

Genetic biomarkers of sorafenib response in patients with hepatocellular carcinoma

Giuseppa Augello¹, Lydia Giannitrapani², Francesca Di Gaudio³, Francesca Scionti⁴, Domenico Ciliberto⁴, Nicoletta Straropoli⁴, Giuseppe Agapito⁴, Mario Cannataro⁴, Pierfrancesco Tassone⁵, Pierosandro Tagliaferri⁵, Aurelio Seidita³, Maurizio Soresi³, Marco Affronti⁶, Gaetano Bertino⁷, Maurizio Russello⁸, Rosaria Ciriminna⁹, Claudia Lino⁹, Francesca Spinnato¹⁰, Francesca Verderame¹⁰, Melchiorre Cervello¹, and Mariamena Arbitrio¹

¹Institute for Biomedical Research and Innovation National Research Council

²Affiliation not available

³University of Palermo

⁴Università degli Studi Magna Graecia di Catanzaro

⁵Magna Graecia University of Catanzaro

⁶Policlinico di Palermo

⁷Azienda Ospedaliero Universitaria Policlinico Vittorio Emanuele Catania

⁸Azienda Ospedaliera ARNAS Garibaldi

⁹CNR

¹⁰Villa Sofia Cervello United Hospitals

July 5, 2023

Abstract

Purpose: The identification of biomarkers for predicting inter-individual sorafenib response variability could allow hepatocellular carcinoma (HCC) patients stratification. SNPs in angiogenesis- and drug absorption, distribution, metabolism, and excretion (ADME)-related genes were evaluated to identify new potential predictive biomarkers of sorafenib response in HCC patients. **Methods:** Five known SNPs in angiogenesis-related genes, including VEGF-A, VEGF-C, HIF-1a, ANGPT2 and NOS3, were investigated in 34 HCC patients (9 sorafenib responders and 25 non-responders). A subgroup of 23 patients was genotyped for SNPs in ADME genes. A machine learning classifier method was used to discover classification rules for our dataset. **Results:** We found that only VEGF-A (rs2010963) C allele and CC genotype were significantly associated with sorafenib response. ADME-related gene analysis identified 10 polymorphic variants in ADH1A (rs6811453), ADH6 (rs10008281), SULT1A2/CCDC101 (rs11401), CYP26A1 (rs7905939), DPYD (rs2297595 and rs1801265), FMO2 (rs2020863) and SLC22A14 (rs149738, rs171248 and rs183574) significantly associated with sorafenib response. We have identified a genetic signature of predictive response which could permit non-responder/responder patient stratification. Angiogenesis- and ADME-related genes correlation was confirmed by cumulative genetic risk score and network and pathway enrichment analysis. **Conclusions:** Our findings provide a proof of concept that need further validation in follow-up studies for HCC patient stratification for sorafenib prescription.

Introduction

Hepatocellular carcinoma (HCC) accounts for about 90% of liver cancers. Most patients with HCC are diagnosed at an advanced tumor stage when treatment options are very limited. Until 2018, when lenvatinib was approved, sorafenib was considered the gold standard in the first-line setting for the treatment of

advanced HCC.¹ However, choosing between these two agents remains challenging due to their low impact on survival and their similar and well-tolerated safety profiles.² Recently, the atezolizumab-bevacizumab combination has emerged as the first-line systemic treatment, but sorafenib is still a relevant choice for refractory patients and those ineligible for immunotherapy.³ Sorafenib exerts its action through the inhibition of tumor cell proliferation and angiogenesis via the targeting of several oncogenic signaling pathways involving serine/threonine and tyrosine kinases (RAF1, BRAF, VEGFR 1, 2, 3, PDGFR, KIT, FLT3, FGFR1, and RET).^{4,5} However, sorafenib resistance remains a major challenge for improving the effectiveness of HCC treatment. The underlying mechanisms for inter-individual variability in response to therapy have not been fully elucidated, and no validated markers have been found that are capable of predicting clinical outcomes or sorafenib tolerability.^{6,7} Thus, the identification of suitable biomarkers for patient stratification for sorafenib response in HCC may potentially help physicians in guiding the selection of tailored treatments.

HCC is a hypervascular tumor in which angiogenesis plays an important role for tumor growth and progression. Among others, VEGF/VEGFR, angiopoietin (ANGPT), endothelial nitric oxide synthase (eNOS or NOS3), and hypoxia-inducible factor-1 α (HIF-1 α) signaling play an important role in regulating tumor angiogenesis.⁸ Single nucleotide polymorphisms (SNPs) in angiogenesis-related genes have been reported to influence outcomes in HCC patients treated with sorafenib.⁹⁻¹¹

In the present study, we selected 5 SNPs in these angiogenesis-related genes for the genotyping of 34 HCC patients, of which 9 showed response (responders) to therapy and 25 no response (non-responders). Additionally, in a subgroup of HCC patients, we evaluated 1,931 SNPs and 5 copy number variations in 231 genes involved in drug absorption, distribution, metabolism, and excretion (ADME) using the DMET Plus microarray assay for the identification of new potential predictive biomarkers of response and outcome.¹²⁻¹⁴ Through a model learning process, we proceeded to apply rules to classify all patients in terms of the detected SNPs and genotypes and according to sorafenib response in order to identify a predictive genetic signature which could allow the stratification of non-responder/responder patients to sorafenib for tailored prescriptions. Furthermore, the correlation between angiogenesis- and ADME-related genes was confirmed by a cumulative genetic risk score (GRS) and by network and pathway enrichment analysis, which demonstrated the association of 8/12 identified genes placed in topological key points of the interaction networks involved in several key common biological pathways correlated to HCC and sorafenib. Our findings should be considered as a “proof of concept” to be further validated in follow-up studies for the stratification of HCC patients towards the improvement of therapeutic choices.

Patients and Methods

Study participants

Thirty-four patients with advanced HCC who had undergone sorafenib treatment were enrolled in our study and distributed into two groups (responders vs non-responders) according to modified Response Evaluation Criteria in Solid Tumors (mRECIST).¹⁵ Among them, patients who showed partial response or stable disease (SD) for more than 6 months were classified as responders, while patients with a progressive disease were considered non-responders. Only two patients showed a complete response. We classified patients from 5 medical centers: the Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties, University Hospital of Palermo, Italy; the Oncology Unit, AOU Mater Domini, Catanzaro, Italy; the Department of Clinical and Experimental Medicine Policlinico “G. Rodolico” University of Catania, Italy; the Liver Unit of ARNAS Garibaldi-Nesima, Catania, Italy; and the Villa Sofia-Cervello Hospital, Palermo, Italy. All patients were enrolled according to inclusion criteria provided by the clinical study protocol. More in detail, since the regulatory approval of the use of sorafenib as a drug for the treatment of HCC is not amenable to locoregional therapy, all the patients who met the inclusion criteria according to international guidelines (AASLD and EASL)^{16,17} for medical therapy and had therefore started therapy with sorafenib, were included in the study after signing the informed consent.

The study was approved by the Ethics Committees of the University Hospital of Palermo, and the Ethics Committees of Section Area Centro (Region of Calabria), as spontaneous study No. 3/2017 and Prot. n. 387,

respectively. All patients gave their approval and signed informed consent according to the recommendations of the Declaration of Helsinki for biomedical research involving human subjects.¹⁸

Sample collection and genotyping

For each subject, a 5 ml peripheral blood sample was collected in EDTA anticoagulant tubes. The sample was centrifuged, and then the plasma was replaced with an equal amount of physiological solution (0.9% sodium chloride) and stored at -80°C. Genomic DNA was extracted from whole blood using QIAamp DNA Blood Mini Kit (Qiagen, Valencia, CA), in accordance with the manufacturer's instructions. For the evaluation of selected SNPs in angiogenesis-related genes, DNA samples of 34 HCC patients were genotyped using the TaqMan allelic discrimination method (StepOne Plus Real-Time PCR System, Applied Biosystems, Monza, Italy) using commercial (VEGF-A rs2010963, VEGF-C rs4604006, HIF-1A rs12434438, ANGPT2 rs55633437, NOS3 rs2070744) genotyping assays (ThermoFisher Scientific, Waltham, MA, USA). DMET Plus assay (ThermoFisher Scientific, Waltham, MA, USA.) was performed as previously described.¹⁹⁻²¹

Genetic risk score

To build a GRS for sorafenib, we estimated the genetic profiles of patients depending on frequencies of genotypes and alleles of examined gene polymorphisms. For each SNP, a score of 0 was defined for homozygous non-responder alleles, 1 for heterozygous responder and non-responder alleles, and 2 for two homozygous responder alleles.²²⁻²⁴ The response-increasing alleles were attributed based on their greater frequency in response subjects according to the literature data for angiogenesis-related genes^{9-11,25} and by data obtained in the present study for ADME-related genes. For each subject, a combined GRS was calculated as the sum of these response-increasing alleles; the score was 0-10 for 5 variants.

Statistical analysis

Normally distributed continuous variables were expressed as mean $\pm$ SD, and the differences between the two groups were assessed with the Mann Whitney U test.

DMET analysis of frequencies was calculated using Fisher's exact test and Bonferroni's correction for multiple comparisons using the DMET-Analyzer tool.²⁶ In all analyses, a p-value < 0.05 was considered statistically significant. ADME genotyping calls from the intensity array were performed with DMET Console software (version 1.1, ThermoFisher Scientific, Waltham, MA, USA). All alleles were tested for Hardy-Weinberg equilibrium. Genotypes with a call rate $< 96\%$, a high rate of "possible rare allele"/"no call," and a concordance of 100% were excluded from further analysis.

Classification rules

To discover classification rules for our dataset, starting from angiogenesis-related SNPs and identified SNPs in ADME genes (sorafenib dataset), we applied a classifier for the identification of rules which could facilitate the stratification of responder/non-responder patients in terms of sorafenib response. Classifiers belong to the non-parametric supervised learning algorithm category, where the machines learn patterns buried in the data using previously labeled training data which is mandatory for supervised learning. Supervised learning aims to produce a model that predicts class variable values by learning simple patterns inferred from the data features. The RandomTree's output is a classification tree easily understandable even by non-domain experts which can be quickly translated into classification rules in the "IF (cond1 & cond2 & ... & condn) THEN class" format.

Network and pathway enrichment analysis

Network analysis was used to determine the influential role of the identified seed genes within the population affected by HCC and treated with sorafenib.

To consolidate the seed genes, we conducted a gene network consolidation approach like the one proposed in Agapito et al.²⁷ The consolidation approach consists of: i) mapping seed genes on the human protein-protein interaction (PPI) network from the Integrated Interactions Database (IID);²⁸ ii) for each mapped seed gene,

the expansion process regards the computing of the community with a radius equal to one, allowing the collection of all genes that share similar functions or involvement in similar biological processes with seed genes.²⁹ We used Cytoscape³⁰ and the Cytoscape plugin CytoHubba³¹ to assess the mapped seed genes' roles in the "seedGeneNetwork," computing the degree of the seed genes. The node degree is the sum of all edges connected to it. A node with a degree equal to d indicates that the neighbor radius nodes linking it is d . To further investigate the seed genes list, we computed pathway enrichment analysis to gain insight into the affected unknown underlying biological mechanisms associated with sorafenib response in the 34 HCC patients.

Reactome enrichment analyses were performed using BiP and pathDIP.^{32,33}

Results

Patient demographic and clinical characteristics

Supporting Information Table S1 shows some demographic and clinical characteristics in the patients studied. The median age was 73 $\pm$ 6 years (range 57-89 years). Men were more prevalent with respect to women (24 males and 10 females). Hepatitis virus C (HCV) infection was present in 19 cases (56%), HBV infection in 6 (18%), and alcoholic- and cryptogenetic/metabolic-related HCC was present in 3 (9%) and 6 (18%) cases, respectively. According to the Barcelona Clinic Liver Cancer (BCLC) classification, 3 patients were BCLC stage A (9%), 21 patients were BCLC stage B (62%) and 10 patients were BCLC stage C (29%). As for the Child-Pugh score, 26 patients (76%) were class A, 6 (18%) were class B, and 2 (6%) were class C. Alpha-fetoprotein (AFP) was $\geq$ 400 ng/dL in 4 cases (12%) and $<$ 400 ng/dL in 30 cases (88%). Twenty-two patients (65%) received locoregional treatments. Six patients (18%) had portal invasion. Finally, 19 (56%) patients maintained their starting dose of sorafenib, whereas 15 (44%) underwent dose reduction.

Allele and genotype distributions

For the study of angiogenesis-related genes, 34 patients (9 responders and 25 non-responders) were genotyped using the TaqMan allelic discrimination method. The distribution of rs2010963, rs4604006, rs12434438, rs55633437, and rs2070744 alleles and genotypes in responder (R) and non-responder (NR) groups is shown in Table 1. Only the C allele ($p = 0.004$) (Table 1), and CC genotype ($p = 0.046$) of *VEGF-A* (rs2010963) were significantly associated with sorafenib response.

The genotyping by DMETTM SNP panel allowed the identification, among 1,936 markers, of 10 SNPs in seven genes as significantly associated with sorafenib response: *ADH1A* (alcohol dehydrogenase 1A), *ADH6* (alcohol dehydrogenase 6), *SULT1A2/CCDC101* (Sulfotransferase Family 1A Member 2)/*CCDC101* (Coiled-Coil Domain-Containing Protein 101), *CYP26A1* (Cytochrome P450 Family 26 Subfamily A Member 1), *DPYD* (Dihydropyrimidine dehydrogenase), *FMO2* (Flavin Containing Dimethylaniline Monooxygenase 2), and *SLC22A14* (Solute Carrier Family 22 Member 14) (Table 2).

The heterozygous genotypes of *SLC22A14* rs171248, rs149738, and rs183574 ($p = 0.007$), and the homozygous genotypes AA in *SULT1A2/CCDC101* rs11401 ($p = 0.019$), *DPYD* rs2297595 ($p = 0.011$), and *FMO2* rs2020863 ($p = 0.004$), and TT in *DPYD*9*rs1801265 ($p = 0.0257$), as well as the GG in *ADH6* rs10008281 ($p = 0.019$) and *CYP26A1* rs7905939 ($p = 0.027$), showed a significant association to a lack of response to sorafenib. Instead, a significant correlation to sorafenib efficacy was found for the CC genotype in *ADH1A* rs6811453 ($p = 0.005$) and the heterozygous genotypes AG in *SULT1A2/CCDC101* rs11401 ($p = 0.019$), *DPYD* rs2297595 ($p = 0.011$), and *FMO2* rs2020863 ($p = 0.004$), and CT in *DPYD*9*rs1801265 ($p = 0.026$). The rs11401 is located on 16p11.2 region which contains a splice site encompassing the *SULT1A2* gene (500B Downstream Variant), and the *CCDC101* gene. In addition, the homozygous genotypes of *SLC22A14* rs171248 (TT), rs149738 (AA), and rs183574 (AA) were found to be significantly associated to responder patients ($p = 0.001$).

SNPs and classification rules related to sorafenib response

Figure 1 shows the classification tree computed from the RandomTree's classifier using the sorafenib dataset.

Transforming the classification tree into classification rules (1-13) obtained by analyzing the input genotype dataset, as shown in Table 3, makes it more straightforward to analyze and understand the meaning of the multiple relations between the SNPs and genotypes responsible for a particular phenotype of sorafenib response.

We identified ten classification rules by which to discriminate patients belonging to the non-responder setting, and three rules for the responder ones, with an accuracy of 69.5652%: a subject could satisfy a rule only if a correspondence existed between their own genotype and a detected SNPs, against every couple of alleles within of a rule. For instance, to verify whether a subject matched, i.e., rule 6 in Table 3, it was necessary that the SNPs (rs171248, rs6811453, rs2010963, rs12434438) assessed in the subject presented as detected genotypes (TT, CT, CC, GG), respectively. Thus, only the subjects that matched all the genotypes within a rule could be classified as “non-responder” according to the matching rule.

Afterwards, we examined the cumulative effects of SNPs obtained from the classification tree, developing a GRS by summing the number of response alleles.²²⁻²⁴ The response-increasing alleles were attributed based on their greater frequency in response subjects according to the literature data for angiogenesis-related genes^{8,10,23} and data obtained in the present study for ADME-related genes. The rs7905939 SNP was excluded from the analysis since a clear response allele was not identified. For each SNP, a score of 0 was defined for homozygous non-response alleles, 1 for heterozygous response and non-response alleles, and 2 for two homozygous response alleles. A higher mean GRS score was significantly associated with responders compared to non-responders, when the sum of the 5 scores for the rs2010963, rs4604006, rs12434438, rs183574, and rs6811453 variants was considered for each patient ($p = 0.008$) (Supporting Information Table S2). The mean of the gene count score was 6.00 ± 0.81 in the responder group, and 4.37 ± 1.36 in the non-responder group.

To explore whether the expression of angiogenesis- and ADME-related genes identified in the decision tree (i.e., SLC22A4, ADH1A, VEGF-A, VEGF-C, HIF-1 α , and CY26A1) might have a role in HCC disease outcome in terms of response to sorafenib, we carried out a bioinformatic analysis of these genes using the public dataset GSE109211, downloaded from GEO, in which data from a subset of HCC patients ($n = 67$) treated with sorafenib are reported. As shown in Figure 2, VEGF-A, HIF-1 α , and ADH1A expression were significantly lower in HCC tissues from sorafenib-responsive patients ($n = 20$), whereas SLC22A14 expression was significantly higher. No significant correlation was found between the expression of VEGF-C and CY26A1 genes and sorafenib response.

Pathway enrichment results

The network analysis highlighted a total of 14 genes with a relevant node degree score. Figure 3 shows the “seedGeneNetwork” with the key interactions among the 14 seed genes and other genes involved in several canonical pathway.

The seed genes enriching multiple pathways highlighted the relationships between seed genes and the affected biological functions involved in HCC and sorafenib. Among all genes, *ADH1A* and *CYP26A1*, along with *VEGF-A* and *VEGF-C*, showed a common involvement in the signal transduction pathways, which are reported as dysregulated in HCC, leading to uncontrolled cell division and metastasis.³⁴

Table 3 reports the top 21 degree-ranked seed genes among which, with a high degree value, are included 8/12 genes identified in the study (in bold). Analyzing Table 4, it is worth noting that a seed gene’s degree is higher than the added genes and refers to the centrality of a node in the network. The higher centrality of the degree reveals the crucial roles of genes in the network. Figure 4 shows the top ten degree-ranked seed genes computed using CytoHubba.

Discussion

HCC patients treated with sorafenib show a highly variable response, and patients experience resistance and adverse events in approximately 30% of cases. The molecular mechanisms underlying inter-individual

variability in sorafenib response have yet to be fully elucidated, and a deeper knowledge of the underlying mechanisms and associated gene variants would allow us to tailor treatment prescriptions for better outcomes.

Thus, with this aim, we studied 5 known candidate SNPs in genes controlling tumor angiogenesis, *VEGF-A* (rs2010963), *VEGF-C* (rs4604006), *HIF-1α* (rs12434438), *ANGPT2* (rs55633437), and *NOS3* (rs2070744), and genotyped a subgroup of 23 HCC patients using the DMET plus platform to identify potential prognostic biomarkers correlated to HCC patient responses to sorafenib treatment. Moreover, in a subgroup of patients ($n = 11$), the serum/plasma concentration of sorafenib was determined. The average steady-state sorafenib concentration was 4.13 mg/L in responders, and 4.41 mg/L in non-responders, independent of sorafenib dosages. No significant correlation was detected between sorafenib concentration and clinical outcome (data not shown), as already reported in the literature.³⁵

In our study, the analysis showed that the allele and genotype frequencies of SNPs in angiogenesis-related genes were significantly correlated with the response to sorafenib only for the rs2010963 C allele ($p = 0.004$) and the CC genotype ($p = 0.046$) of the *VEGF-A* gene, in accordance with the results of the retrospective multicenter study ALICE-2, in which the rs2010963 C allele and CC/CG genotype were significantly associated with a higher median overall survival of HCC patients receiving sorafenib.¹⁰ It is likely that *VEGF-A*-related genetic variants could influence the level of circulating VEGF,^{36,37} therefore affecting sorafenib response. Also, in the SHARP trial,³⁸ it was reported that a low VEGF-A plasma baseline level, as a prognostic independent factor, can predict outcomes in patients with advanced HCC, both in the entire patient population and in the placebo cohort.³⁷ Consistent with these results, in our GEO analysis, patients expressing lower levels of VEGF-A mRNA showed a better response to sorafenib therapy (GSE109211 dataset). Our findings confirmed the prominent role of the rs2010963 gene variant and *VEGF-A* expression as significant predictive factors for sorafenib response in HCC patients.^{10,39}

DMET genotyping showed a statistically significant association of the sorafenib “non-responder” phenotype with the heterozygous genotypes of *SLC22A14* rs171248, rs149738, and rs183574, and the homozygous genotypes AA in *SULT1A2/CCDC101* rs11401, *DPYD* rs2297595, *FMO2* rs2020863, the TT in *DPYD*9* rs1801265, the GG in *ADH6* rs10008281, and *CYP26A1* rs7905939. Instead, the sorafenib “responder” phenotype was associated with the genotypes CC in *ADH1A* rs6811453, the AG in *SULT1A2/CCDC101* rs11401, *DPYD* rs2297595, *FMO2* rs2020863, and CT in *DPYD*9* rs1801265, as well as with the homozygous genotypes of *SLC22A14* rs171248 (TT), rs149738 (AA) and rs183574 (AA). These results demonstrate that the ADME genotype is correlated with different responses to sorafenib, underlying the role of the reference allele or variant in the effect on treatment response.

Moreover, to verify whether a correlation of SNPs in angiogenesis- and ADME-related genes might help to discriminate responder/non-responder patients, we applied a classifier to mine classification rules able to figure out the principal signatures for discriminating among patients belonging to the responder/non-responder to sorafenib phenotypes. The novelty of our study lies in the identification of 10 rules in different genotype associations for the identification of the non-responder phenotype, and 3 rules for the responder type. These rules may represent a genetic signature which could allow the stratification of patients who are fit for sorafenib treatment. We found that the genetic signature including 3 ADME-SNPs, *SLC22A14* (rs171248), *ADH1A* (rs6811453) and *CYP26A1* (rs7905939), and 3 known SNPs in angiogenesis-related genes, *VEGF-A* (rs2010963), *VEGF-C* (rs4604006), and *HIF-1A* (rs12434438), was correlated to sorafenib response in our dataset, allowing the discrimination between responders/non-responders according to mRECIST criteria. The signature of response, identified by a decision tree, was also validated by the GRS analysis.

However, while the role of angiogenesis related genes is well known in HCC patients treated with sorafenib, little information is reported in the literature regarding the role of selected ADME genes in this context. *SLC22A14*, also known as organic cation transporter-like 2 (OCTL2), is a gene encoding a member of the organic-cation transporter family and anions (OATs), whose expression is high in the liver. As for the other SLCs, it is involved in regulating the homeostasis of metabolites and the uptake of a wide range of molecules and the disposition of drugs, as well as in promoting cell proliferation, migration, and invasion in HCC.⁴⁰ The function of SLC members in sorafenib resistance is not clear, and only recently have studies begun to

investigate their role in chemoresistance,⁴¹⁻⁴⁴ highlighting the role of aberrant variants or SNPs in organic cation transporters during liver carcinogenesis, with effects on the ability of HCC to respond to sorafenib.

The *ADH1A* gene catalyzes the oxidation of alcohols to aldehydes and belongs to the superfamily of dehydrogenase enzymes. Several reports provide a correlation of *ADH1A* and other *ADH* s expression with increased risk of liver cancer, with an impact on the prognosis for HCC patients.^{20,45}

CYP26A1 , a member of the cytochrome P450 enzyme superfamily, is mainly involved in retinoic acid metabolism and the synthesis of cholesterol, steroids, and other lipids, and it contributes to the development and progression of multiple cancers.⁴⁶⁻⁴⁸ Previous studies in HCC have shown that *CYP26A1* mRNA is downregulated in tumor tissue compared to paired-matched non-tumor tissues,⁴⁹ but the role of *CYP26A1* in HCC is not entirely clear despite being reported as hypovitaminosis. A, as potential result from a *CYP26A1* depletion, could be correlated to higher risk of carcinogenesis.⁵⁰

To support the correlation between ADME- and angiogenesis-related genes, the network and PEA analysis highlighted the association of 8/12 identified genes in topological key points with a relevant node degree score in important pathways underlying biological mechanisms implicated in HCC and sorafenib. In fact, *VEGF-A*, together with *ADH1A* , *CYP26A1* , and *VEGF-C* , showed a common interaction in “signal transduction pathways” which are known to be dysregulated in HCC, with consequent uncontrolled cell division and metastasis,⁴⁴ alteration of intracellular regulators or extracellular signals with abnormal epigenetic modification, and gene expression in the tumor microenvironment. Moreover, all seed genes identified were involved in multiple pathways, showing their involvement in significantly affected by biological functions.

Given that the most recent AASLD guidelines on HCC systemic therapy have approved the use of drugs like bevacizumab+atezolizumab and lenvatinib as first-line therapies, in addition to sorafenib, the opportunity to verify a score like the GRS observed in our responder patients or the genetic signature identified in our retrospective study might help clinicians to select patients with higher chances to benefit from sorafenib treatment. In the presence of a favorable GRS, physicians could treat advanced HCC patients with sorafenib as soon as possible. Conversely, patients with an unfavorable GRS might not be excellent candidates for sorafenib.

Conclusion

The novelty of our study is the identification of genetic signatures through classification rules for SNPs in angiogenesis and ADME genes associated with sorafenib responses in HCC patients, through which it may be possible to personalize prescription. The application of GRS could allow a better stratification of patients. In addition, the network analysis conducted in this study supports the association of 8/12 analyzed genes in topological key points involved in several common biological pathways correlated to HCC and sorafenib. However, our study has some limitations. The sample size was relatively small, and further investigations in a larger sample size may be needed. The opportunity to test the classification rules on predictive biomarkers of response to sorafenib in an independent and larger validation set would give more robustness to our findings. Therefore, our findings had an exploratory aim and are intended as a “proof-of-concept” research to be further validated in a larger dataset to allow sorafenib-tailored prescriptions through predictive biomarkers of response and outcome in HCC patients. The involvement of seed genes in multiple biological pathways related to sorafenib and HCC, as well as the common interactions of *ADH1A* , *CYP26A1* , *VEGF-A* , and *VEGF-C* in signal transduction pathways, should allow future studies on the simultaneous targeting of different signaling pathways or common downstream proteins involved in HCC control and sorafenib response with the aim of personalizing treatment for this still incurable disease.

Acknowledgments

The authors are grateful to Mrs. Antonina Azzolina for the technical support provided.

Conflict of Interest

All authors declare they have no conflicts of interest.

Funding Information

This work was supported by the PSN2014 - Assessorato della Salute - Regione Siciliana prot./Area int.2/n.85890 to M. Cervello., F.D.G., R.C., and L.G.

Data Availability

The data sets analyzed during the current study are available upon request.

Author Contributions

G. Augello, M. Arbitrio, L.G., M.R.E., V.D.S. and M. Cervello designed the research and wrote the manuscript; G. Augello, M. Arbitrio, F. Scionti, D.C., R.C., C.L., F.D.G. and N.S. performed the research; G. Agapito, M. Cervello and G. Augello analyzed the data; L.G., G.B., M.R., F.V., F. Spinnato, M. Affronti, A. Seidita collected samples and clinical data; P. Tagliaferri, P. Tassone, M.S., M. Catanzaro and G.M. critically revised the work. G. Augello, M. Arbitrio and M. Cervello conceived and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

References

1. Cervello M, Emma MR, Augello G, Cusimano A, Giannitrapani L, Soresi M, Akula SM, Abrams SL, Steelman LS, Gulino A, Belmonte B, Montalto G, McCubrey JA. New landscapes and horizons in hepatocellular carcinoma therapy. *Aging (Albany NY)* 2020; 12: 3053-3094. doi: 10.18632/aging.102777
2. Luo J, Gao B, Lin Z, Fan H, Ma W, Yu D, Yang Q, Tian J, Yang X, Li B. Efficacy and safety of lenvatinib versus sorafenib in first-line treatment of advanced hepatocellular carcinoma: A meta-analysis. *Front Oncol.* 2022; 12:1010726. doi: 10.3389/fonc.2022.1010726
3. Finn RS, Qin S, Ikeda M, Galle PR, Ducreux M, Kim TY, Kudo M, Breder V, Merle P, Kaseb AO, Li D, Verret W, Xu DZ, Hernandez S, Liu J, Huang C, Mulla S, Wang Y, Lim HY, Zhu AX, Cheng AL; IMbrave150 Investigators. Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinom. *N Engl J Med* . 2020; 382:1894-1905. doi: 10.1056/NEJMoa1915745
4. Wilhelm S, Carter C, Lynch M, Lowinger T, Dumas J, Smith RA, Schwartz B, Simantov R, Kelley S. Discovery and development of sorafenib: a multikinase inhibitor for treating cancer, *Nat Rev Drug Disc* . 2006; 5:835-44. doi: 10.1038/nrd2130
5. Gong L, Giacomini MM, Giacomini C, Maitland ML, Altman RB, Klein TE. PharmGKB summary: sorafenib pathways, *Pharmacogenet Genomics* . 2017; 27:240-246. doi: 10.1097/FPC.0000000000000279
6. Terada T, Noda S, Inui K. Management of dose variability and side effects for individualized cancer pharmacotherapy with tyrosine kinase inhibitors. *Pharmacol Ther* . 2015; 152:125-134. doi: 10.1016/j.pharmthera.2015.05.009
7. Marisi G, Cucchetti A, Ulivi P, Canale M, Cabibbo G, Solaini L, Foschi FG, De Matteis S, Ercolani G, Valgiusti M, Frassinetti GL, Scartozzi M, Casadei Gardini A. Ten Years of Sorafenib in Hepatocellular Carcinoma: Are there any Predictive and/or Prognostic Markers? *World J Gastroenterol* . 2018; 24:4152-4163. doi: 10.3748/wjg.v24.i36.4152
8. Lugano R, Ramachandran M, Dimberg A. Tumor angiogenesis: causes, consequences, challenges and opportunities. *Cell Mol Life Sci.* 2020; 77:1745-1770. doi: 10.1007/s00018-019-03351-7
9. Casadei Gardini A, Marisi G, Faloppi L, Scarpi E, Foschi FG, Iavarone M, Lauletta G, Corbelli J, Valgiusti M, Facchetti F, Della Corte C, Neri LM, Tamberi S, Cascinu S, Scartozzi M, Amadori D, Nanni O, Tenti E, Ulivi P, Frassinetti GL. ENOS Polymorphisms and Clinical Outcome in Advanced HCC Patients Receiving Sorafenib: Final Results of the ePHAS Study. *Oncotarget* 2016; 7:27988-27999. doi: 10.18632/oncotarget.8569
10. Faloppi L, Puzzone M, Casadei Gardini A, Silvestris N, Masi G, Marisi G, Vivaldi C, Gadaleta CD, Ziranu P, Bianconi M, Loretelli C, Demurtas L, Lai E, Giampieri R, Galizia E, Ulivi P, Battelli N, Falcone A, Cascinu S, Scartozzi M. Angiogenesis Genotyping and Clinical Outcomes in Patients with Advanced Hepatocellular Carcinoma Receiving Sorafenib: The ALICE-2 Study. *Target Oncol* . 2020; 15:115-126. doi: 10.1007/s11523-020-00698-x

11. Casadei-Gardini A, Marisi G, Dadduzio V, Gramantieri L, Faloppi L, Ulivi P, Foschi FG, Tamburini E, Vivaldi C, Rizzato MD, Ielasi L, Canale M, Conti F, Rudnas B, Fornaro L, Silvestris N, Silletta M, Cardellino GG, Lonardi S, Fornari F, Orsi G, Rovesti G, Zagonel V, Cascinu S, Scartozzi M. Association of NOS3 and ANGPT2 Gene Polymorphisms with Survival in Patients with Hepatocellular Carcinoma Receiving Sorafenib: Results of the Multicenter Prospective INNOVATE Study. *Clin Cancer Res* . 2020; 26:4485-4493. doi: 10.1158/1078-0432.CCR-19-3897
12. Arbitrio M, Di Martino MT, Scionti F, Agapito G, Guzzi PH, Cannataro M, Tassone P, Tagliaferri P. DMET (Drug Metabolism Enzymes and Transporters): A pharmacogenomic platform for precision medicine. *Oncotarget* 2016; 8:54028-54050. doi: 10.18632/oncotarget.9927
13. Arbitrio M, Scionti F, Di Martino MT, Caracciolo D, Pensabene L, Tassone P, Tagliaferri P. Pharmacogenomics Biomarker Discovery and Validation for Translation in Clinical. *Practic Clin Transl Sci* . 2021; 14: 113-119. doi: 10.1111/cts.12869
14. Scionti F, Di Martino MT, Caracciolo D, Pensabene L, Tagliaferri P, Arbitrio M. Tools in Pharmacogenomics Biomarker Identification for Cancer Patients. *Methods Mol Biol*. 2022; 2401:1-12. doi: 10.1007/978-1-0716-1839-4_1
15. Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* . 2010; 30:52-60. doi: 10.1055/s-0030-1247132
16. Bruix J, Sherman M. American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma: an update. *Hepatology* . 2011; 53:1020-1022. doi: 10.1002/hep.24199
17. European Association For The Study Of The Liver; European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma, J. Hepatol. 56 (2012) 908-943. doi: 10.1016/j.jhep.2011.12.001
18. Scionti FPL, Di Martino MT, Arbitrio M, Tagliaferri P. Ethical Perspectives on Pharmacogenomic Profiling. Elsevier. Amsterdam, The Netherlands, 2021.
19. Scionti F, Di Martino MT, Sestito S, Nicoletti A, Falvo F, Roppa K, Arbitrio M, Guzzi PH, Agapito G, Pisani A, Riccio E, Concolino D, Pensabene L. Genetic variants associated with Fabry disease progression despite enzyme replacement therapy, *Oncotarget* . 2017; 8:107558-107564. doi: 10.18632/oncotarget.22505
20. Scionti F, Agapito G, Caracciolo D, Riillo C, Grillone K, Cannataro M, Di Martino MT, Tagliaferri P, Tassone P, Arbitrio M. Risk Alleles for Multiple Myeloma Susceptibility in ADME Genes. *Cells* . 2022; 11:189. doi: 10.3390/cells11020189
21. Arbitrio M, Scionti F, Altomare E, Di Martino MT, Agapito G, Galeano T, Staropoli N, Iuliano E, Grillone F, Fabiani F, Caracciolo D, Cannataro M, Arpino G, Santini D, Tassone P, Tagliaferri P. Polymorphic Variants in NR1H3 and UGT2B7 Predict Taxane Neurotoxicity and Have Prognostic Relevance in Patients With Breast Cancer: A Case-Control Study, *Clin Pharmacol Ther*. 2019; 106:422-431. doi: 10.1002/cpt.1391
22. Cargnin S, Viana M, Sances G, Cantello R, Tassorelli C, Terrazzino S. Using a Genetic Risk Score Approach to Predict Headache Response to Triptans in Migraine Without Aura. *J Clin Pharmacol* . 2019; 59:288-294. doi: 10.1002/jcph.1320
23. Du B, Zhang C, Yue L, Ren B, Zhao Q, Li D, He Y, Zhang W. Prediction model for the efficacy of folic acid therapy on hyperhomocysteinaemia based on genetic risk score methods. *Br J Nutr* . 2019; 122:39-46. doi: 10.1017/S0007114519000783
24. Gagno S, D'Andrea MR, Mansutti M, Zanusso C, Puglisi F, Dreussi E, Montico M, Biason P, Cecchin E, Iacono D, Russo S, Cinausero M, Saracchini S, Gasparini G, Sartori D, Bari M, Collovà E, Meo R, Merkabaoui G, Spagnoletti I, Pellegrino A, Gianni L, Sandri P, Cretella E, Vattemi E, Rocca A, Serra P, Fabbri MA, Benedetti G, Foghini L, Medici M, Basso U, Amoroso V, Riccardi F, Baldelli AM, Clerico M, Bonura S, Saggia C, Innocenti F, Toffoli G. A New Genetic Risk Score to Predict the Outcome of Locally Advanced or Metastatic Breast Cancer Patients Treated With First-Line Exemestane: Results From a Prospective Study. *Clin Breast Cancer* . 2019; 19:137-145. doi: 10.1016/j.clbc.2018.11.009
25. Faloppi L, Casadei Gardini A, Masi G, Silvestris N, Loretelli C, Ulivi P, Vivaldi C, Bianconi M, Giampieri R, Bittoni A, Andrikou K, Del Prete M, Scartozzi M, Cascinu S. Angiogenesis polymorphisms

- profile in the prediction of clinical outcome of advanced HCC patients receiving sorafenib: Combined analysis of VEGF and HIF-1 α Final results of the ALICE-2 study. *J Clin Oncol* . 2016; 34:280. doi: 10.1200/jco.2016.34.4.suppl.280
26. Guzzi PH, Agapito G, Di Martino MT, Arbitrio M, Tassone P, Tagliaferri P, Cannataro M. DMET-Analyzer: automatic analysis of Affymetrix DMET Data. *BMC Bioinformatics* . 2012; 13:258. doi: 10.1186/1471-2105-13-258
 27. Agapito G, Milano M, Cannataro M. A statistical network pre-processing method to improve relevance and significance of gene lists in microarray gene expression studies. *BMC Bioinformatics* . 2022; 23:1-20. doi: 10.1186/s12859-022-04936-z
 28. Kotlyar M, Pastrello C, Malik Z, Jurisica I. IID 2018 update: context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res* . 2019; 47:581-589. doi: 10.1093/nar/gky1037
 29. Lee S, Liu Y, Li J, Friedman C, Lussier YA. Discovery of protein interaction networks shared by diseases. *Pac Symp Biocomput* . 2007; 76-87.
 30. Paul S, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* . 2003; 13:2498-2504. doi: 10.1101/gr.1239303
 31. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. Cytohubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8 Suppl 4(Suppl 4):S11 doi: 10.1186/1752-0509-8-S4-S11
 32. Agapito G, Cannataro U. Using BioPAX-Parser (BiP) to enrich lists of genes or proteins with pathway data. *BMC Bioinformatics* . 2021; 22:1-35. doi: 10.1186/s12859-021-04297-z
 33. Rahmati S, Abovsky M, Pastrello CM, Kotlyar M, Lu R, Cumbaa CA, Rahman P, Chandran V, Jurisica I. pathDIP 4: an extended pathway annotations and enrichment analysis resource for human, model organisms and domesticated species. *Nucleic Acids Res* . 2020; 48:479-488. doi: 10.1093/nar/gkz989
 34. Farzaneh Z, Vosough M, Agarwal T, Farzaneh M. Critical signaling pathways governing hepatocellular carcinoma behavior; small molecule-based approaches. *Cancer Cell Int* . 2021; 21: 208 doi: 10.1186/s12935-021-01924-w
 35. Kobayashi K, Higai K, Matsuo K, Bannai Y, Horie H, Otakara Y, Li W, Koike K, Yanagino S, Watanabe K, Yoshio T, Wakui N, Nagai H. Quantitative Measurements of Sorafenib in Patients with Advanced Hepatocellular Carcinoma. *Int J Res Stud Med Health Sci*. 2018; 3:14-19.
 36. Ruggiero D, Dalmasso C, Nutile T, Sorice R, Dionisi L, Aversano M, Bröet P, Leutenegger AL, Bourgain C, Ciullo M. Genetics of VEGF serum variation in human isolated populations of cileto: importance of VEGF polymorphisms. *PLoS One* . 2011; 6:e16982. doi: 10.1371/journal.pone.0016982
 37. Al-Habboubi HH, Mahdi N, Abu-Hijleh TM, Abu-Hijleh FM, Sater MS, Almawi WY. The relation of vascular endothelial growth factor (VEGF) gene polymorphisms on VEGF levels and the risk of vasoocclusive crisis in sickle cell disease, *Eur J Haematol* . 2012; 89:403-409. doi: 10.1111/ejh.12003
 38. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc JF, de Oliveira AC, Santoro A, Raoul JL, Forner A, Schwartz M, Porta C, Zeuzem S, Bolondi L, Greten TF, Galle PR, Seitz JF, Borbath I, Häussinger D, Giannaris T, Shan M, Moscovici M, Voliotis D, Bruix J; SHARP Investigators Study Group. SHARP Investigators Study Group Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* . 2008; 24:378-390. doi: 10.1056/NEJMoa0708857
 39. Gardini AC, Faloppi L, Aprile G, Brunetti O, Caparello C, Corbelli J, Chessa L, Bruno D, Ercolani G, Leonetti A, Stefano G, Farella N, Foschi FG, Lanzi A, Dadduzio V, Marisi G, Masi G, Negri FV, Pagan F, Santini D, Scarpi E, Silletta M, Silvestris N, Tamburini E, Tassinari D, Vivaldi C, Gentilucci UV, Zagonel V, Calvetti L, Cascinu S, Frassinetti GL, Scartozzi M. Multicenter prospective study of angiogenesis polymorphism validation in HCC patients treated with sorafenib. An INNOVATE study protocol. *Tumori* . 2018; 104:476-479. doi: 10.5301/tj.5000704
 40. Fang X, Liu Y, Xiao W, Zhao N, Zhu C, Yu D, Zhao Y. Prognostic SLC family genes promote cell proliferation, migration, and invasion in hepatocellular carcinoma. *Acta Biochim Biophys Sin. (Shanghai)* . 2021; 53:1065-1075. doi: 10.1093/abbs/gmab076

41. Arimany-Nardi C, Minuesa G, Keller T, Erkizia I, Koepsell H, Martinez-Picado J, Pastor-Anglada M. Role of Human Organic Cation Transporter 1 (hOCT1) Polymorphisms in Lamivudine (3TC) Uptake and Drug-Drug Interactions. *Front Pharmacol* . 2016; 7:175. doi: 10.3389/fphar.2016.00175
42. Zhao Q, Zheng B, Meng S, Xu Y, Guo J, Chen LJ, Xiao J, Zhang W, Tan ZR, Tang J, Chen L, Chen Y. Increased expression of SLC46A3 to oppose the progression of hepatocellular carcinoma and its effect on sorafenib therapy. *Biomed Pharmacother* . 2019; 114:108864. doi: 10.1016/j.biopha.2019.108864
43. Herraiz E, Lozano E, Macias RI, Vaquero J, Bujanda L, Banales JM, Marin JJ, Briz O. Expression of SLC22A1 variants may affect the response of hepatocellular carcinoma and cholangiocarcinoma to sorafenib. *Hepatology* . 2013; 58:1065-1073. doi: 10.1002/hep.26425
44. Geier A, Macias RI, Bettinger D, Weiss J, Bantel H, Jahn D, Al-Abdulla R, Marin JJ. The lack of the organic cation transporter OCT1 at the plasma membrane of tumor cells precludes a positive response to sorafenib in patients with hepatocellular carcinoma. *Oncotarget* . 2017; 8:15846-15857. doi: 10.18632/oncotarget.15029
45. Liu X, Huang R, Liu X, You H, Kong F, Tang R. Prognostic implications of alcohol dehydrogenases in hepatocellular carcinoma. *BMC Cancer* . 2020; 20:1204. doi: 10.1186/s12885-020-07689-1
46. Osanai M, Lee GH. Increased expression of the retinoic acid-metabolizing enzyme CYP26A1 during the progression of cervical squamous neoplasia and head and neck cancer. *BMC Res Notes* . 2014; 7:697. doi: 10.1186/1756-0500-7-697
47. Osanai M, Lee GH. The retinoic acid-metabolizing enzyme CYP26A1 upregulates fascin and promotes the malignant behavior of breast carcinoma cells. *Oncol Rep* . 2015; 34:850-858. doi: 10.3892/or.2015.4042
48. Yu Y, Wang Y, Zou Y, Yu Y. CYP26A1 Is a Novel Cancer Biomarker of Pancreatic Carcinoma: Evidence from Integration Analysis and In Vitro Experiments. *Dis Markers* . 2022; 2022:5286820. doi: 10.1155/2022/5286820
49. Kanki K, Akechi Y, Ueda C, Tsuchiya H, Shimizu H, Ishijima N, Toriguchi K, Hatano E, Endo K, Hirooka Y, Shiota G. Biological and clinical implications of retinoic acid-responsive genes in human hepatocellular carcinoma cells. *J Hepatol* . 2013; 59:1037-1044. doi: 10.1016/j.jhep.2013.06.024
50. Lotan R. Retinoids as modulators of tumor cells invasion and metastasis. *Semin Cancer Biol*. 199;1 2: 197-208.

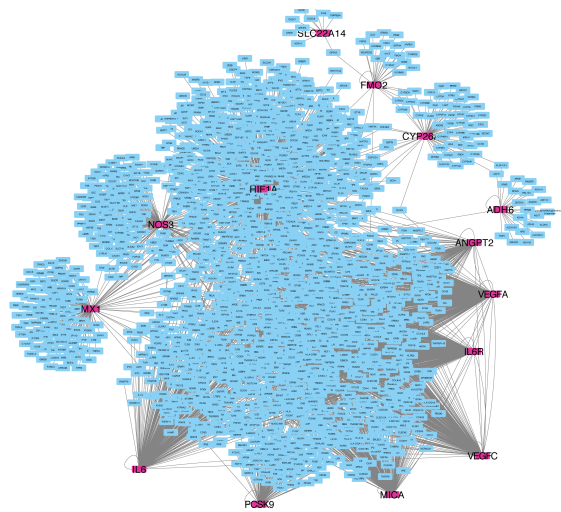
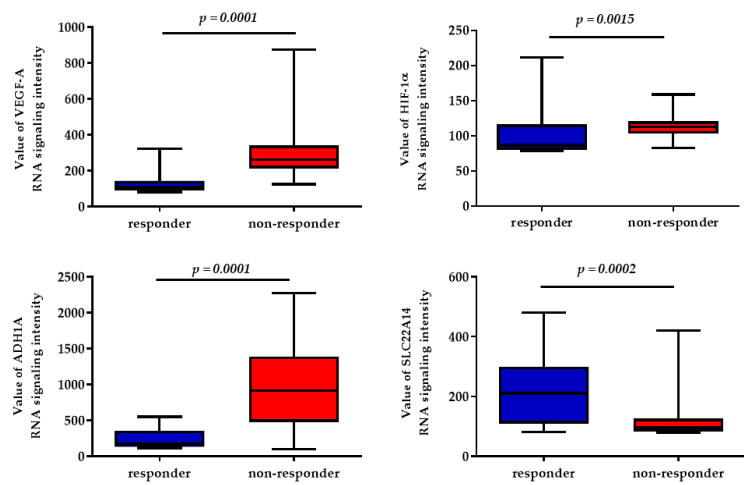
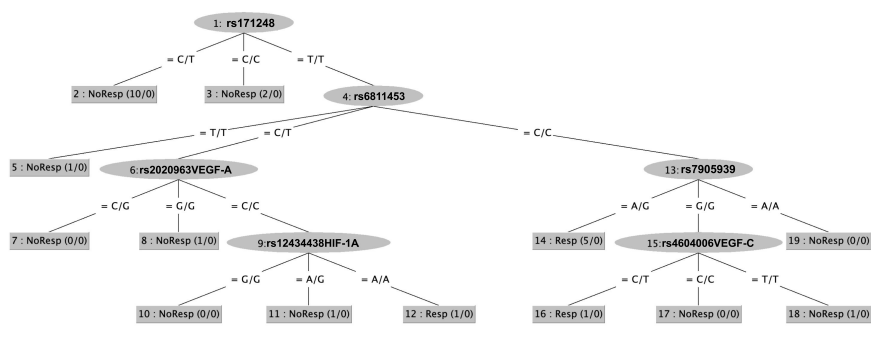
Figure legends

Figure 1 . The Decision Tree computed from the RandomTree classifier, available in Weka, using the sorafenib dataset. The RandomTree's parameters were set up as follows: weka.classifiers.trees. RandomTree -K 0 -M 1.0 -V 0.001 -S 1, the selected Test models 10-fold cross-validation, reaching an accuracy of 69.5652%.

Figure 2. Expression of genes identified by decision tree analysis in HCC patients treated with sorafenib. Data are expressed as mean $\pm$ SD, and the differences between the two groups were assessed with the Mann Whitney U test.

Figure 3. The consolidated network obtained for each seed gene by computing the neighborhoods with a radius of one.

Figure 4 . Top 10 genes by their degree of relevance in the seed network, computed using CytoHubba: high significance values are represented by red, orange, and yellow, while all bluish colors represent less significant values.



Hosted file

Figure 4.tif available at <https://authorea.com/users/635972/articles/653126-genetic-biomarkers-of-sorafenib-response-in-patients-with-hepatocellular-carcinoma>

Hosted file

Table 1.docx available at <https://authorea.com/users/635972/articles/653126-genetic-biomarkers-of-sorafenib-response-in-patients-with-hepatocellular-carcinoma>

Hosted file

Table 2.docx available at <https://authorea.com/users/635972/articles/653126-genetic-biomarkers-of-sorafenib-response-in-patients-with-hepatocellular-carcinoma>

Hosted file

Table 3.docx available at <https://authorea.com/users/635972/articles/653126-genetic-biomarkers-of-sorafenib-response-in-patients-with-hepatocellular-carcinoma>

Hosted file

Table 4.docx available at <https://authorea.com/users/635972/articles/653126-genetic-biomarkers-of-sorafenib-response-in-patients-with-hepatocellular-carcinoma>