10.1158/0008-5472.CAN-04-2012

112. Kirby DM, McFarland R, Ohtake A, et al. Mutations of the mitochondrial ND1 gene as a cause of MELAS. *J Med Genet*. 2004;41(10):784-789. doi:10.1136/jmg.2004.020537

113. Valentino ML, Barboni P, Rengo C, et al. The 13042G --> A/ND5 mutation in mtDNA is

pathogenic and can be associated also with a prevalent ocular phenotype. *J Med Genet*. 2006;43(7):e38. doi:10.1136/jmg.2005.037507

114. Andreu AL, Hanna MG, Reichmann H, et al. Exercise intolerance due to mutations in the cytochrome b gene of mitochondrial DNA. *N Engl J Med*. 1999;341(14):1037-1044. doi:10.1056/NEJM199909303411404

115. Rubio-Gozalbo ME, Dijkman KP, van den Heuvel LP, Sengers RCA, Wendel U, Smeitink JAM. Clinical differences in patients with mitochondriocytopathies due to nuclear versus mitochondrial DNA mutations. *Hum Mutat*. 2000;15(6):522-532. doi:10.1002/1098-1004(200006)15:6<522::AID-HUMU4>3.0.CO;2-Y

116. Olivier M, Hollstein M, Hainaut P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb Perspect Biol*. 2010;2(1):a001008-a001008. doi:10.1101/cshperspect.a001008

117. Kamp WM, Wang P-Y, Hwang PM. TP53 mutation, mitochondria and cancer. *Curr Opin Genet Dev*. 2016;38:16-22. doi:10.1016/j.gde.2016.02.007

118. Kim J, Yu L, Chen W, et al. Wild-Type p53 Promotes Cancer Metabolic Switch by Inducing PUMA-Dependent Suppression of Oxidative Phosphorylation. *Cancer Cell*. 2019;35(2):191-203.e8. doi:10.1016/j.ccell.2018.12.012

119. Itahana Y, Itahana K. Emerging Roles of p53 Family Members in Glucose Metabolism. *Int J Mol Sci*. 2018;19(3). doi:10.3390/ijms19030776

120. Liu J, Zhang C, Hu W, Feng Z. Tumor suppressor p53 and metabolism. Verma CS, ed. *J Mol Cell Biol*. 2019;11(4):284-292. doi:10.1093/jmcb/mjy070

121. Gomes AS, Ramos H, Soares J, Saraiva L. p53 and glucose metabolism: an orchestra to be directed in cancer therapy. *Pharmacol Res*. 2018;131:75-86. doi:10.1016/J.PHRS.2018.03.015

122. Safdar A, Khrapko K, Flynn JM, et al. Exercise-induced mitochondrial p53 repairs mtDNA mutations in mutator mice. *Skelet Muscle*. 2015;6(1):7. doi:10.1186/s13395-016-0075-9

123. Niu Z-S, Niu X-J, Wang W-H. Genetic alterations in hepatocellular carcinoma: An update. *World J Gastroenterol*. 2016;22(41):9069-9095. doi:10.3748/wjg.v22.i41.9069

124. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu TCGAR, Cancer Genome Atlas Research Network DA, Roberts LR. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell*. 2017;169(7):1327-1341.e23. doi:10.1016/j.cell.2017.05.046

125. Liu Y, Wang W. Aflatoxin B1 impairs mitochondrial functions, activates ROS generation, induces apoptosis and involves Nrf2 signal pathway in primary broiler

hepatocytes. *Anim Sci J*. 2016;87(12):1490-1500. doi:10.1111/asj.12550

126. Rotimi OA, Rotimi SO, Goodrich JM, Adelani IB, Agbonihale E, Talabi G. Time-Course Effects of Acute Aflatoxin B1 Exposure on Hepatic Mitochondrial Lipids and Oxidative Stress in Rats. *Front Pharmacol*. 2019;10:467. doi:10.3389/fphar.2019.00467

127. Zsurka G, Peeva V, Kotlyar A, Kunz WS. Is There Still Any Role for Oxidative Stress in Mitochondrial DNA-Dependent Aging? *Genes (Basel)*. 2018;9(4). doi:10.3390/genes9040175

128. Gammage PA, Frezza C. Mitochondrial DNA: the overlooked oncogenome? *BMC Biol*. 2019;17(1):53. doi:10.1186/s12915-019-0668-y

129. Lynch M. The lower bound to the evolution of mutation rates. *Genome Biol Evol*. 2011;3:1107-1118. doi:10.1093/gbe/evr066

130. Ju YS, Alexandrov LB, Gerstung M, et al. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *Elife*. 2014;3. doi:10.7554/eLife.02935

131. Zheng W, Khrapko K, Coller HA, Thilly WG, Copeland WC. Origins of human mitochondrial point mutations as DNA polymerase γ-mediated errors. *Mutat Res Mol Mech Mutagen*. 2006;599(1-2):11-20. doi:10.1016/j.mrfmmm.2005.12.012

132. Faith JJ, Pollock DD. Likelihood analysis of asymmetrical mutation bias gradients in vertebrate mitochondrial genomes. *Genetics*. 2003;165(2):735-745. http://www.ncbi.nlm.nih.gov/pubmed/14573484. Accessed October 22, 2019.

133. Yasukawa T, Yang M-Y, Jacobs HT, Holt IJ. A Bidirectional Origin of Replication Maps to the Major Noncoding Region of Human Mitochondrial DNA. *Mol Cell*. 2005;18(6):651-662. doi:10.1016/j.molcel.2005.05.002

134. Yasukawa T, Reyes A, Cluett TJ, et al. Replication of vertebrate mitochondrial DNA entails transient ribonucleotide incorporation throughout the lagging strand. *EMBO J*. 2006;25(22):5358-5371. doi:10.1038/sj.emboj.7601392

135. Pavlov YI, Mian IM, Kunkel TA. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol*. 2003;13(9):744-748. doi:10.1016/s0960-9822(03)00284-7

136. Ma X, Shao Y, Tian L, et al. Analysis of error profiles in deep next-generation sequencing data. *Genome Biol*. 2019;20(1):50. doi:10.1186/s13059-019-1659-6

137. Meyer JN, Leung MCK, Rooney JP, et al. Mitochondria as a Target of Environmental Toxicants. *Toxicol Sci*. 2013;134(1):1-17. doi:10.1093/toxsci/kft102

138. Petrizzo A, Mauriello A, Tornesello ML, Buonaguro FM, Tagliamonte M, Buonaguro L. Cellular prognostic markers in hepatitis-related hepatocellular carcinoma. *Infect*

Agent Cancer. 2018;13(1):10. doi:10.1186/s13027-018-0183-8

139. Tian M-X, Liu W-R, Wang H, et al. Tissue-infiltrating lymphocytes signature predicts survival in patients with early/intermediate stage hepatocellular carcinoma. *BMC Med*. 2019;17(1):106. doi:10.1186/s12916-019-1341-6

140. Gabrielson A, Wu Y, Wang H, et al. Intratumoral CD3 and CD8 T-cell Densities Associated with Relapse-Free Survival in HCC. *Cancer Immunol Res*. 2016;4(5):419-430. doi:10.1158/2326-6066.CIR-15-0110

141. Jonckheere AI, Smeitink JAM, Rodenburg RJT. Mitochondrial ATP synthase: architecture, function and pathology. *J Inherit Metab Dis*. 2012;35(2):211-225. doi:10.1007/s10545-011-9382-9

142. Lu J, Sharma LK, Bai Y. Implications of mitochondrial DNA mutations and mitochondrial dysfunction in tumorigenesis. *Cell Res*. 2009;19(7):802-815. doi:10.1038/cr.2009.69

143. Yang X, Li Y, Zou L, Zhu Z. Role of Exosomes in Crosstalk Between Cancer-Associated Fibroblasts and Cancer Cells. *Front Oncol*. 2019;9:356. doi:10.3389/fonc.2019.00356

144. Richards KE, Zeleniak AE, Fishel ML, Wu J, Littlepage LE, Hill R. Cancer-associated fibroblast exosomes regulate survival and proliferation of pancreatic cancer cells. *Oncogene*. 2017;36(13):1770-1778. doi:10.1038/onc.2016.353

145. Tai Y-L, Chen K-C, Hsieh J-T, Shen T-L. Exosomes in cancer development and clinical applications. *Cancer Sci*. 2018;109(8):2364-2374. doi:10.1111/cas.13697

146. He M, Qin H, Poon TCW, et al. Hepatocellular carcinoma-derived exosomes promote motility of immortalized hepatocyte through transfer of oncogenic proteins and RNAs. *Carcinogenesis*. 2015;36(9):1008-1018. doi:10.1093/carcin/bgv081

147. Wu Q, Zhou L, Lv D, Zhu X, Tang H. Exosome-mediated communication in the tumor microenvironment contributes to hepatocellular carcinoma development and progression. *J Hematol Oncol*. 2019;12(1):53. doi:10.1186/s13045-019-0739-0

148. Zhao H, Yang L, Baddour J, et al. Tumor microenvironment derived exosomes pleiotropically modulate cancer cell metabolism. *Elife*. 2016;5:e10250. doi:10.7554/eLife.10250

149. Sansone P, Savini C, Kurelac I, et al. Packaging and transfer of mitochondrial DNA via exosomes regulate escape from dormancy in hormonal therapy-resistant breast cancer. *Proc Natl Acad Sci U S A*. 2017;114(43):E9066-E9075. doi:10.1073/pnas.1704862114

150. Jeong J, Suh Y, Jung K. Context Drives Diversification of Monocytes and Neutrophils in Orchestrating the Tumor Microenvironment. *Front Immunol*. 2019;10:1817.

doi:10.3389/fimmu.2019.01817

151. Woolbright BL, Jaeschke H. *Xenobiotic and Endobiotic Mediated Interactions Between the Cytochrome P450 System and the Inflammatory Response in the Liver*. 1st ed. Elsevier Inc.; 2015. doi:10.1016/bs.apha.2015.04.001

152. Margetts J, Ogle LF, Chan SL, et al. Neutrophils: driving progression and poor prognosis in hepatocellular carcinoma? *Br J Cancer*. 2018;118(2):248-257. doi:10.1038/bjc.2017.386

153. Qian B, Deng Y, Im JH, et al. A Distinct Macrophage Population Mediates Metastatic Breast Cancer Cell Extravasation, Establishment and Growth. Bereswill S, ed. *PLoS One*. 2009;4(8):e6562. doi:10.1371/journal.pone.0006562

154. Qian B-Z, Pollard JW. Macrophage Diversity Enhances Tumor Progression and Metastasis. *Cell*. 2010;141(1):39-51. doi:10.1016/j.cell.2010.03.014

155. Zhang Q, Liu L, Gong C, et al. Prognostic Significance of Tumor-Associated Macrophages in Solid Tumor: A Meta-Analysis of the Literature. Hoque MO, ed. *PLoS One*. 2012;7(12):e50946. doi:10.1371/journal.pone.0050946

156. Feng F, Zheng G, Wang Q, et al. Low lymphocyte count and high monocyte count predicts poor prognosis of gastric cancer. *BMC Gastroenterol*. 2018;18(1):148. doi:10.1186/s12876-018-0877-9

157. Lin G-N, Peng J-W, Xiao J, Liu D-Y, Xia Z-J. Prognostic impact of circulating monocytes and lymphocyte-to-monocyte ratio on previously untreated metastatic non-small cell lung cancer patients receiving platinum-based doublet. *Med Oncol*. 2014;31(7):70. doi:10.1007/s12032-014-0070-0

158. Fu X-T, Song K, Zhou J, et al. Tumor-associated macrophages modulate resistance to oxaliplatin via inducing autophagy in hepatocellular carcinoma. *Cancer Cell Int*. 2019;19(1):71. doi:10.1186/s12935-019-0771-8

159. Chen R, Xu X, Tao Y, Qian Z, Yu Y. Exosomes in hepatocellular carcinoma: a new horizon. *Cell Commun Signal*. 2019;17(1):1. doi:10.1186/s12964-018-0315-1

160. Kurelac I, Iommarini L, Vatrinet R, et al. Inducing cancer indolence by targeting mitochondrial Complex I is potentiated by blocking macrophage-mediated adaptive responses. *Nat Commun*. 2019;10(1):903. doi:10.1038/s41467-019-08839-1

161. Tseng L-M, Yin P-H, Chi C-W, et al. Mitochondrial DNA mutations and mitochondrial DNA depletion in breast cancer. *Genes, Chromosom Cancer*. 2006;45(7):629-638. doi:10.1002/gcc.20326

162. van Osch FHM, Voets AM, Schouten LJ, et al. Mitochondrial DNA copy number in colorectal cancer: between tissue comparisons, clinicopathological characteristics and

survival. *Carcinogenesis*. 2015;36(12):bgv151. doi:10.1093/carcin/bgv151

163. Barbosa IA, Machado NG, Skildum AJ, Scott PM, Oliveira PJ. Mitochondrial remodeling in cancer metabolism and survival: Potential for new therapies. *Biochim Biophys Acta - Rev Cancer*. 2012;1826(1):238-254. doi:10.1016/J.BBCAN.2012.04.005

164. Carew JS, Huang P. Mitochondrial defects in cancer. *Mol Cancer*. 2002;1(1):9. doi:10.1186/1476-4598-1-9

165. Gasparre G, Iommarini L, Porcelli AM, et al. An inherited mitochondrial DNA disruptive mutation shifts to homoplasmy in oncocytic tumor cells. *Hum Mutat*. 2009;30(3):391-396. doi:10.1002/humu.20870

166. Yin Z, Yang L, Wu F, et al. Reactive Oxygen Species-Mediated Cezanne Inactivation by Oxidation of its Catalytic Cysteine Residue in Hepatocellular Carcinoma. *Oncol Res Featur Preclin Clin Cancer Ther*. 2019;27(9):1069-1077. doi:10.3727/096504019X15566157027506

167. Kumari S, Badana AK, G MM, G S, Malla R. Reactive Oxygen Species: A Key Constituent in Cancer Survival. *Biomark Insights*. 2018;13:1177271918755391. doi:10.1177/1177271918755391

168. Overturf K, al-Dhalimy M, Ou CN, Finegold M, Grompe M. Serial transplantation reveals the stem-cell-like regenerative potential of adult mouse hepatocytes. *Am J Pathol*. 1997;151(5):1273-1280. http://www.ncbi.nlm.nih.gov/pubmed/9358753. Accessed October 28, 2019.

169. Alison MR, Choong C, Lim S. Application of liver stem cells for cell therapy. *Semin Cell Dev Biol*. 2007;18(6):819-826. doi:10.1016/J.SEMCDB.2007.09.016

170. Fellous TG, Islam S, Tadrous PJ, et al. Locating the stem cell niche and tracing hepatocyte lineages in human liver. *Hepatology*. 2009;49(5):1655-1663. doi:10.1002/hep.22791

171. Taylor RW, Barron MJ, Borthwick GM, et al. Mitochondrial DNA mutations in human colonic crypt stem cells. *J Clin Invest*. 2003;112(9):1351-1360. doi:10.1172/JCI19435

172. Parsons BL. Multiclonal tumor origin: Evidence and implications. *Mutat Res Mutat Res*. 2018;777:1-18. doi:10.1016/J.MRREV.2018.05.001

173. Liu J, Dang H, Wang XW. The significance of intertumor and intratumor heterogeneity in liver cancer. *Exp Mol Med*. 2018;50(1):e416-e416. doi:10.1038/emm.2017.165

Annex 1

Patient ID	Sample	Type	mtDNA Coverage	Mean read depth	Best predicted haplogroup
LIHC_1	LIHC_1_T	Primary Tumor	100.00%	84.82	F1a1a
	LIHC_1_B	Blood Derived Normal	100.00%	63.35	F1a1a
LIHC_2	LIHC_2_T	Primary Tumor	100.00%	429.51	H6c
	LIHC_2_NST	Solid Tissue Normal	100.00%	139.8	H6c
LIHC_3	LIHC_3_T	Primary Tumor	100.00%	51.84	A7
	LIHC_3_B	Blood Derived Normal	100.00%	34.79	A7
LIHC_4	LIHC_4_T	Primary Tumor	100.00%	55.55	G1a1a
	LIHC_4_B	Blood Derived Normal	99.00%	31.32	G1a1a
LIHC_5	LIHC_5_T	Primary Tumor	100.00%	104.55	B4b'd'e'j
	LIHC_5_B	Blood Derived Normal	100.00%	38.71	B4b'd'e'j
LIHC_6	LIHC_6_T	Primary Tumor	100.00%	59.29	F1a1b
	LIHC_6_B	Blood Derived Normal	100.00%	37.85	F1a1b
LIHC_7	LIHC_7_T	Primary Tumor	100.00%	61.82	F1b1a1a2
	LIHC_7_B	Blood Derived Normal	100.00%	40.54	F1b1a1a2
LIHC_8	LIHC_8_T	Primary Tumor	100.00%	44.58	D4j3
	LIHC_8_B	Blood Derived Normal	100.00%	43.28	D4j3
LIHC_9	LIHC_9_T	Primary Tumor	100.00%	51.58	B4a1b1a
	LIHC_9_B	Blood Derived Normal	100.00%	35.03	B4a1b1a
LIHC_10	LIHC_10_T	Primary Tumor	100.00%	85.47	B4b1b
	LIHC_10_B	Blood Derived Normal	100.00%	44.39	B4b1b
LIHC_11	LIHC_11_T	Primary Tumor	100.00%	63.21	A5b
	LIHC_11_B	Blood Derived Normal	100.00%	37.96	A5b
LIHC_12	LIHC_12_T	Primary Tumor	100.00%	56.94	M9a

	LIHC_12_B	Blood Derived Normal	99.00%	31.09	M9a
LIHC_13	LIHC_13_T	Primary Tumor	100.00%	40.05	C4b
	LIHC_13_B	Blood Derived Normal	100.00%	32.42	C4b
LIHC_14	LIHC_14_T	Primary Tumor	100.00%	52.72	D4a1h
	LIHC_14_B	Blood Derived Normal	99.00%	32.68	D4a1h
LIHC_15	LIHC_15_T	Primary Tumor	100.00%	47.09	F1c
	LIHC_15_B	Blood Derived Normal	100.00%	36.4	F1c
LIHC_16	LIHC_16_T	Primary Tumor	100.00%	78.59	B4b1b
	LIHC_16_B	Blood Derived Normal	99.00%	37.57	B4b1b
LIHC_17	LIHC_17_T	Primary Tumor	100.00%	51.74	B4c1b2c2
	LIHC_17_B	Blood Derived Normal	99.00%	42.16	B4c1b2c2
LIHC_18	LIHC_18_T	Primary Tumor	100.00%	53.15	B4b1a2a
	LIHC_18_B	Blood Derived Normal	99.00%	34.82	B4b1a2a
LIHC_19	LIHC_19_T	Primary Tumor	100.00%	135.97	M7c1a
	LIHC_19_B	Blood Derived Normal	100.00%	45.24	M7c1a
LIHC_20	LIHC_20_T	Primary Tumor	100.00%	42	B4g2
	LIHC_20_B	Blood Derived Normal	100.00%	66.1	B4g2
LIHC_21	LIHC_21_T	Primary Tumor	100.00%	43.11	M12a1b
	LIHC_21_B	Blood Derived Normal	100.00%	40.94	M12a1b
LIHC_22	LIHC_22_T	Primary Tumor	100.00%	82.63	C4a1a_195
	LIHC_22_B	Blood Derived Normal	100.00%	38.43	C4a1a_195
LIHC_23	LIHC_23_T	Primary Tumor	100.00%	167.17	D4j
	LIHC_23_B	Blood Derived Normal	100.00%	48.14	D4j
LIHC_24	LIHC_24_T	Primary Tumor	100.00%	42.64	N9a9
	LIHC_24_B	Blood Derived Normal	100.00%	41.13	N9a9
LIHC_25	LIHC_25_T	Primary Tumor	100.00%	30.57	Z4a1a1
	LIHC_25_B	Blood Derived Normal	100.00%	31.4	Z4a1a1

LIHC_26	LIHC_26_T	Primary Tumor	100.00%	43.99	D6a1
	LIHC_26_B	Blood Derived Normal	99.00%	34.48	D6a1
LIHC_27	LIHC_27_T	Primary Tumor	100.00%	125.09	B4a1c3a
	LIHC_27_B	Blood Derived Normal	100.00%	47.98	B4a1c3a
LIHC_28	LIHC_28_T	Primary Tumor	100.00%	39.36	D4b2b1
	LIHC_28_B	Blood Derived Normal	100.00%	30.02	D4b2b1
LIHC_29	LIHC_29_T	Primary Tumor	100.00%	75.52	Z4a1a1
	LIHC_29_B	Blood Derived Normal	100.00%	33.46	Z4a1a1
LIHC_30	LIHC_30_T	Primary Tumor	100.00%	36.13	D4f1
	LIHC_30_B	Blood Derived Normal	100.00%	46.54	D4f1
LIHC_31	LIHC_31_T	Primary Tumor	100.00%	67.66	D4a2
	LIHC_31_B	Blood Derived Normal	99.00%	30.57	D4a2
LIHC_32	LIHC_32_T	Primary Tumor	100.00%	157.63	B5a2a1_16129
	LIHC_32_B	Blood Derived Normal	100.00%	32.31	B5a2a1_16129
LIHC_33	LIHC_33_T	Primary Tumor	100.00%	138.94	B5b3a
	LIHC_33_B	Blood Derived Normal	100.00%	39.36	B5b3a
LIHC_34	LIHC_34_T	Primary Tumor	100.00%	59.44	A5a
	LIHC_34_B	Blood Derived Normal	100.00%	30.42	A5a
LIHC_35	LIHC_35_T	Primary Tumor	100.00%	332.04	M74a
	LIHC_35_B	Blood Derived Normal	99.00%	40.07	M74a
LIHC_36	LIHC_36_T	Primary Tumor	100.00%	45.57	B4m
	LIHC_36_B	Blood Derived Normal	100.00%	40.53	B4m
LIHC_37	LIHC_37_T	Primary Tumor	100.00%	46.26	N10b
	LIHC_37_B	Blood Derived Normal	100.00%	40.58	N10b
LIHC_38	LIHC_38_T	Primary Tumor	100.00%	180.33	L3e1e
	LIHC_38_NST	Solid Tissue Normal	100.00%	291.85	L3e1e
LIHC-39	LIHC_39_T	Primary Tumor	100.00%	269.01	H13a1a1a

	LIHC_39_B	Blood Derived Normal	100.00%	92.3	H13a1a1a
	LIHC_39_NST	Solid Tissue Normal	100.00%	342.23	H13a1a1a
LIHC_40	LIHC_40_T	Primary Tumor	100.00%	193.68	J1c3
	LIHC_40_NST	Solid Tissue Normal	100.00%	438.9	J1c3
LIHC_41	LIHC_41_T	Primary Tumor	99.00%	35.7	U5b1d1b
	LIHC_41_NST	Solid Tissue Normal	100.00%	54.67	U5b1d1b
LIHC_42	LIHC_42_T	Primary Tumor	100.00%	50.87	B5a1b1
	LIHC_42_B	Blood Derived Normal	100.00%	43.62	B5a1b1
LIHC_43	LIHC_43_T	Primary Tumor	100.00%	319.4	K1a_150
	LIHC_43_B	Blood Derived Normal	100.00%	72.47	K1a_150
	LIHC_43_NST	Solid Tissue Normal	100.00%	642.2	K1a_150
LIHC_44	LIHC_44_T	Primary Tumor	100.00%	169.38	L3d1a1
	LIHC_44_B	Blood Derived Normal	100.00%	99.97	L3d1a1
LIHC_45	LIHC_45_T	Primary Tumor	100.00%	201.82	J1c1b1a1
	LIHC_45_B	Blood Derived Normal	100.00%	59.09	J1c1b1a1
	LIHC_45_NST	Solid Tissue Normal	100.00%	336.56	J1c1b1a1
LIHC_46	LIHC_46_T	Primary Tumor	100.00%	183.41	B4a1c
	LIHC_46_NST	Solid Tissue Normal	100.00%	526	B4a1c
LIHC_47	LIHC_47_T	Primary Tumor	100.00%	85.38	J2b1a4
	LIHC_47_NST	Solid Tissue Normal	100.00%	381.78	J2b1a4
LIHC_48	LIHC_48_T	Primary Tumor	100.00%	254.18	H1af2
	LIHC_48_B	Blood Derived Normal	100.00%	168.05	H1af2
LIHC_49	LIHC_49_T	Primary Tumor	100.00%	61.47	N9a8
	LIHC_49_NST	Solid Tissue Normal	100.00%	47.66	N9a8
LIHC_50	LIHC_50_T	Primary Tumor	99.61%	43.08	L2a1e1
	LIHC_50_B	Blood Derived Normal	99.54%	35.69	L2a1e1
LIHC_51	LIHC_51_T	Primary Tumor	100.00%	69.49	H1at1

	LIHC_51_B	Blood Derived Normal	100.00%	41.38	H1at1
LIHC_52	LIHC_52_T	Primary Tumor	99.83%	52.78	H2a2a1
	LIHC_52_B	Blood Derived Normal	99.35%	33.97	H2a2a1
LIHC_53	LIHC_53_T	Primary Tumor	100.00%	108.47	H5a1g1a
	LIHC_53_B	Blood Derived Normal	100.00%	39.88	H5a1g1a
LIHC_54	LIHC_54_T	Primary Tumor	100.00%	33.28	HV16
	LIHC_54_B	Blood Derived Normal	100.00%	38.24	HV16
LIHC_55	LIHC_55_T	Primary Tumor	100.00%	97.71	U5a1d2b
	LIHC_55_B	Blood Derived Normal	100.00%	89.89	U5a1d2b
LIHC_56	LIHC_56_T	Primary Tumor	99.90%	82.43	H13a1a1b
	LIHC_56_B	Blood Derived Normal	100.00%	71.37	H13a1a1b
LIHC_57	LIHC_57_T	Primary Tumor	100.00%	83.79	J1c2h
	LIHC_57_B	Blood Derived Normal	99.80%	52.45	J1c2h
LIHC_58	LIHC_58_T	Primary Tumor	99.90%	84.82	T1a1_152
	LIHC_58_B	Blood Derived Normal	100.00%	97.22	T1a1_152
LIHC_59	LIHC_59_T	Primary Tumor	100.00%	126.47	J1b1a1
	LIHC_59_NST	Solid Tissue Normal	100.00%	425.63	J1b1a1
LIHC_60	LIHC_60_T	Primary Tumor	99.90%	62.45	H7d3a
	LIHC_60_NST	Solid Tissue Normal	100.00%	600.57	H7d3a
LIHC_61	LIHC_61_T	Primary Tumor	100.00%	353.12	H1e1a
	LIHC_61_NST	Solid Tissue Normal	100.00%	406.85	H1e1a
LIHC_62	LIHC_62_T	Primary Tumor	100.00%	394.99	K1a4a1f
	LIHC_62_NST	Solid Tissue Normal	100.00%	396.78	K1a4a1f
LIHC_63	LIHC_63_T	Primary Tumor	100.00%	105.54	H6a1b3a
	LIHC_63_NST	Solid Tissue Normal	100.00%	473.06	H6a1b3a
LIHC_64	LIHC_64_T	Primary Tumor	100.00%	111.57	H2a5b2
	LIHC_64_NST	Solid Tissue Normal	99.30%	67.82	H2a5b2

LIHC_65	LIHC_65_T	Primary Tumor	100.00%	204.46	T16223C
	LIHC_65_NST	Solid Tissue Normal	100.00%	111.36	T16223C
LIHC_66	LIHC_66_T	Primary Tumor	100.00%	35.56	K1c2
	LIHC_66_B	Blood Derived Normal	99.80%	51.81	K1c2
LIHC_67	LIHC_67_T	Primary Tumor	99.90%	55.32	N21_195
	LIHC_67_B	Blood Derived Normal	100.00%	70.13	N21_195
LIHC_68	LIHC_68_T	Primary Tumor	100.00%	67.73	J1b1a1a
	LIHC_68_B	Blood Derived Normal	99.80%	42.07	J1b1a1a
LIHC_69	LIHC_69_T	Primary Tumor	99.70%	125.89	I1b
	LIHC_69_NST	Solid Tissue Normal	99.10%	79.26	I1b
LIHC_70	LIHC_70_T	Primary Tumor	99.80%	47.63	U5b3b1
	LIHC_70_B	Blood Derived Normal	99.40%	35.99	U5b3b1
LIHC_71	LIHC_71_T	Primary Tumor	100.00%	120.27	D1c
	LIHC_71_B	Blood Derived Normal	99.70%	60.67	D1c
LIHC_72	LIHC_72_T	Primary Tumor	99.60%	61.61	H2a2b4
	LIHC_72_B	Blood Derived Normal	99.70%	35.93	H2a2b4
LIHC_73	LIHC_73_T	Primary Tumor	99.90%	97.83	A2k1a
	LIHC_73_B	Blood Derived Normal	99.70%	51.96	A2k1a
LIHC_74	LIHC_74_T	Primary Tumor	100.00%	122.89	T2e
	LIHC_74_B	Blood Derived Normal	99.90%	41.63	T2e
LIHC_75	LIHC_75_T	Primary Tumor	100.00%	49.28	V2a1
	LIHC_75_B	Blood Derived Normal	99.10%	30.62	V2a1
LIHC_76	LIHC_76_T	Primary Tumor	99.20%	41.78	R11b
	LIHC_76_B	Blood Derived Normal	99.40%	42.36	R11b
LIHC_77	LIHC_77_T	Primary Tumor	100.00%	81.07	M7c1a3a
	LIHC_77_B	Blood Derived Normal	100.00%	142.09	M7c1a3a
LIHC_78	LIHC_78_T	Primary Tumor	100.00%	404.32	M10a1b

	LIHC_78_B	Blood Derived Normal	100.00%	112.73	M10a1b
LIHC_79	LIHC_79_T	Primary Tumor	99.30%	39.98	D4a
	LIHC_79_B	Blood Derived Normal	99.90%	44.53	D4a
LIHC_80	LIHC_80_T	Primary Tumor	99.70%	46.93	K1c1a
	LIHC_80_B	Blood Derived Normal	99.90%	36.51	K1c1a
LIHC_81	LIHC_81_T	Primary Tumor	100.00%	199.89	L2b1a3
	LIHC_81_B	Blood Derived Normal	99.90%	33.62	L2b1a3
LIHC_82	LIHC_82_T	Primary Tumor	100.00%	82.15	H1b1g
	LIHC_82_B	Blood Derived Normal	99.90%	45.94	H1b1g
LIHC_83	LIHC_83_T	Primary Tumor	100.00%	83.5	H1c1_16093
	LIHC_83_B	Blood Derived Normal	99.90%	40.6	H1c1_16093
LIHC_84	LIHC_84_T	Primary Tumor	99.90%	50.64	H13a1a1
	LIHC_84_B	Blood Derived Normal	100.00%	42.9	H13a1a1
LIHC_85	LIHC_85_T	Primary Tumor	99.90%	86.54	U5b2b3a1a
	LIHC_85_B	Blood Derived Normal	99.80%	45.42	U5b2b3a1a
LIHC_86	LIHC_86_T	Primary Tumor	100.00%	91.01	H3g3
	LIHC_86_B	Blood Derived Normal	99.90%	58.55	H3g3
LIHC_87	LIHC_87_T	Primary Tumor	99.86%	67.15	C1b
	LIHC_87_B	Blood Derived Normal	98.10%	32.34	C1b
LIHC_88	LIHC_88_T	Primary Tumor	99.10%	34.91	C1b
	LIHC_88_B	Blood Derived Normal	99.50%	39.7	C1b
LIHC_89	LIHC_89_T	Primary Tumor	100.00%	77.34	B2
	LIHC_89_B	Blood Derived Normal	100.00%	93.43	B2
LIHC_90	LIHC_90_T	Primary Tumor	99.50%	42.74	D1
	LIHC_90_B	Blood Derived Normal	99.80%	36.37	D1
LIHC_91	LIHC_91_T	Primary Tumor	99.80%	44.55	A2
	LIHC_91_B	Blood Derived Normal	99.50%	46.76	A2

LIHC_92	LIHC_92_T	Primary Tumor	100.00%	87.16	H1b2
	LIHC_92_B	Blood Derived Normal	99.80%	41.19	H1b2
LIHC_93	LIHC_93_T	Primary Tumor	99.20%	34.96	B4c1b2a
	LIHC_93_B	Blood Derived Normal	99.90%	40.52	B4c1b2a
LIHC_94	LIHC_94_T	Primary Tumor	100.00%	82.15	L3e3b1
	LIHC_94_B	Blood Derived Normal	100.00%	100.01	L3e3b1
LIHC_95	LIHC_95_T	Primary Tumor	100.00%	69.29	U5a1a2b
	LIHC_95_B	Blood Derived Normal	99.90%	49.86	U5a1a2b
LIHC_96	LIHC_96_T	Primary Tumor	99.80%	49.47	K1a4
	LIHC_96_B	Blood Derived Normal	100.00%	63.55	K1a4
LIHC_97	LIHC_97_T	Primary Tumor	100.00%	123.83	K1a4a1h
	LIHC_97_B	Blood Derived Normal	99.90%	72.09	K1a4a1h
LIHC_98	LIHC_98_T	Primary Tumor	100.00%	110.72	W1c
	LIHC_98_B	Blood Derived Normal	99.30%	40.72	W1c
LIHC_99	LIHC_99_T	Primary Tumor	99.80%	45.42	J1c3b
	LIHC_99_B	Blood Derived Normal	99.50%	36.05	J1c3b
LIHC_100	LIHC_100_T	Primary Tumor	99.70%	35.45	T2b3_151
	LIHC_100_B	Blood Derived Normal	99.80%	40.41	T2b3_151
LIHC_101	LIHC_101_T	Primary Tumor	100.00%	140.37	T2a1a
	LIHC_101_B	Blood Derived Normal	99.90%	41.22	T2a1a
LIHC_102	LIHC_102_T	Primary Tumor	99.60%	32.12	L2b1a3
	LIHC_102_B	Blood Derived Normal	99.80%	42.23	L2b1a3
LIHC_103	LIHC_103_T	Primary Tumor	99.90%	67.72	K1d
	LIHC_103_B	Blood Derived Normal	99.90%	35.89	K1d
LIHC_104	LIHC_104_T	Primary Tumor	100.00%	124.58	H5a1
	LIHC_104_B	Blood Derived Normal	99.50%	34.65	H5a1

T: tumor; B: blood; NST: normal solid tissue