Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

ONCOLOGIA, EMATOLOGIA E PATOLOGIA

Ciclo XXXII

Settore Concorsuale: 06/D3

Settore Scientifico Disciplinare: MED06

NGS AND MEDICALLY-DRIVEN INTEGRATIVE BIOINFORMATICS APPLICATIONS IN GASTROINTESTINAL STROMAL TUMORS TO OVERCOME THE TKI RESISTANCE.

Presentata da: Giuseppe Tarantino

Coordinatore Dottorato **Supervisore**

Prof. Pier-Luigi Lollini **Prof.ssa Maria Abbondanza Pantaleo**

Esame finale anno 2020

Sommario

Preface

Thesis outline

Much of the work presented in this thesis comes from prior publications. In some cases, chapters are heavily based on scientific publications, in other cases prior work is only mentioned in passing. Naturally, not all of the ideas and work presented are my own. Besides the presented background material, many of the results and ideas in this thesis have been developed through collaboration with various colleagues, especially my tutor Maria Abbondanza Pantaleo and with my colleagues Valentina Indio, Annalisa Astolfi and Milena Urbini.

In particular, much of the data presented here have been adapted from these three scientific manuscript:

- Maria A. Pantaleo, Giuseppe Tarantino, Claudio Agostinelli, Milena Urbini, Margherita Nannini, Maristella Saponara, Chiara Castelli, Silvia Stacchiotti, Elena Fumagalli, Lidia Gatto, Donatella Santini, Antonio De Leo, Teresa Marafioti, Ayse Akarca, Elena Sabattini, Andrea Pession, Andrea Ardizzoni, Valentina Indio & Annalisa Astolfi (2019): Immune microenvironment profiling of gastrointestinal stromal tumors (GIST) shows gene expression patterns associated to immune checkpoint inhibitors response, OncoImmunology, doi: 10.1080/2162402X.2019.1617588. © 2019 The Author(s). Published with license by Taylor & Francis Group, LLC.
- Indio, V.; Astolfi, A.; Tarantino, G.; Urbini, M.; Patterson, J.; Nannini, M.; Saponara, M.; Gatto, L.; Santini, D.; Do Valle, I.F.; Castellani, G.; Remondini, D.; Fiorentino, M.; Von Mehren, M.; Brandi, G.; Biasco, G.; Heinrich, M.C.; Pantaleo, M.A. Integrated Molecular Characterization of Gastrointestinal Stromal Tumors (GIST) Harboring the Rare D842V Mutation in PDGFRA Gene. Int. J. Mol. Sci. 2018, 19, 732.
- Nannini M, Tarantino G, Indio V, Ravegnini G, Astolfi A, Urbini M, De Leo A, Santini D, Ceccarelli C, Gruppioni E, Altimari A, Castellucci P, Fanti S, Di Scioscio V, Saponara M, Gatto L, Pession A, Martelli PL, Casadio R, Pantaleo MA. Molecular modelling evaluation of exon 18 His845_Asn848delinsPro PDGFRα mutation in a metastatic GIST patient responding to imatinib. Sci Rep. 2019 Feb 18;9(1):2172. doi: 10.1038/s41598-018-38028-x.

Abstract

The complex molecular biology involved in gastrointestinal stromal tumors (GIST) progressive disease led to a growing urgency and interest in developing new strategies to overcome TKI resistance. In this study, we have used Next generation sequencing data and several bioinformatics approaches to investigate the GITS infiltrating immune-cell subpopulations and to perform an integrated molecular characterization of D842V mutant GIST, with the aim evaluating the potential of an anti-PD-L1 treatment in combination with tyrosine kinase inhibitors (TKI) and to identify new possible target. Gene expression profile and immunohistochemistry (IHC) of a cohort of 31 GIST KIT and PDGFRA mutated supported the presence of immune infiltrate in GIST, with dominance of CD4+ and CD8+ T cells and M2 macrophages showing a remarkable similarity with melanoma microenvironment. The expanded IFN-γ-induced immune signature (EIIS) genes were expressed in most of GIST samples and positively correlated with PD-L1 abundance ($p < .0001$). Moreover, the median TIS score for GIST was between the 65th and 70th percentile of the Cancer Genome Atlas dataset, in the same range of tumors responding to anti-PD-1/PD-L1. Analysis of the Gene Expression Omnibus database GIST samples pre- and post-treatment confirmed that imatinib downregulates PD-L1 and IRF1 expression through the inhibition of KIT and PDGFRA, thus contributing to counteract the suppressed adaptive immune response against GIST. The presence of a rich immune infiltrate in GIST along with the presence of TIS and EIIS suggests that GIST may benefit from immunotherapy along with tyrosine kinase inhibitors. Analysis of whole exome sequencing data of 19 D842V mutant GIST samples did not show any actionable recurrent molecular events, beyond D842V, of therapeutic significance. Molecular modeling however, suggests that the D842V mutant protein binds imatinib with lower affinity with respect to wild-type structure, showing higher stability during the interaction with other type I TKIs (like crenolanib).

Prerequisite knowledge

The thesis aims to be fairly self contained, however some knowledge of mathematics, statistics and theoretical computer science is assumed. From mathematics the reader should be familiar with linear algebra, calculus, basic set theory and logic.

INTRODUCTION

Gastrointestinal stromal tumors (GIST) despite being the most common mesenchymal malignant form of the digestive tract, belong to rare tumors with an incidence equal to 1.5 new cases / 100,000 inhabitants / year. GISTs are a subgroup of soft-tissue sarcoma tumors characterized in about 85% of cases by mutations of the KIT proto-oncogene (KIT) or the Platelet derived growth factor receptor alpha (PDGFRA) genes with consequent constitutive activation of downstream signaling cascade (Figure 1; Figure 2). KIT and PDGFRA are members of the type III class of tyrosine kinase receptors that are targets of tyrosine kinase inhibitors (TKI) [1,2,3]. The use of KIT/PDGFRA tyrosine kinase inhibitors (TKIs; imatinib, sunitinib, and regorafenib) has revolutionized the medical treatment of GIST patients. It is known that the sensitivity to TKIs is strictly correlated with the various types of KIT/PDGFRA mutations, with the longest progression-free and overall survival associated to patients whose GIST harbors exon 11 mutations [4,5,6,7,8,9]. Among PDGFRA mutant GIST, different mutations have been described with a variable spectrum of sensitivity to TKIs. Some PDGFRA mutations (e.g., V561D or deletion DIMH842-845) are associated to a clear sensitivity to imatinib that is the first line treatment, in vitro and in clinical studies, whereas other alterations confer treatment resistance in vitro (e.g., PDGFRA D842V, PDGFRA D842Y, or PDGFRA DI842-843IM). Patients KIT / PDGFRA WT (10-15% of GIST) and PDGFRA patients with a D842V mutation are resistant to Imatinib [4,8]. Unfortunately, patients who respond to Imatinib also develop resistance to treatment after a median time of 24 months. For this reason, second and third generation TKI such as sunitinib and regorafenib targeting also kinases involved in tumor-related angiogenesis, were developed and approved for imatinib-refractory GIST treatment. However, the benefit of these second and third generations of TKI is very limited because patients again develop resistance after a median time of 4–6 months. Recently, new pan-TKI inhibitors demonstrated good and interesting results in the prolongation of survivals in early phase trials [10,11]. However, the molecular background of GIST resistant to TKI is very complex due to the acquisition of new mutations and several genome alterations and currently there are no effective, approved treatments available for patients with PDGFRA D842V mutant GIST. Patients' prolonged life expectancy associated with the complex biology

involved in progressive disease led to a growing urgency and interest in developing new strategies to overcome the TKI resistance. Thanks to the introduction of the Next generation sequencing (NGS) techniques it became possible to sequence the entire human genome, the exome (Exome-seq) or the transcriptome (RNA-seq) with relatively short times and costs. Together these technologies have the potential to identify all the genetic anomalies (mutations, fusions between genes, insertions, deletions, rearrangements) of cancer cells in a sensitive and efficient way. They also allow us to study over- and down-expressed genes and enriched or turned off pathways. In addition, the re-introduction of immunotherapy in oncology with the monoclonal antibodies against the immunological checkpoints PD-1 and CTLA-4 revolutionized the treatment of cancer improving survival rates. Many studies on predictive factors of treatment response such as PD-1/PD-L1 expression, tumor mutational burden (TMB), immunogenic antigens or tumor-infiltrating immune cell presence and activation are under evaluation. Only few studies, in preclinical and clinical setting, were conducted investigating the immunological profile of GIST. The identification and characterization by NGS of molecular alterations in GIST tumors and the evaluation of alternative treatments that can allow us to overcome resistance represents the rationale of the present thesis, which in recent years has been developed into two analytical phases. As a first step we dissect the immunological landscape of GIST to provide further rationale for immune-based approach and eventually develop basis to improve the treatment of GIST. Secondly, we proceeded with the study of the molecular background of the subgroup of primary and metastatic D842V mutant GIST using whole exome sequencing (WES) analysis, in order to describe the molecular signature and to pinpoint any additional genomic event potentially relevant for the treatment of these patients.

Gastrointestinal stromal tumor

GIST represents only 0.2% of all neoplasms of the gastrointestinal tract. They are the most common mesenchymal malignant form of the digestive tract but since their incidence is equal to 1.5 new cases of 100,000 inhabitants per year, they are classified as rare neoplasms. The age of onset is quite wide and in most cases GIST arise within the gastrointestinal tract wall, in particular 50-60% originate from the stomach, 20-30% from the small intestine, 5-10% from rectum, colon or esophagus. Typically, the advanced disease is characterized by hepatic and/or peritoneal spread. The pathological diagnosis of

GIST is based on the combination of morphological evaluation, immunohistochemistry (IHC) and molecular analysis. In IHC, more than 90% of GISTs show widespread cytoplasmic positivity for KIT (CD117) [12]. Furthermore, one third of the cases shows positivity to DOG1, which today is considered a new sensitive and specific marker for GISTs [13]. Other immunophenotypic markers, variably expressed but not specific for GISTs are CD34, smooth muscle actin, h-caldesmon and, rarely, desmin. Finally, the molecular analysis of the KIT and PDGFRA mutation status is helpful for the diagnostic confirmation. At the molecular level, about 85% of adult sporadic GISTs are characterized by the presence of mutations against genes encoding the KIT or PDGFRA receptor tyrosine kinases, with consequent constitutive activation of the downstream signaling cascade (Figure 2) . While the remaining 10-15% of cases are defined GIST WT as they do not carry mutations against KIT or PDGFRA [14].

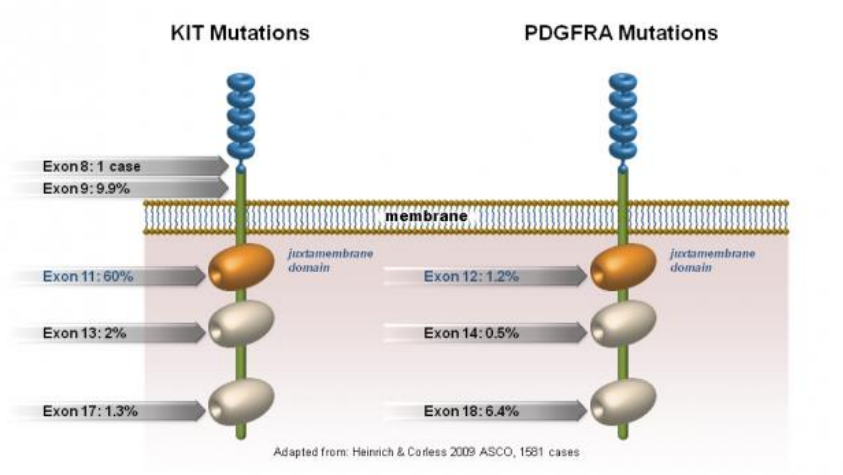


Figure 1*. Schematic representation of the KIT and PDGFRA receptors with highlighted percentages of mutations for exons. Image adapted from http://www.gistsupport.org/about-gist/mutation-analysis-kit-and-pdgfra/*

GIST-KIT^{mut}: 70-80% of GISTs are characterized by the presence of mutations against KIT, proto-oncogene belonging to the family of type III tyrosine kinase receptors. At the structural level, KIT consists of an extracellular portion with 5 Ig-like domains, a transmembrane domain and an intracellular portion containing two tyrosine kinase domains: the first has a binding pocket for ATP and the second has a phosphotransferase region (activation loop) [15]. Following binding to its ligand, SCF, there is dimerization and autophosphorylation of the receptor with a resulting activation of downstream effector

molecules, such as the RAS / MAPK, JAK / STAT, PI3K / AKT pathways involved in the transmission of proliferative stimuli and anti-apoptotic [16]. All the mutations found in KIT in GISTs lead to constitutive activation of the receptor. Most mutations involve the trans-membrane portion of the receptor, encoded by exon 11. These mutations include in-frame deletions, insertions, substitutions, or combinations of these [17]. Deletions are associated with lower progression-free survival [18,19,20]. In particular, deletions involving codon 557 and/or 558 seem to confer a more aggressive biological behavior [21].

GIST-PDGFRA^{mut}: Platelet-Derived Growth Factor Receptor Alpha (PDGFRA) mutations occur in approximately 5–7% of cases. PDGFRA belongs to the type III tyrosine kinase (TK) receptor family. This family is characterized by a specific molecular structure comprising an extracellular (EC) domain and a cytoplasmic domain with a juxtamembrane (JM) region and a TK domain. The EC and cytoplasmatic domain are connected by a transmembrane region. The activation of the receptor occurs as a result of the binding of ligands in the EC domain that lead to dimerization and to a phosphorylation cascade of tyrosine residues in multiple downstream signalling molecules. Inside the TK domain an activation loop (A-loop) has been described, and it conformationally regulates the ATP-binding pocket and leads to kinase activation. Oncogenic PDGFRA mutations activate receptor TKs, resulting in a constitutive phosphorylation. Mutations in the EC domain lead to spontaneous receptor dimerization. Mutations in the cytoplasmic domain instead mainly affect the A-loop encoded by exon 18 (~5%), or more rarely the JM domain encoded by exon 12 (~1%), or the ATP binding domain encoded by exon 14 (<1%). Over half of all PDGFRA mutations are represented by the substitution at position 842 in the A-loop of an aspartic acid (D) with a valine (V), recognized as D842V, conferring primary resistance to imatinib in vitro as well as in clinical observations due to the conformation of the kinase domain, which negatively affects imatinib binding. Most of the PDGFRA mutated GISTs have clinical-pathological and molecular characteristics that distinguish them from the GIST KIT^{mut}. In particular, they are GIST mainly with gastric localization and generally have a low potential for malignancy [22; 23]. Therefore, both the location and the nature of the mutation affect the affinity or binding of imatinib to the kinase and thus are important for imatinib sensitivity.

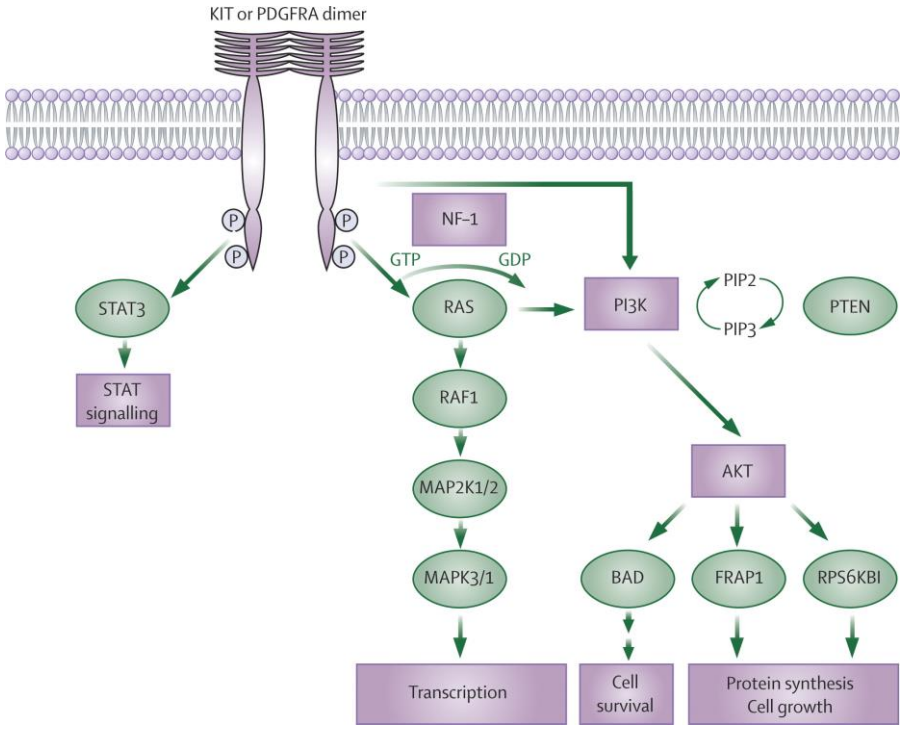


Figure 2. *Transcription factors and pathways activated by KIT and PDGFRA receptors.Image taken from the work of Lancet of Rubin et al. DOI:https://doi.org/10.1016/S0140-6736(07)60780-6*

Another group of GIST consists of the PDGFRA/KIT^{wt} but it is not the object of study of this work.

Treatment of advanced disease

The discovery and introduction in clinical practice of tyrosine kinase inhibitors (TKI) have represented a remarkable progress in the treatment of GIST modifying radically the history and prognosis of patients suffering from this disease. The first line treatment is represented by the Imatinib. An inhibitor of the tyrosine kinase receptors KIT, PDGFRA and Bcr / Abl, approved by the FDA in February 2001 for inoperable or metastatic localized GISTs. Imatinib binds the kinase receptor to the ATP pocket preventing the hydrolysis of the tyrosine kinase domain and blocking all downstream signal transduction pathways [24]. Approximately 80% of GISTs show a primary response to Imatinib. However, response levels are closely influenced by the mutational status of KIT and PDGFRA. Tumors with mutation in exon 11 of KIT are the most responsive to Imatinib (70-85% of responses) as the receptor undergoes a conformational change that favors the binding of the drug [25]. Instead, tumors with mutation in exon 9 have an intermediate response level (25-48%), while those with mutation in exon 13 or 17 are not very responsive [25,26]. Furthermore,

approximately 10-15% of patients are primarily resistant to treatment with Imatinib, or develop resistance within 6 months of starting treatment. These are mainly cases with the D842V mutation in exon 18 of PDGFRA. Finally, around 40-50% of initially responsive cases may develop secondary resistance phenomena [27]. Generally they occur within 24 months of starting treatment with Imatinib and can develop as a result of different mechanisms, like pharmacokinetic variables or the acquisition of structural and functional abnormalities in KIT and PDGFRA receptors or the acquisition of chromosomal alterations. For this reason, the Sunitinib as a second-line treatment for cases not responsive to Imatinib was introduced. The Sunitinib has a broader spectrum of action. Indeed, it is a multitirosin kinase inhibitor approved in 2006, capable of inhibiting: KIT, PDGFRA, PDGFRβ, VEGFR1, VEGFR2 and VEGFR3, FLT3, CSF-1R and RET. Unfortunately, in these cases, the duration of response to treatment with Sunitinib is however limited (generally one year) [28]. Another TKI used for the treatment of GIST is the Regorafenib. It is a multitirosin kinase inhibitor capable of inhibiting: KIT, PDGFR, VEGFR1, VEGFR2, VEGFR3, RET, FGF1R, B-RAF and MAPK. It was introduced thanks to a multicenter phase III trial (GRID Trial) conducted on 199 progressing metastatic patients pre-treated with Imatinib and Sunitinib. A significant advantage of the Regorafenib treatment arm in terms of disease-free survival (4.8 vs. 0.9 months, $p < 0.0001$) compared to placebo was demonstrated in this series [29]. Based on these data, Regorafenib was introduced as a third-line treatment, in case of progression after treatment with Imatinib and Sunitinib. However, even in this case the benefit introduced is poor because again patients after five months of treatments develop a resistance. Due to the enormous variability in the treatment response and to the high frequency of onset of resistance, arise the need to develop new studies to understand the molecular causes of these phenomena and to exploit the potential of new therapeutic approaches.

Next Generation Sequencing (NGS)

In the last decade many progress in the molecular characterization of tumors has been made thanks to the advent of new-generation genomic sequencing technologies (NGS). The NGS technologies have accelerated the production of data, also leading to a consequent reduction in costs. The main innovation consisted in moving from sequencing techniques targeted to a few genes in a few samples, to the possibility of studying thousands of genes

at a time with massive sequencing of entire exomes and even of the entire genome. Generally samples are fragmented and amplified, obtaining DNA library that will then be immobilized on a solid support or on marbles in order to sequence all the fragments in parallel [30]. In the sequencing process, the incorporation signals of the different nuceotids are converted to read. Sequences of a few hundred bases that appear in a given chromosomal region. This "base-calling" procedure also detects a series of parameters, such as signal intensity, background noise and the presence of non-specific signals, to generate the nucleotide sequence and assign quality scores to each base (quality scores). These are related to the probability of error and are a useful tool to eliminate bases or exclude reads that show low quality. The reads are then aligned on the reference genome and then thanks to specific bioinformatics pipelines data are analyzed to identify molecular alterations. the Whole-Exome Sequencing (WES) allows us to analyze only the coding regions of the genome, that is the exome, reducing the complexity, the size of data and the costs of the analysis. With WES data is possible to identify single nucleotide variants (SNV) and insertions or deletions (INDELs) in the coding regions; furthermore it is possible to estimate the copy number alteration (gain or loss), using the reads coverage data in the sequenced regions. Whole-Transcriptome Sequencing (RNA-seq) instead allows sequencing the entire transcriptome starting from mRNA library. It allows to quantify the gene expression profile, to identify the presence of rearrangements of fusion genes and to evaluate alternative splicing. Furthermore it is possible, with specific bioinformatics pipelines, to identify the SNV and INDELs even if with lower accuracy with respect to the WES technique. The main limitation of the transcriptome analysis is represented by the impossibility to identify the presence of variants in the regions not expressed in the sample. These massive sequencing techniques are today the ideal tools for the characterization of the alterations present in the different tumor types, allowing for comprehensive analyzes at relatively low costs and in a short time.

Molecular Docking

The explosion of the next generation sequencing technology has resulted in an increasing number of new therapeutic targets for drug discovery. At the same time, high-throughput protein purification, crystallography and nuclear magnetic resonance spectroscopy techniques have been developed and contributed to many structural details of proteins

and protein–ligand complexes. These advances allow the computational strategies to permeate all aspects of drug discovery today. These computational strategies have the advantages of low cost and effective screening. One of the most common methods is the Molecular Docking. Indeed it has acquired an increasingly importance for pharmaceutical research. Early clarification of ligand-receptor binding mechanism was the lock-and-key theory proposed by Fischer [31] in which the ligand fits into the receptor like lock and key. Then the induced-fit theory created by Koshland [32] takes the lock-and-key a step further. According to this theory the active site of the protein is continually reshaped by interactions with ligands as the ligands interact with the protein. The docking process involves two basic steps: the sampling of the ligand in the active site of the protein, sampling algorithms should be able to reproduce the experimental binding mode; and the assessment of the binding affinity [33]. For the sampling, is important to know the location of the binding site because it increase a lot the efficiency, this is often achieved using a know ligand. Various sampling algorithms have been developed and widely used in molecular docking software. Among the most used there are the stochastic methods like Monte Carlo (MC) and genetic algorithm, which search the conformational space by randomly modifying a ligand conformation or a population of ligands. The conformations are then ranked using a scoring function. The scoring functions are used to delineate the correct poses from the incorrect ones. These functions however compute just an estimation of the binding affinity between the protein and ligand, adopting various assumptions and simplifications. Successful application examples show that computational approaches have the power to screen hits from a huge database and design novel small molecules. Obviously, we have to consider that we are dealing with simulations, the realistic interactions between small molecules and receptors are still relied on experimental technology. Today patients with metastatic/advanced D842V mutant GIST do not benefit from standard TKIs therapy therefore the docking analysis could be used for the screening of all the new TKIs for the identification of novel treatment strategy. In particular in this study we have evaluated the potential of the crenolanib a promising anti-proliferative TKI that is a potent inhibitor of PDGFRA and PDGFRB.

The re-introduction of Immunotherapy in Oncology

Immunotherapy is a cancer treatment that boosts our Immune system to fight cancer. Our immune system protects us from disease, killing bacteria and viruses. One main type of immune cell that does this is called a T cell. T cells have proteins on them that turn ON the immune response and other proteins that turn it OFF. These proteins called checkpoints. Drugs that block checkpoint proteins are called checkpoint inhibitors. These drugs works releasing the brake on the T-cell turning the immune system back ON and making the T cells able to find and attack the cancer cells. Recently, the re-introduction of the Tcell-based immunotherapy in oncology with the monoclonal antibodies against the immunological checkpoints PD-1 and CTLA-4 revolutionized the treatment of cancer improving the survival rates in Melanoma and in the Non Small Cell Lung Carcinoma (NSCLC). However, these molecules do not work in all patients: at present, in fact, we do not have specific criteria to identify ideal patients who will respond to this specific therapy. Many studies on predictive factors of treatment response such as PD-1/PD-L1 expression, tumor mutational burden (TMB), immunogenic antigens or tumor-infiltrating immune cell presence and activation are under evaluation. Responses to immunotherapy preferentially occur in tumors with a preexisting antitumor T-cell response. For this reason, we can define three possibilities for immune phenotypes in tumors (Figure 3). There are “cold” tumors (T cell noninflamed) called cold tumors or immune desert tumors in which the army of T cells is not present and there is a total lack of immune response in the tumor. The second situation is one in which we can distinguish an army of cells ready to attack the tumor but which is not found entirely within the tumor. This means that there is a pre-existing immune response but the T cells are unable to enter the tumor microenvironment. In this situation, inhibiting PD-L1 alone may not be sufficient to stimulate T cell-mediated immunity. The third possibility is that of an T-cell inflamed (or hot) tumor with an army of cancer cell cells, this is what happens in some tumors like melanoma and lung cancer. The immunotherapeutic approaches with PD-L1 and CTLA4 inhibitors have been reported indirectly before for patients with a pre-existing immune response. However, inhibition of PD-L1 or CTLA4 can stimulate T cell-mediated immunity, but does not address other tumor escape mechanisms. Clearly these are simplifications, the real situation is a phenotypic spectrum that goes from the immunity tumor all inflamed.

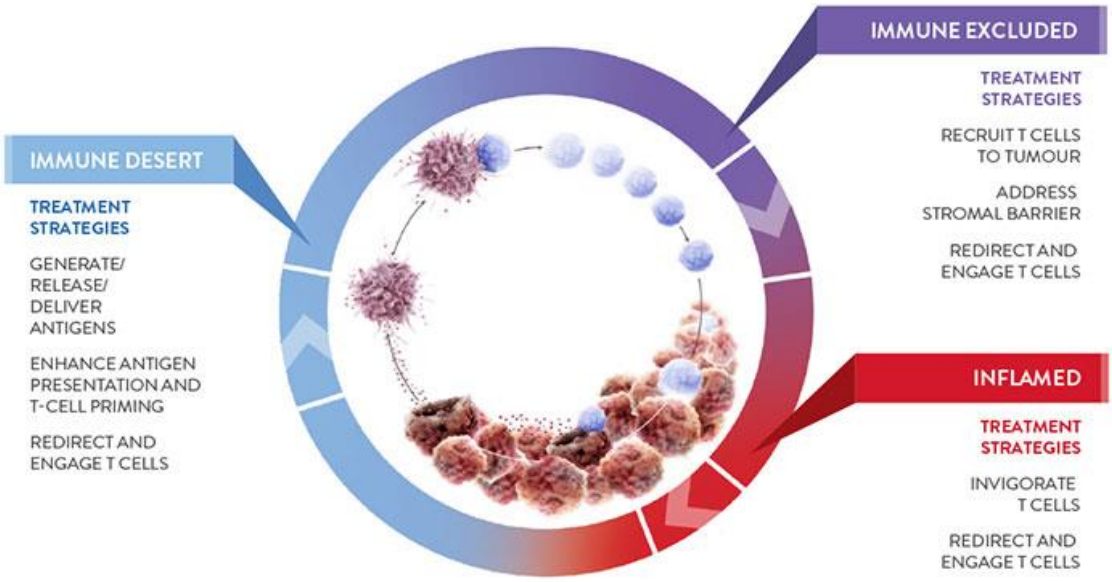


Figure 3. *Phenotypic spectra of tumors and specific treatment strategies. Image obtained from https://www.researchcancerimmunotherapy.com/approach-development/treatment-strategies*

Only few studies, in preclinical and clinical setting, were conducted investigating the immunological profile of GIST. Available data showed that tumor-infiltrating immune cells are present in GIST, and seem to play a role in disease outcome and in increasing the antitumor activity of imatinib. Indeed, In a transgenic mouse GIST model imatinib could polarize tumor-associated macrophages (TAM) to an M2-like phenotype [34]. The TAM were described in the microenvironment of GIST, in particular M2 macrophages were most enriched in metastatic and in imatinib-treated cases. Tumor-infiltrating lymphocytes (TIL) are the second most enriched immune cell population in GIST samples [34,35,36,37].Few data on PD-1/PD-L1 expression in GIST are available showing that PD-L1 tumor expression by immunohistochemistry (IHC) was higher in GIST than in other sarcoma and PD-L1 expression at the mRNA level was heterogeneous across tumors [36,37,38,39]. With regard to GIST patients, a clinical trial on the combination of KIT and CTLA-4 blockade with dasatinib and ipilimumab in GIST reported no synergistic activity, but the number of patients was limited to derive any conclusion [40].

AIM OF THE WORK

About 85% of the gist is characterized by mutations of genes encoding the tyrosine kinase receptors KIT and PDGFRA. Patients with an exon 11 mutation of KIT are sensitive to treatment with Imatinib which is the first-line tyrosine kinase inhibitor. While KIT and PDGFRA patients wt and patients with d842v mutation are resistant to treatment. In addition, even sensitive patients after a median treatment time of 24 months develop resistance. For this reason, second and third generation TKI have been developed, such as sunitinib and regorafenib, which, however, have led to a reduced benefit as patients, after an average treatment period of 5 months, develop resistance. Patients' prolonged life expectancy associated with the complex biology involved in progressive disease led to a growing urgency and interest in developing new strategies to overcome the TKI resistance. For this reason, we developed different analysis adopting several bioinformatics procedure to evaluate the potential of novel treatments:

1. The first one was to dissect the immunological landscape of GIST KIT and PDGFRA muated (about 85% of GIST) to characterize the GIST microenvironment and to provide further rationale for immune-based approach. Since the current medical treatment of GIST relies on multiple and different generations of TKI, it is mandatory that any new treatment approach is conceived in combination with a TKI and therefore evaluating the potential of the checkpoint inhibitors treatment in combination with Imatinib in GIST samples.

2. The second one was to examine the molecular background of the GIST subgroup of primary and metastatic D842V mutant using whole exome sequencing (WES) analysis, in order to describe the molecular signature and to pinpoint any additional genomic event potentially relevant for the treatment of these patients. Exploiting also the in silico evaluation of the effects of the D842V mutation on the PDGFRA protein structure to evaluate the potential of novel available TKI.

MATERIALS AND METHODS

Patients and tumor samples

Both the studies were approved by the local institutional ethical committee of Azienda Ospedaliero-Universitaria Policlinico S.Orsola-Malpighi (approval number 113/2008/U/Tess, 30 September 2008 approval code, approval date) and all patients provided written informed consent. Regarding the study of the molecular background of primary and metastatic D842V PDGFRA mutant GIST on a total of 19 tumor samples and their matched normal counterpart (peripheral blood or stomach) was performed the Whole Exome Sequencing (WES). The 19 tumors collected belong to 14 unique patients: 12 patients with one single tumor sample, one patient with two samples (T04 and T05 corresponding to patient P04), one patient with five samples (T07, T08, T09, T10, and T11 corresponding to patient P06). Patient's multiple samples analyzed correspond to different geographically distinct specimens available. Instead, to investigate the immune landscape of GIST KIT and PDGFRA mutant a total of 31 tumor samples were retrospectively collected and analyzed. All cases were revised by two pathologists with expertise in GIST diagnosis, and all samples were characterized by the presence of KIT or PDGFRA mutation. KIT/PDGFRA WT GIST cases were excluded. Twenty-six samples were from primary tumors and five from metastases. In 26/31 patients, the analysis was done in absence of TKI therapy and only in 5 patients after TKI therapy (2 cases after imatinib and 3 cases after imatinib and sunitinib). Fresh tumor tissue was snap-frozen in liquid nitrogen and stored at −80°C until RNA extraction. This retrospective dataset was obtained coupling samples analyzed with two different analysis. In particular for 19 specimens the analysis was done on HG-U133Plus 2.0 Affymetrix arrays while in 12 tumors by Whole-Transcriptome RNA Sequencing on Illumina platform. , the analysis was done on HG-U133Plus 2.0 Affymetrix arrays while in 12 tumors by Whole-Transcriptome RNA Sequencing on Illumina platform. The tumor and clinical characteristics of both the dataset are reported in Table 1 and Table 2 for the GIST KIT/PDGFRA mutant and for the GIST subgroup PDGFRA mutant D842V respectively.

***Table 1**. 31 GIST Patients analyzed with Microarray and RNA-seq*

Patient	*Sample*	*Sex*	*Age*	*Tissue sample*	*Site*	*size*	*MI*	*KIT/PDGFRA mutation*	*Platform*
P01	GIST002	F	85	Primary	Stomach	8	<5	KIT exon 11 V560D	Microarray
P02	GIST004	M	79	Primary	Stomach	5	7	KIT exon 9 ins 502–503	Microarray
P03	GIST005	M	68	Primary	Stomach	7	4	PDGFRA exon 12 del/ins SPDGHE566–571RIQ	Microarray
P04	GIST008	M	62	Primary	Stomach	2	4	KIT exon 11 V559D	Microarray
P05	GIST009	M	54	Primary	Stomach	3	<5	KIT exon 11 ins TQLPYDHKWEFP 574–585	Microarray
P06	GIST011	M	77	Metastasis	Stomach	20	>10	KIT exon 11 del WK557–558	RNA-seq
P07	GIST012	F	66	Primary	Stomach	4	<5	PDGFRA exon 14 K646E	Microarray
P08	GIST013	M	46	Primary	Ileum	7	<5	KIT exon 11 V559D	Microarray
P09	GIST015	F	64	Primary	Stomach	5.5	<5	PDGFRA exon 18 del DIMH842-845	Microarray
P10	GIST016	F	62	Primary	Stomach	3.7	NA	KIT exon 11 L576P	Microarray
P11	GIST018	M	NA	NA	NA	>5	NA	KIT exon 11 V559G	Microarray
P12	GIST020	M	38	Metastasis	Ileum	NA	NA	KIT exon 11 del MYEQW552–557	Microarray
P13	GIST022	F	76	NA	Stomach	NA	NA	PDGFRA exon18 D842V	Microarray
P14	GIST025	M	84	NA	NA	NA	NA	KIT exon 11 del/ins WKV557–559F	Microarray
P15	GIST026	M	49	Metastasis	Stomach	NA	NA	PDGFRA exon 12 V561D	Microarray
P16	GIST121	M	71	Primary	Stomach	5.5	4	KIT exon 11 V559D	Microarray
P17	GIST124	M	73	Metastasis	Stomach	17	>10	KIT exon 11 ins 1765–1766	RNA-seq
P18	GIST125	F	48	Primary	Stomach	6	5	KIT exon 11 W557R	Microarray
P19	GIST129	M	59	Primary	Stomach	5	5	KIT exon 11 del/ins YEVQWKV553–559L	Microarray
P20	GIST130	F	79	Primary	Stomach	7	>10	KIT exon 9 ins 502-503	Microarray
P21	GIST131	M	68	Metastasis	Ileum	6	3	KIT exon 11 del VYIDPTQLPY569-578	RNA-seq
P22	GIST135	F	61	Primary	Stomach	3.5	6	KIT exon 11 del WKVVE557-561	Microarray
P23	GIST136	M	76	Primary	Stomach	4.5	6	PDGFRA exon 18 D842V	RNA-seq
P24	GIST138	F	75	Primary	Stomach	7	8	PDGFRA exon 18 D842V	RNA-seq
P25	GIST140	F	45	Primary	Stomach	2	2	PDGFRA exon 18 D842V	RNA-seq

P26	GIST142	M	68	Primary	Stomach	3	5	PDGFRA exon 18 D842V	RNA-seq
P27	GIST150	F	58	Primary	Stomach	7	68	KIT exon 11 del PMYE551-554	RNA-seq
P28	GIST165	M	51	Primary	Stomach	12	<2	PDGFRA exon 18 D842V	RNA-seq
P29	GIST174	M	61	Primary	Stomach	6	14	KIT exon 11 L576P	RNA-seq
P30	GIST178	F	70	Primary	Stomach	10	>5	KIT exon 11 V559D	RNA-seq
P31	GIST188	F	62	Primary	Duodenum	4	>5	KIT exon 11 del NGNNYVYIDPTQL564-576	RNA-seq

***Table 2**. Patients and Tumor samples of the GIST PDGFRA D842V mutant*

Patient	***Sample***	***Sex***	***Age***	***Tissue sample***	***Site***	***size***	***MI***	***KIT/PDGFRA mutation***	***Platform***
P01	T01	M	62	Metastasis	Stomach	11	300	PDGFRA exon 18 D842V	WES
P02	T02	F	74	Metastasis	Stomach	16.2	5	PDGFRA exon 18 D842V	WES
P03	T03	M	51	Metastasis	Stomach	5	55	PDGFRA exon 18 D842V	WES
P04	T04	M	51	Metastasis	Stomach	6.2	7	PDGFRA exon 18 D842V	WES
	T05	M	51	Metastasis	Stomach	6.2	7	PDGFRA exon 18 D842V	WES
P05	T06	M	53	Primary	NA	NA	NA	PDGFRA exon 18 D842V	WES
P06	T07	M	56	Metastasis	Stomach	30	high	PDGFRA exon 18 D842V	WES
	T08	M	56	Metastasis	Stomach	30	high	PDGFRA exon 18 D842V	WES
	T09	M	56	Metastasis	Stomach	30	high	PDGFRA exon 18 D842V	WES
	T10	M	56	Metastasis	Stomach	30	high	PDGFRA exon 18 D842V	WES
	T11	M	56	Metastasis	Stomach	30	high	PDGFRA exon 18 D842V	WES
P07	T12	F	63	NA	Stomach	10.5	19	PDGFRA exon 18 D842V	WES
P08	T13	M	76	Primary	Stomach	NA	NA	PDGFRA exon 18 D842V	WES
P09	T14	M	30	Primary	Stomach	NA	NA	PDGFRA exon 18 D842V	WES
P10	T15	F	50	Primary	Stomach	1.8	2	PDGFRA exon 18 D842V	WES
P11	T16	F	75	Primary	Stomach	7	8	PDGFRA exon 18 D842V	WES

P12	T17	M	68	Primary	Stomach	3	5	PDGFRA exon 18 D842V	WES
P13	T18	M	76	Primary	Stomach	4.5	6	PDGFRA exon 18 D842V	WES
P14	T19	M	51	Primary	Stomach	12	2	PDGFRA exon 18 D842V	WES

Whole Exome Sequencing (WES)

WES was done on 19 DNA isolated from fresh frozen and FFPE tumor tissue og GIST D842V mutant and from matched normal peripheral blood or stomach DNA. Whole exome libraries were developed applying different protocols and using two different sequencing platforms: Nextera Rapid Capture Exome Enrichment (Illumina) was adopted on five out of 19 samples that were sequenced on Illumina HiScanSQ at 2 × 100 bp read length; eight out of 19 libraries were prepared with Nimblegen SeqCap v02 (Roche, Pleasanton, CA, USA), and six out of 19 with Nimblegen SeqCap v03 (Roche, Pleasanton, CA, USA) and were run on HiSeq2000 Illumina platform at 100 bp in single-end. For all the three capturing methods, the exome enrichment was performed according to manufacturer's protocols. The figure below (Figure 4) shows the operations that have been implemented in the analysis pipeline.

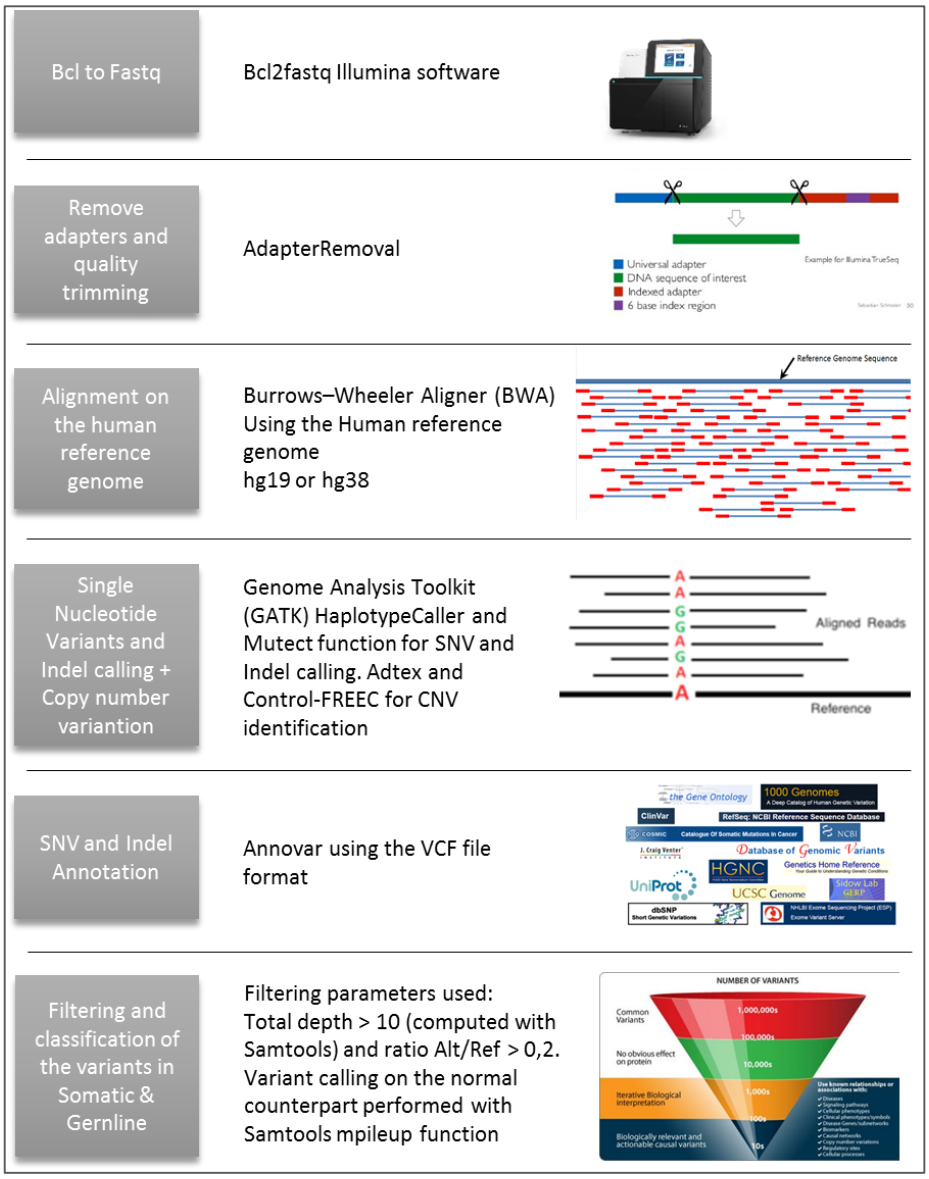


***Figure 4**. Pipeline of analysis of Whole exome sequencing data.*

The first step was the conversion from the Bcl obtained with the Illumina platforms to the fastqfiles. This step was performed with the software bcl2fastq of Illumina. After the conversion shorts reads were processed to remove sequencing adapters and to filter or trim the reads for sequence quality (minimum Phred quality of 10 and minimum length of trimmed sequence of 30). Both these steps were performed with AdapterRemoval v.1.5.4 tool [41]. The cleaned reads were then mapped on the Human reference genome hg19 using the Burrows–Wheeler Aligner (BWA v0.7.12) [42]. The whole exome data were then analyzed with the aim to detect point mutations and copy number variations. First, the realignment around insertions and deletions (InDels) and the base quality recalibration

were performed with Genome Analysis Toolkit v3.3-0 (GATK) [43]. For WES tumor samples, Mutect v1.1.7 [44] and GATK v3.3-0 (HaplotypeCaller function) were then used to call the single nucleotide variants (SNVs) and InDels respectively, adopting the default parameters for both algorithms. Among the total set of called variants, we selected those with total depth > 10, Ratio > 0.2 (ratio between the depth of coverage of the alternate event and the total depth of coverage), within coding exons and having a non-silent effect on protein sequence (non-synonymous and nonsense SNVs, frameshift and non-frameshift InDels). The alterations were filtered on databases of human variability (dbSNP: https://www.ncbi.nlm.nih.gov/SNP/), 1000Genomes: http://www.internationalgenome.org/; EVS: http://evs.gs.washington.edu/EVS/; ExAC: http://exac.broadinstitute.org/) in order to discard polymorphism and keep novel or very rare variants (population frequency < 0.01). All the gene-based and filter-based annotation steps were performed with Annovar v2015Jun16 software tool [45]. Finally, the resulting list of variants were differentiated between somatic or germline by calling the alternate events on the normal counterpart alignments using Samtools v1.4 mpileup function. The somatic variants were handled with SnpSift dbNSFP v4.1 [46], an integrated database of functional computational tools to predict the alteration effects of on protein function and stability. Moreover, on paired tumor/normal WES data, the analysis of amplifications and large deletions was performed applying two different tools, Control FREEC v7.2 [47] and ADTEX v2.0 [48]. A consensus method was realized with the aim of selecting the overlapping regions given by the two algorithms (overlap of gain or loss ≥ 80%) followed by a downstream filtering procedure that takes into account the uncertainty value given by Control FREEC (uncertainty < 80%). Finally, the Database of human Genomic Variants (http://dgv.tcag.ca/dgv/app/home) was adopted to screen out the polymorphic copy number variants.

Creation and processing of the PDGFRA D842V structure for docking analysis

The crystalized structure of the human tyrosine kinase domain of the PDGFRA wild type receptor (PDB: 5K5X) present in the protein data bank (PDB: https://www.rcsb.org/) was used as template for the building by homology of the D842V mutant. The Building by homology was performed with Modeller v9.8 [49]. Instead, the 3D model of the PDGFRA receptor in its active conformation was built adopting as template the kinase domain of

the human receptor KIT (PDB: 1PKG). Since there were not any crystalyzed PDGFRA in its active conformation, we use the KIT receptor as template that shares a sequence identity of 61%. The next step was he generation of the pairwise alignment using ClustalW2 in the PIR format that was used to create 20 different models with modeller. The 3D structure with the lower energy was then selected for the docking procedure.

Autodock and the Lamarkian Genetic Algorithm

AutoDock (version 4.2.6) [50] is a suite of automated docking tools that is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. It involves of two main programs: AutoDock for docking of the ligand to a set of grids describing the target protein; and AutoGrid for pre-calculating these grid maps of interaction energies. Autodock performs well in predicting relative quantities and rankings of series of similar molecules. Like the majority of the docking software it lies on two related steps: The sampling of the ligand in the ligand binding pocket of the protein; and the creation of the ranking of these conformations via a scoring functions. Autodock allows to indicate between different sampling algorithms. Instead, for the assessment of the binding affinity it use the classical force-field-based scoring functions with some extensions. As sampling algorithm of the docking analysis, was used the lamarkian genetic algorithm (LGA). This algorithm like the Monte Carlo is a stochastic process. The idea of the genetic algorithm stems from Darwin's theory of evolution. Degrees of freedom of the ligand are programmed as binary strings called "genes". These genes make up the "chromosome" which actually represents the pose of the ligand. Mutation and crossover are two kinds of genetic operators in GA. Mutation makes random changes to the genes while crossover exchange genes between two chromosomes. When genetic operators affect the genes, the consequence is a new ligand structure (Schematic representation in Figure 5). New structures will be assessed by scoring function, and the ones that survived (i.e., exceeded a threshold) can be used for the next generation.

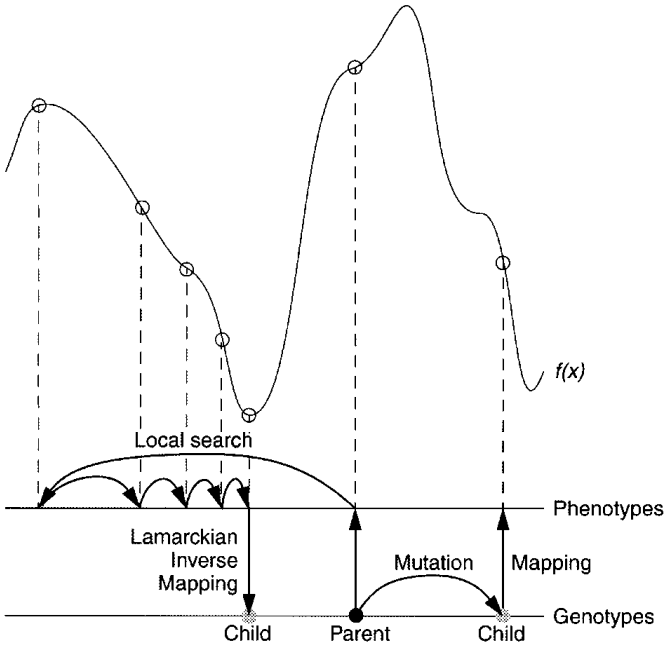


Figure 5. *This figure illustrates genotypic and phenotypic search. Genotypes are mapped to phenotypes by a developmental mapping function. Sufficient iterations of the local search arrive at a local minimum.*

The lamarkian extension combine a local search (LS) with a global one to perform even an energy minimization. The LS method has the advantage that it does not require gradient information about the local energy landscape, facilitating torsional space search. The hybrid of the GA method with the adaptive LS method together form the so-called lamarckian genetic algorithm (LGA) which has enhanced performance relative to the GA alone.

In silico docking procedure

The Rigid Molecular Docking of crenolanib to the modeled PDGFRA was performed with Autodock v4.2. The pdb file was converted in the pdbqt format after the determination of the optimal energetic-states of protonation. This format possesses necessary information for the Docking like the partial charges, and the torsional freedom. Subsequently in order to identify the features of the receptor site, the potential grid maps were created using the program autogrid, fixing the search on the residues of the active site (Figure 6). The creation of the grid is a fundamental step because it allows to register the energetic values of the van der Waals and electrostatic interactions of each atom between ligand and target. This

helps in speeding up the calculations but it impedes to take into account the flexibility of the active site.

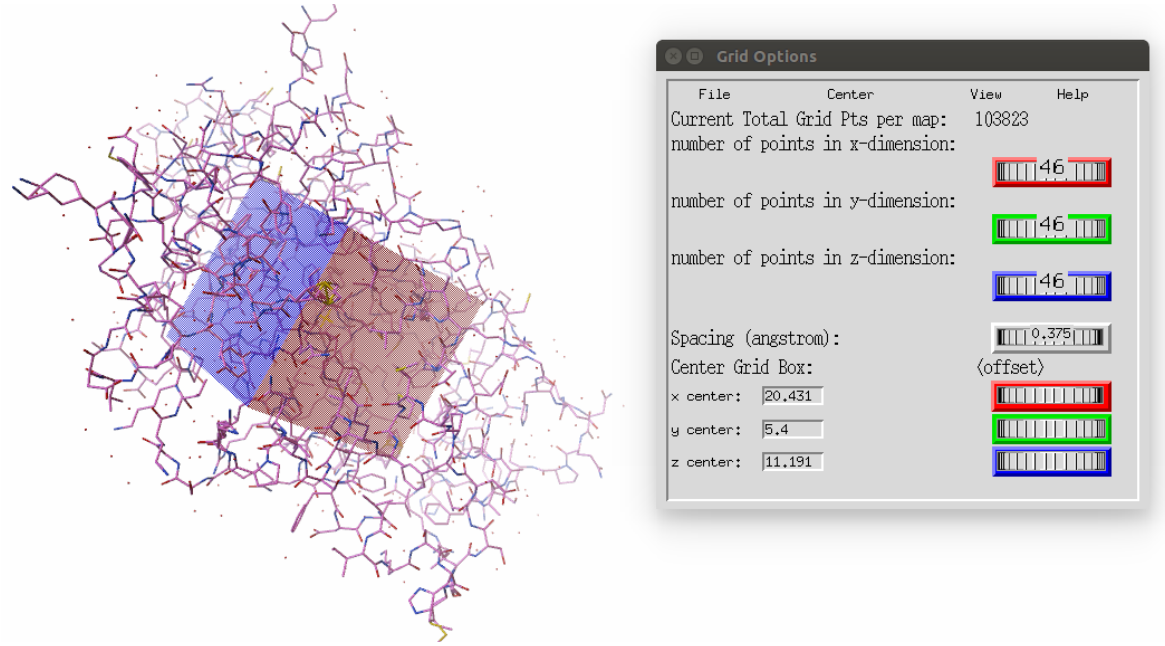


Figure 6. *Creation of the potential grid maps on the LBP.*

The 3D grid box (dimensions 60 × 60 × 60 unit in number of grid points; grid spacing 0.375 Å) centered into the kinase ATP binding pocket was created using autogrid v4.2. In addition, the crenolanib coordinates were generated with PRODRG server [51] and five active torsion angles were set. The Crenolanib was docked using the Lamarkian genetic algorithm (LGA), performing 1500 LGA runs. A generation consists of five stages: mapping and fitness evaluation, selection, crossover, mutation, and elitist selection, in that order. In the LGA, each generation is followed by local search, being performed on a user-defined proportion of the population. This algorithm, like many other docking algorithms, consists of two steps: the sampling of the space and the scoring of the complexes ligand-protein obtained. In order to obtain a global minimum, it mimics the evolutionary process creating clusters of conformations. As a compromise between being too selective or not enough selective I have used a RMSD value of cluster tolerance of 1.0 Å. This value has a huge importance for the creation of the clusters and for the accuracy of the analysis. Also the analysis was performed with an initial population of 300 conformations, a cutoff of 27,000 generations, and with rates of mutation and crossover set to 0.02 and 0.8, respectively. Then for the

ligand the best pose of the best cluster was chosen in terms of “Binding energy”, that can be described as the difference between the benefit of the ligand to be in the binding site rather than in the solvent.

Gene Expression Profiles

As regards the analysis of gene expression, a retrospective dataset previously used in another scientific work by Nannini et al. [52] was used. Gene expression data as described before were obtained from RNA-seq and Microarray analysis. In detail, RNeasy Mini Kit (Qiagen, Milan, Italy) was used to extract the whole RNA from tumor samples. For the RNA-seq samples, the cDNA libraries were created starting from 250 ng total RNA with TruSeq RNA Sample Prep Kit v2 (Illumina, San Diego, CA, USA) ensuing the manufacturer’s protocol. The HiScanSQ sequencer (Illumina) was used to generate sequences at 75bp in paired-end mode obtaining an average of 61 million mapped reads for sample, reaching an average depth of coverage of 44X. The gene expression was quantified using kallisto adopting the transcript per million (TPM) normalization. In particular, Kallisto allow to rapidly quantify the abundances of transcripts performing a pseudoalignment that determine the compatibility of reads with targets. Pseudoalignment of reads preserves the main information needed for quantification making kallisto not only fast but also accurate. For microarray subset, the RNA was quality-controlled and labeled as indicated by the Affymetrix expression technical manual and then hybridized to HG-U133Plus 2.0 arrays. Gene expression data were normalized and quantified by the RMA algorithm using the R package oligo of R-bioconductor.

Prediction of Infiltrating Immune subpopulations

The analytical tool CIBERSORT (Cell-type Identification By Estimating Relative Subsets Of RNA Transcripts) was used for the analysis of 31 tumors samples that were studied either with Affymetrix Array (19 samples) or Illumina whole transcriptome sequencing (12 samples) as described in Table 2. CIBERSORT is a machine learning method based on support vectors machine, trained using gene expression profiles of pure immune

subpopulations isolated by flow cytometry. In such a way that from the complex and convoluted transcriptomic profile of the tumor mass succeeds to identify the immune infiltrate if present and to quantify in a relative and absolute manner 22 immune subpopulations (Schematic representation of training and testing procedure of CIBERSORT in Figure 7).

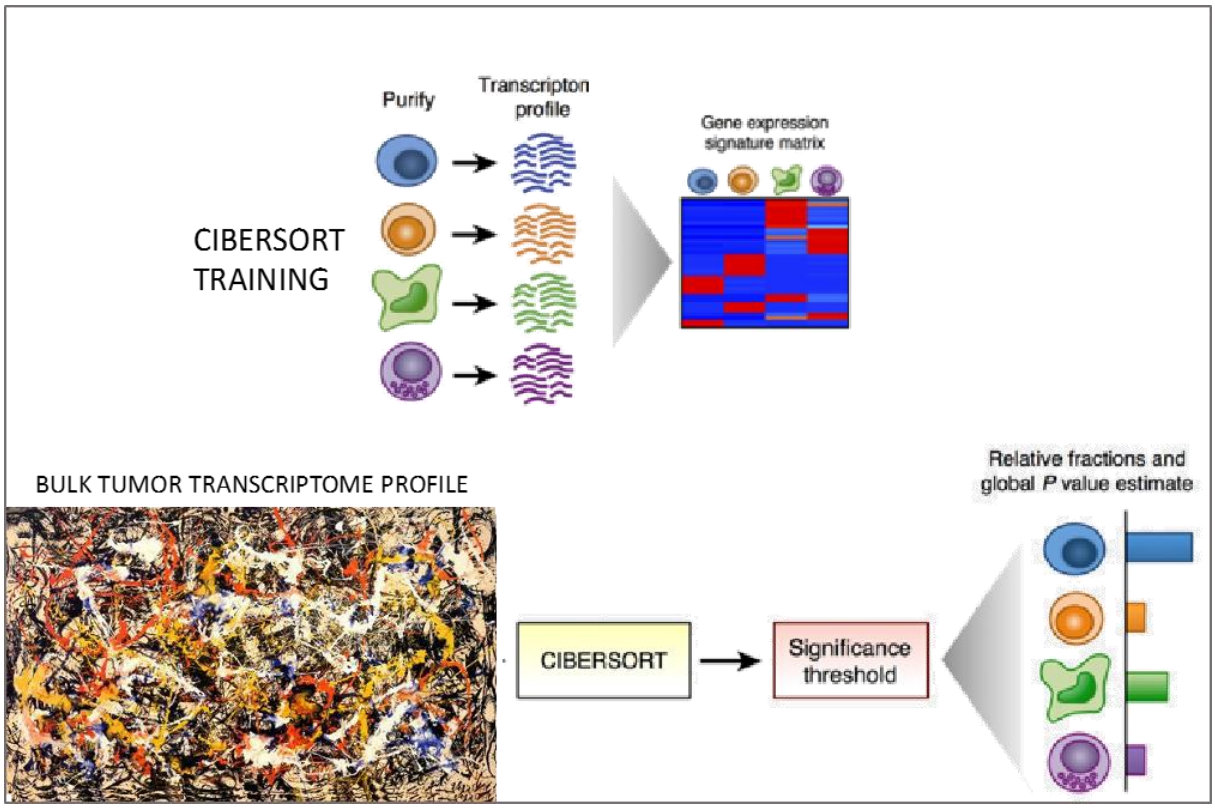


***Figure 7**. This figure shows how cibersort has been trained and what can be done with this software. This figure was adapted from the https://cibersort.stanford.edu home figure.*

Even if the CIBERSORT algorithm was originally developed using microarray data, it was declared as "platform agnostic" [53] and, therefore, applicable to both Affymetrix and Illumina data. The analysis was performed distinctly for the two set of data obtained with different techniques (microarray and RNA-seq). For each set, an unsupervised hierarchical clustering analysis was implemented using the CIBERSORT absolute estimation with the aim to assess the variability of the main cells subpopulations in the tumor microenvironment. Instead, the CIBERSORT relative abundance results were compared between GIST and 18 different solid tumors previously investigated by Gentles et al. [54]. For this purpose, we adopted the microarray subset of GIST because the reference authors based their study on

microarray data as well. More specifically, the 22 immune subpopulations were grouped into 11 cell classes following the representation given by Gentles et al. then the similarity between GIST and other solid tumor was assessed applying two different unsupervised approaches: the hierarchical clustering method and PCA (R-bioconductor, stats package). All the hierarchical clustering were performed adopting the ComplexHeatmap package in R-bioconductor (distance: Euclidean, clustering method: average linkage).

Making RNA-seq and Microarray gene expression data comparable

For both microarray and RNA-seq samples, the transcriptome data were also adopted to measure the expression of specific signatures known to be related to the cancer immune landscape. Since these are two different techniques, in order to use the gene expression data together we had to normalize and $\log_2$ transform the data either with quantile normalization or $\log_2$ TPM calculation respectively for microarray and RNA-seq data. Also an additional normalization was performed by subtracting the arithmetic mean of the $\log_2$ expression data of ten housekeeping gene expression (Table 3).

$$Expression\ GENE = \log_2 x - \frac{\sum_{hk=1}^{10} \log_2 x_i}{10}$$

This set of genes was previously recognized as low variance within a huge set of cancer types and was adopted as a normalization factor for each sample [55]. The gene expression data were used to assess the expression of expanded IFN-γ-induced immune signature (EIIS), a 18 genes signature defined by Ayers et al. (Table 3) [55]. The total EIIS score for each sample was calculated by averaging the expression values of the 18 genes. The achieved scores were correlated with the corresponding PD-L1 expression by sample using the R package stats.

***Table 1**. Housekeeping genes and genes belonging to the expanded IFN-g induced signature (EIIS) defined by Ayers et al.*

Housekeeping genes	***The expanded IFN-γ-induced signature (EIIS)***

Gene name	Gene name	Gene description
ABCF1	CCL5	C-C motif chemokine ligand 5
G6PD	CD2	CD2 molecule
NRDE2	CD3D	CD3d molecule
OAZ1	CD3E	CD3e molecule
POLR2A	CIITA	class II major histocompatibility complex transactivator
SDHA	CXCL10	C-X-C motif chemokine ligand 10
STK11IP	CXCL13	C-X-C motif chemokine ligand 13
TBC1D10B	CXCR6	C-X-C motif chemokine receptor 6
TBP	GZMB	granzyme B
UBB	GZMK	granzyme K
	HLA-DRA	major histocompatibility complex, class II, DR alpha
	HLA-E	major histocompatibility complex, class I, E
	IDO1	indoleamine 2,3-dioxygenase 1
	IL2RG	interleukin 2 receptor subunit gamma
	LAG3	lymphocyte activating 3
	NKG7	natural killer cell granule protein 7
	STAT1	signal transducer and activator of transcription 1
	TAGAP	T cell activation RhoGTPase activating protein

T cell-Inflamed Score (TIS) and analysis of public data

We have also evaluated the T-cell inflamed score (TIS) (Table 4). This signature quantify, using various technology platforms, different but highly correlated gene transcripts associated with the presence of an adaptive immune response that is peripherally suppressed. Measuring a phenotype that seems to be necessary, although not sufficient, for clinical benefit from PD-1/PD-L1 blockade. It consist of genes related to antigen presentation, chemokine expression, cytotoxic activity, and adaptive immune resistance. Starting from a huge tumor data set spanning nine cancer types (KEYNOTE-012 and -028 studies), the authors defined the TIS as a 18 genes signature they determined the TIS score as a linear function defined as:

$$TIS = \sum_{i=1}^{18} x_i w_i$$

where x_i is the ith gene's expression value obtained as described above and w_i is a predefined weight for the ith gene derived by Ayers et al. using logistic regressions models (weights are available in the google document of the patent of Ayers and collaborators). Lately, Danaher et al. computed the TIS score exploiting the gene expression data of the tumors included in TCGA [56]. They found that median TIS scores were higher in tumor types with higher rates of response to PD-1/PD-L1 inhibitors (e.g., melanoma, renal cell cancer), and cancers with high mutation load (e.g., non-small cell lung cancer [NSCLC]), however within each tumor type there was considerable inter-sample variability (Figure 8).

***Table 2**. The T cell Inflamed 18 gene signature defined by Ayers et al.*

Gene name	***Gene description***
CCL5	C-C motif chemokine ligand 5
CD27	CD27 molecule
CD274	CD274 molecule
CD276	CD276 molecule
CD8A	CD8a molecule
CMKLR1	chemerin chemokine-like receptor 1
CXCL9	C-X-C motif chemokine ligand 9
CXCR6	C-X-C motif chemokine receptor 6
HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1
HLA-DRB1	major histocompatibility complex, class II, DR beta 1
HLA-E	major histocompatibility complex, class I, E
IDO1	indoleamine 2,3-dioxygenase 1
LAG3	lymphocyte activating 3
NKG7	natural killer cell granule protein 7
PDCD1LG2	programmed cell death 1 ligand 2
PSMB10	proteasome subunit beta 10
STAT1	signal transducer and activator of transcription 1
TIGIT	T cell immunoreceptor with Ig and ITIM domains

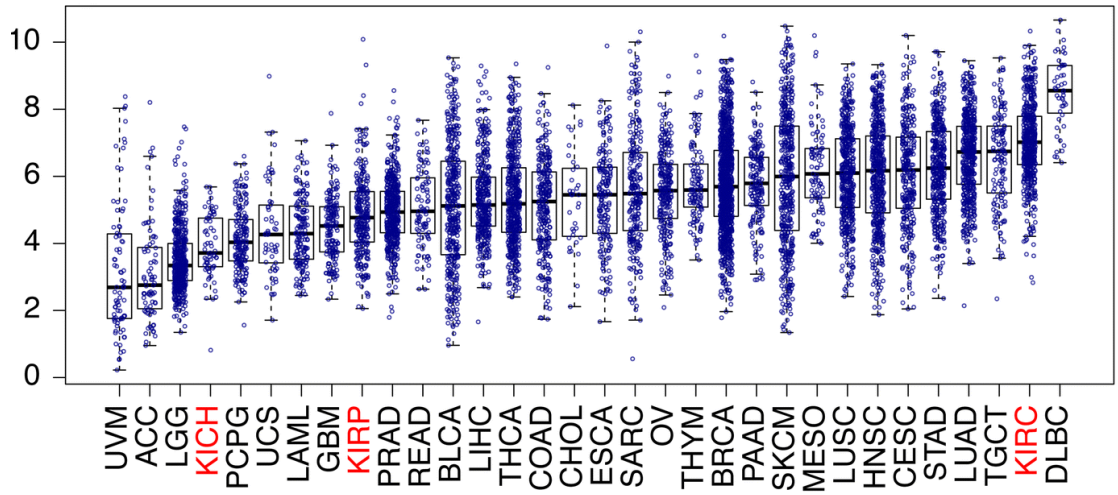


***Figure 8**. TIS scores in all TCGA patients obtained by Danaher et al.*

Following the approach of Danaher and colleagues, we assessed the TIS score in our GIST series with respect to the TCGA cancer subtypes. The whole TCGA was obtained as transcript per million from the Google Cloud Pilot RNA-Sequencing public available web platform (https://osf.io/gqrz9/). In which they have processed over 10,000 RNA-Seqeuencing samples from the Cancer Genome Atlas using kallisto. Further source of open data were explored with the aim of identify imatinib effect on PD-L1 and IRF1 expression. Especially we used the GSE15966 dataset (GEO), consisting of 18 coupled GIST samples pre- and post-imatinib treatment obtained with microarray platform [57]. The LOESS normalized expression data of the coupled GIST samples was achieved with GEO2R tool and the differential expression of PD-L1 and IRF1 between pre- and post-imatinib-treated GIST was estimated by a paired t-test analysis.

Immunohistochemistry

Immunohistochemistry validation was performed on eight GIST FFPE samples by a collaboration with the group of prof. Claudio Agostinelli of the University of Bologna. Antibodies used were the following: anti-CD8 (dilution 1:100, clone 144b, Dako), anti-TIA-1 (dilution 1:300, clone 2G9, Immunotech), FOXP3 (dilution 1:100, clone SP97, Abnova),

anti-CD163 (dilution 1:100 clone 10D6, Leica), and anti-PD-L1 (dilution 1:100, clone E1L3N, Cell Signaling). The immunostained slides were scanned at 200× magnification by Olympus Dot-slide microscope digital system equipped with image analysis software VS-ASW and then assessed in representative areas: each microenvironment marker was independently scored both in the core of the tumor and in the invasive margin of tumor. Results were first reported as mean number of positive tumor-infiltrating immune cells/high power field and than mathematically referred to 1 mm^2. The PD-L1 immunostaining in neoplastic cells was scored as 1+ when <5% positive cells were counted; 2+ when the percentage of stained cells was >5% and <50%; and 3+ when the number of stained cells was >50%.

RESULTS

GIST immune microenvironment

The abundance of 22 immune cell subpopulations infiltrating the tumor was assessed on 31 GIST retrospectively collected samples, in which gene expression profiles (GEP) were previously obtained with microarray (19 samples) and RNA-seq (12 samples) techniques. The two analysis conducted separately highlighted the prevalence of T cells, both CD4+ and CD8+, and M2 macrophages (Figure 9). Overall, the absolute abundance estimation of tumor-infiltrating cells revealed some degree of variability, therefore we investigate if this difference was due to the different molecular classes or to the platform employed but we did not find a significative difference neither between the KIT/ PDGFRA and between Microarray/RNA-seq samples in our dataset (Supplementary Figure 1). In addition, Cibersort results obtained with the two different methods, on a small set of samples analyzed with both the platform, are highly correlated (Pearson r = 0.82–0.89) and cluster together (Supplementary Figure 2). Regarding the abundance of the immune subpopulations we found that macrophages negatively correlated with T cells presence (CD4+ and CD8+ together) supporting the presence of an infiltrating immune microenvironment and likely of an adaptive mechanism of immune escape, as described in other oncological settings (Figure 10).

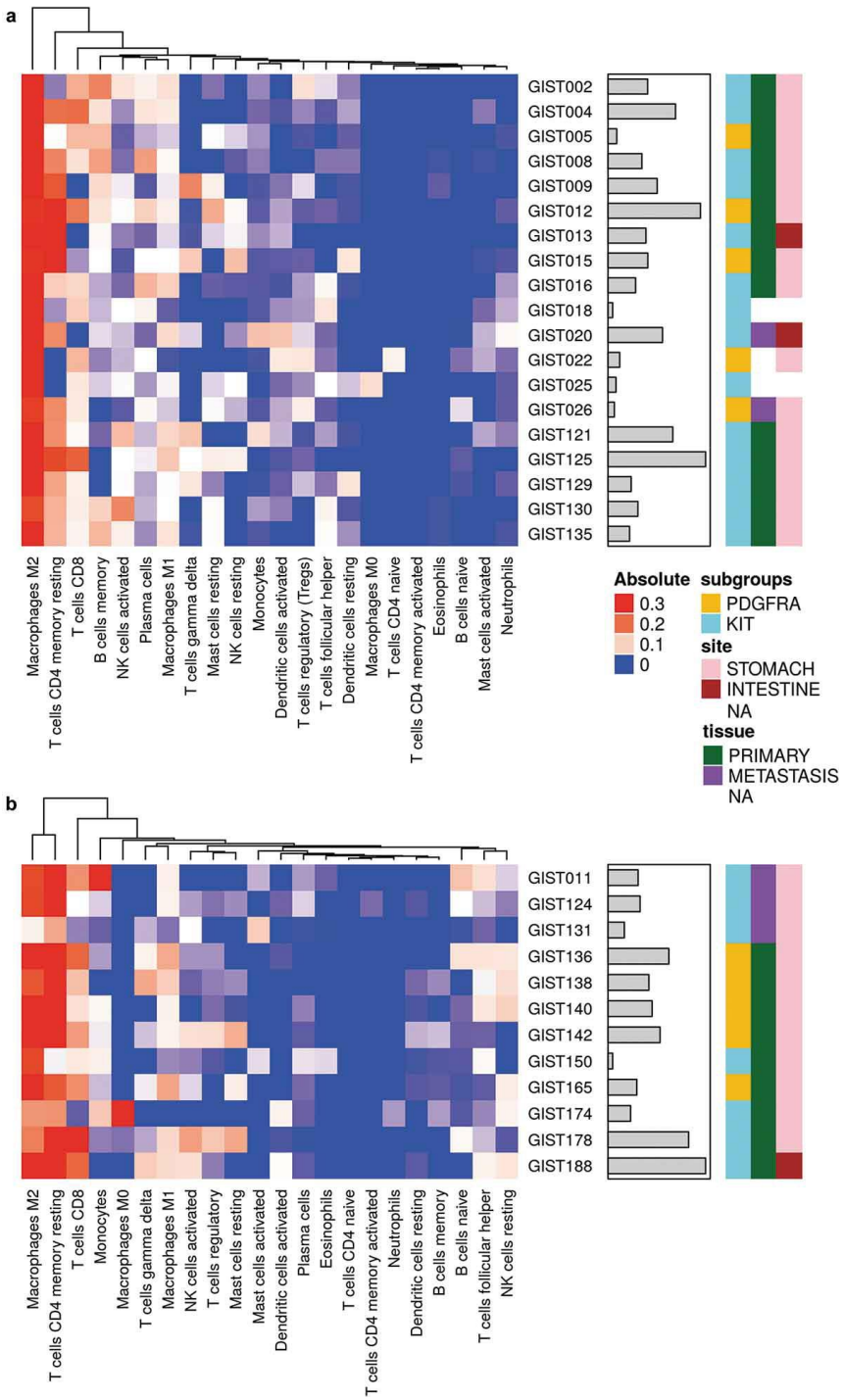


Figure 9. *Heatmap representing the composition of the immune infiltrate signatures by microarray (a) and RNA-seq (b) data with CIBERSORT analysis (absolute abundance). Hierarchical clustering was performed on the infiltrating immune populations using Euclidean distance as a metric of similarity and average linkage as clustering method. The gray bars indicate the total absolute score for each sample. KIT- and PDGFRA-mutant GIST are labeled in cyan and yellow respectively. Tissue samples are labeled in green for primary tumors and purple for metastasis. The tumor site instead is represented with pink and brown boxes for stomach and intestine respectively. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

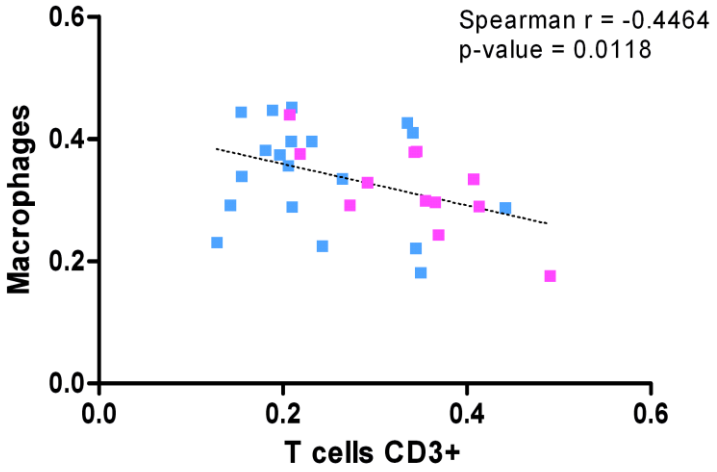


Figure 10. *Correlation between Macrophages and T cell CD3+ (CD4+ and CD8+). Data derived from the two analytical methods microarray and RNA-seq are colored in blue and violet respectively. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

We compared the rate of tumor-infiltrating leukocyte populations between GISTs and several types of solid tumors from AJ Gentle et al. Nat. Med. 2015 using unsupervised methods. GIST samples showed a tumor microenvironment similar to that of metastatic and primary melanoma, one of the tumors that mostly benefits from immunotherapeutic approaches (Figure 11A). The hierarchical clustering technique highlight that GIST display a high abundance of infiltrating $CD8^+$ T cells, similarly to primary and metastatic melanoma, where it is known to be predominantly. This evidence was also supported by another unsupervised approach like the principal component analysis (PCA) (Figure 11B).

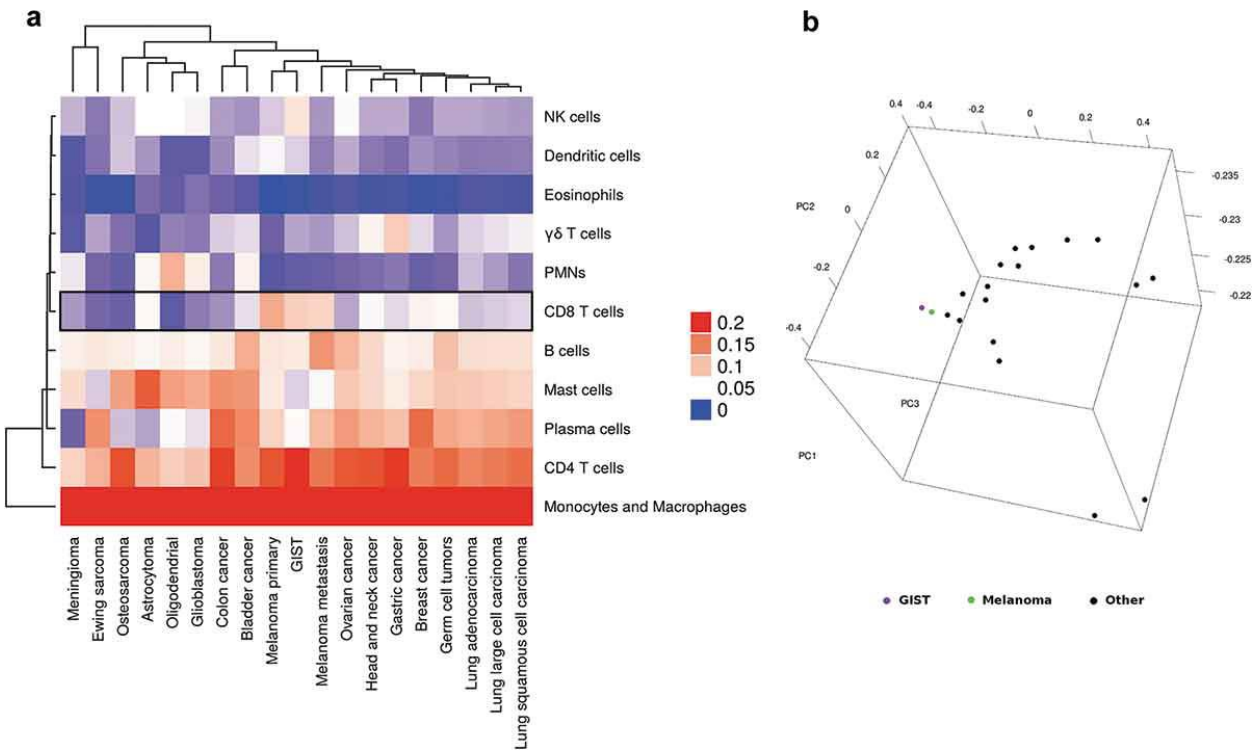


Figure 11. *(a) Unsupervised hierarchical clustering of the tumor-infiltrating composition of GIST and other solid tumor types. The heatmap shows that CD8+ T cells are particularly enriched in GIST and melanoma (primary and metastatic). (b) Principal component analysis of CIBERSORT results of GIST (in purple) and other solid tumors. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

The IHC analysis supported the significant presence of an intra- and peri-tumoral immune infiltrate in GIST identifying as the most abundant subpopulations the T cells $CD8^+$ and M2 macrophages $CD163^+$ (Figure 12). The number of $CD8^+$ lymphocytes and $CD163^+$ elements on average were 17.5 mm^2 and 27.2 mm^2 respectively (Supplementary Table 1). These immune populations were also detected at the invasive margin of the tumors (Supplementary Table 2). It was observed that T cell $CD8^+$ express markers of cytotoxicity like the Tia-1 (Figure 12). The Natural Killer subpopulation $CD16^+$/granulysin$(GNLY)^+$ was also detected in most of the GIST samples (5 of 8 samples) both in the core and in the invasive margin of the tumors (Supplementary Table 1; Supplementary figure 3). Intra-tumoral $FOXP3^+$ T-regulatory lymphocytes were revealed in 4/8 cases. Furthermore, the IHC study revealed a PD-L1 protein expression on neoplastic cells in 50% of the samples analyzed (score +2) (Supplementary Table 1).

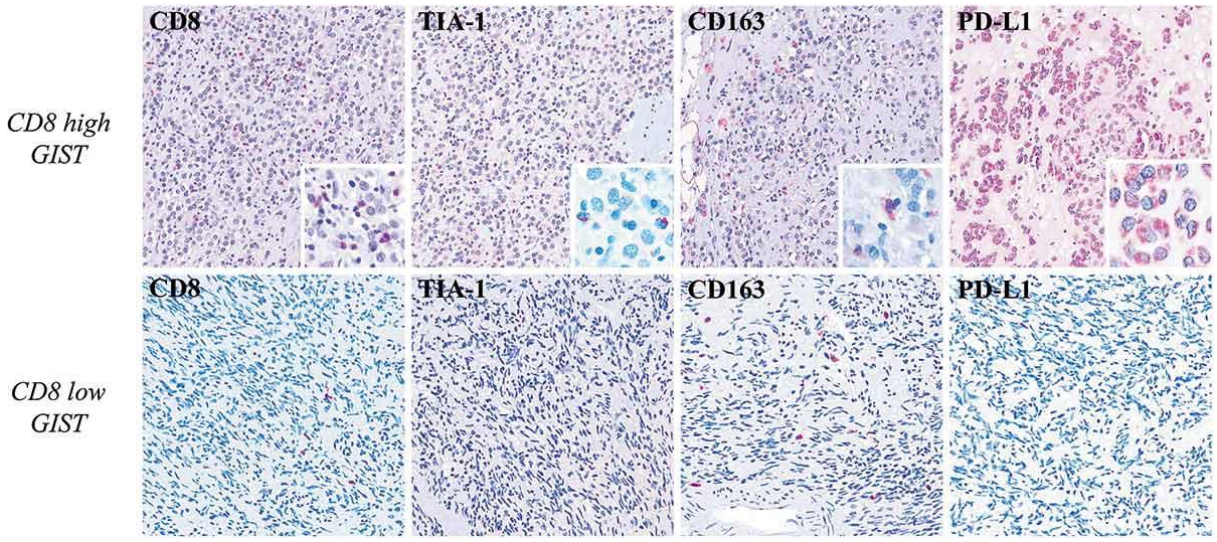


***Figure 12**. Immunohistochemical characterization of GIST samples. In the upper row, one high-CD8+ GIST shows high number of Tia-1+ (x100) (inset: x400) cell of microenvironment, presence of M2 CD163+ macrophages (x100) (inset: x400), and PD-L1 positivity (x100) (inset: x400) in the neoplastic population. The lower row shows one low-CD8+ sample that is characterized instead by a very low number Tia-1+ (x100) cells, presence of M2 CD163+ macrophages (x100), and PD-L1 negativity (x100) in the neoplastic population. The comparison between the two GIST groups high-CD8+ versus low-CD8+ highlights significant differences in the CD8 and TIA1 proteins expression (p = 0.01 and p = 0.02 respectively); differently, there are no significant differences in terms of CD163 and PD-L1 (p = 0.37 and p = 0.71 respectively). Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

GIST express immune signatures predictive of immune checkpoint inhibitor response

Gene expression data were also interrogated in order to evaluate the expression in the 31 GIST samples of several genes representing different immune parameters like the expanded IFN-γ-induced immune signature (EIIS). It was important to verify the presence of this signature because it was recently identified as a predictor of immunotherapy response in head and neck squamous cell carcinoma and melanoma (Table 3). This signature includes 18 genes associated to cytolitic activity, inflammatory cytokines, T-cell markers, antigen presentation, and immunomodulatory factors. We found that this signature was clearly expressed in the majority of GIST samples and, interestingly, it positively correlated with PD-L1 expression ($p < 0.0001$) (Figure 13AB). Remarkably, even

in melanoma samples from the Cancer Genome Atlas (TCGA), the EIIS signature correlates with PD-L1 expression (Figure 17D).

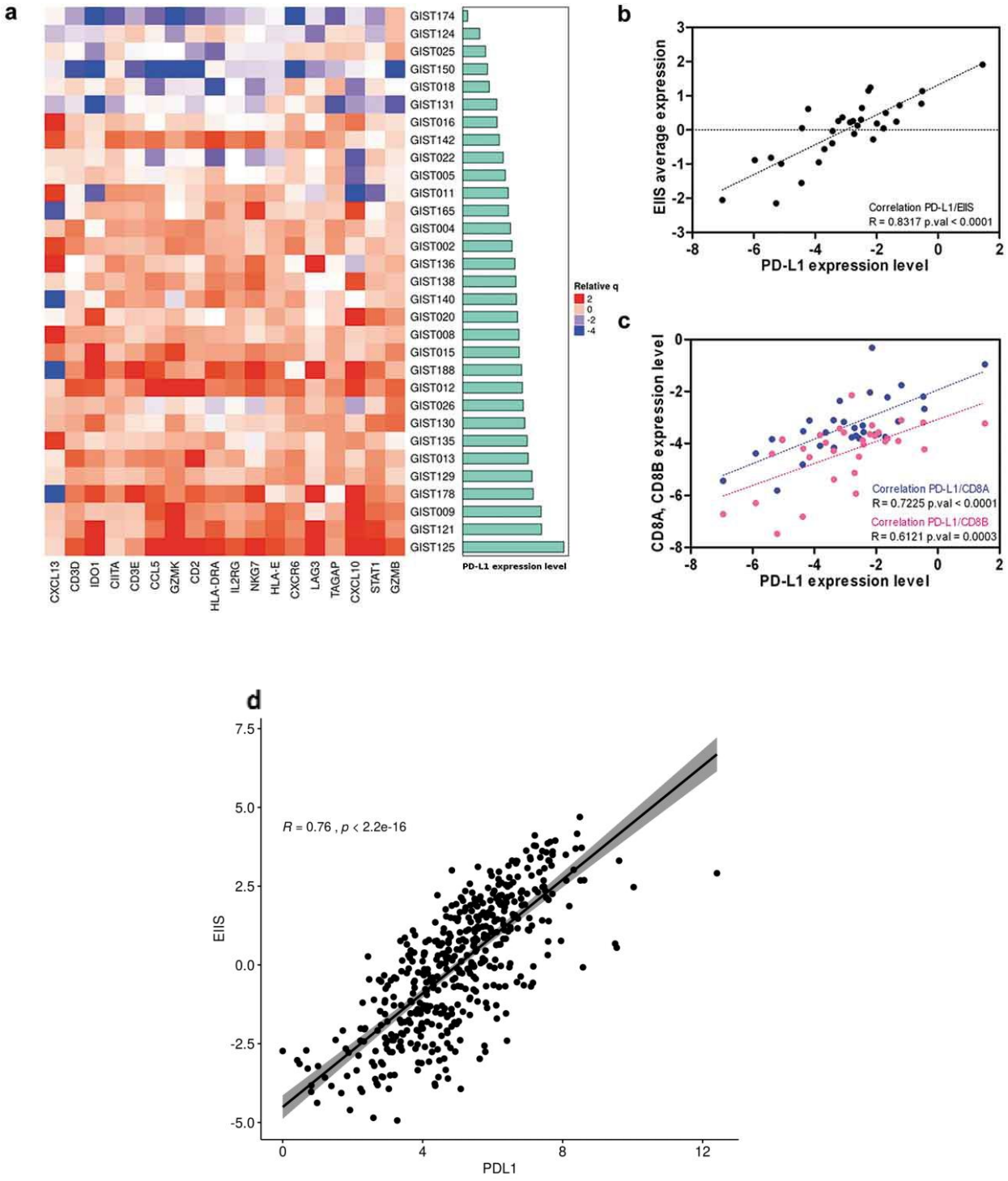


Figure 13. *Figure 4. (a) Heatmap representing the positive correlation between the expanded IFN-γ-induced immune signature and PD-L1 expression (cyan bars) in GIST sample. (b) Scatterplot between the average expression of the EIIS per sample and the PDL1 expression. (c) Positive correlation between PD-L1 and both CD8A and CD8B expressions. (d) EIIS correlation with PD-L1*

expression in Skin cutaneous melanoma (SKCM) samples of the TCGA. Figures adapted from the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588

Moreover, the PD-L1 transcript abundance also positively correlated with the expression of CD8 receptor alpha and beta chains (respectively $p < 0.0001$ and $p = 0.0003$) that are expressed on cytotoxic T cells (Figure 13C). Gene expression data were also investigated to quantify another immune signature related to clinical response to checkpoint inhibitor treatments, the T-cell-inflamed signature (TIS) proposed by Ayers et al. as an improvement of the EIIS described above. This signature measures the presence of an adaptive immune response that was previously shown to be enriched in patients responsive to PD-1 inhibitor treatment (Table 4). Adopting the TIS algorithm we compared the TIS value between our GIST series and other tumors from TCGA dataset. Results showed that GIST have a high median TIS score (8.22), a score very close to tumor types with the highest response rates to anti-PD-1/PD-L1 inhibitors, such as head and neck squamous cell carcinoma (8.10), lung squamous cell carcinoma (8.26), and kidney clear cell carcinoma (8.79) (Figure 14).

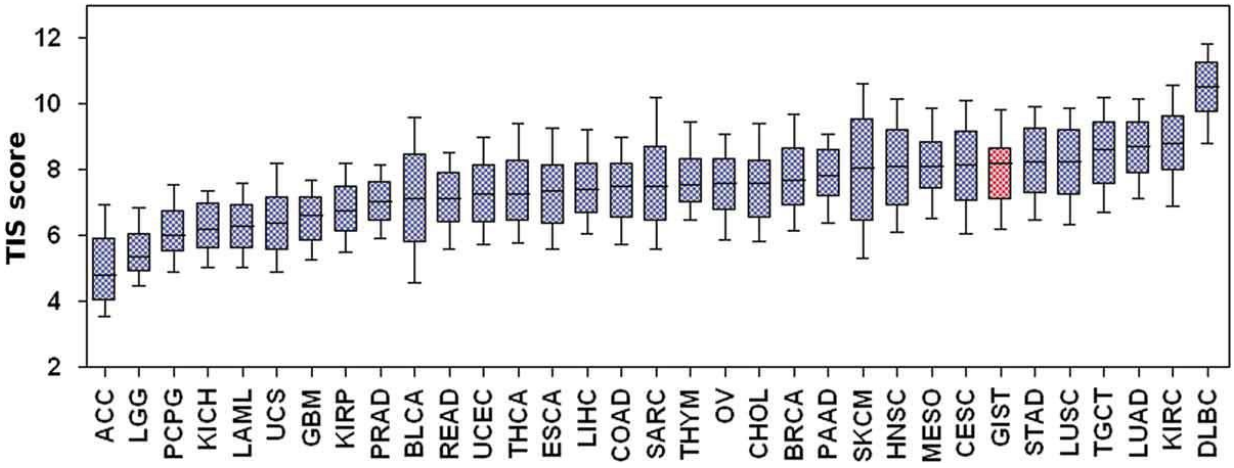


***Figure 14**. T-cell-inflamed signature score of GIST and other solid tumor types from TCGA. ACC: adrenocortical carcinoma; BLCA: bladder urothelial carcinoma; BRCA: breast invasive carcinoma; CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL: cholangiocarcinoma; COAD: colon adenocarcinoma; DLBC: lymphoid neoplasm diffuse large B-cell lymphoma; ESCA: esophageal carcinoma, GBM: glioblastoma multiforme; HNSC: head and neck squamous cell carcinoma; KICH: kidney chromophobe; KIRC: kidney renal clear cell carcinoma; KIRP: kidney renal papillary cell carcinoma; LAML: acute myeloid leukemia; LGG: brain lower grade*

glioma; LIHC: liver hepatocellular carcinoma; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma, MESO: mesothelioma; OV: ovarian serous cystadenocarcinoma; PAAD: pancreatic adenocarcinoma; PCPG: pheochromocytoma and paraganglioma; PRAD: prostate adenocarcinoma; READ: rectum adenocarcinoma; SARC: sarcoma; SKCM: skin cutaneous melanoma; STAD: stomach adenocarcinoma; TGCT: testicular germ cell tumors; THCA: thyroid carcinoma; THYM: thymoma (THYM); UCS: uterine carcinosarcoma; UVM: uveal melanoma. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588

Imatinib downregulates PD-L1 expression in GIST samples

Since the existing medical treatment of GIST relies on several generations of TKI, it is mandatory that any novel treatment approach is conceived in combination with a TKI. It was previously showed by Seifert et al. that imatinib can exert immune modulatory effects and that PD-1/PD-L1 blockade enhances the antitumor efficacy of imatinib in murine GIST model. Also it was demonstrated that JAK inhibitors and imatinib cause a reduction of the expression of PD-L1 in vitro in GIST882 and GIST-T1 cell lines. For this reason, we investigated the possible modulatory effect of imatinib on the expression of the immune checkpoint inhibitors targets in a The Gene Expression Omnibus database (GEO) dataset of GIST tumor samples pre- and post-imatinib treatment. Since have been reported several evidences in the STRING database of the interaction between KIT and PDGFRA with STAT1 (Figure 15) it is conceivable that their signaling may control the expression of PD-L1 through the phosphorylation of STAT1 and the upregulation of IRF1 (Figure 16), as it was described in melanoma [58].

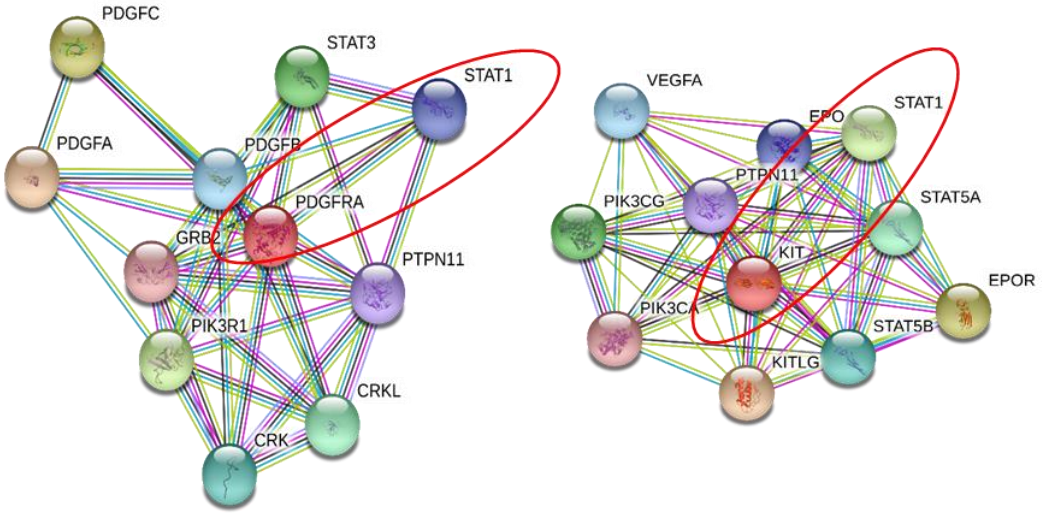


Figure 15. *STRING interactions networks of PDGFRA and KIT. Know interactions are represented with cyan and purple connection, respectively for interactions obtained from curated database and determined experimentally. Instead, predicted ones in green red and blue. The red ellipse was adopted to highlight the know interactions between PDGFRA and KIT with STAT1.*

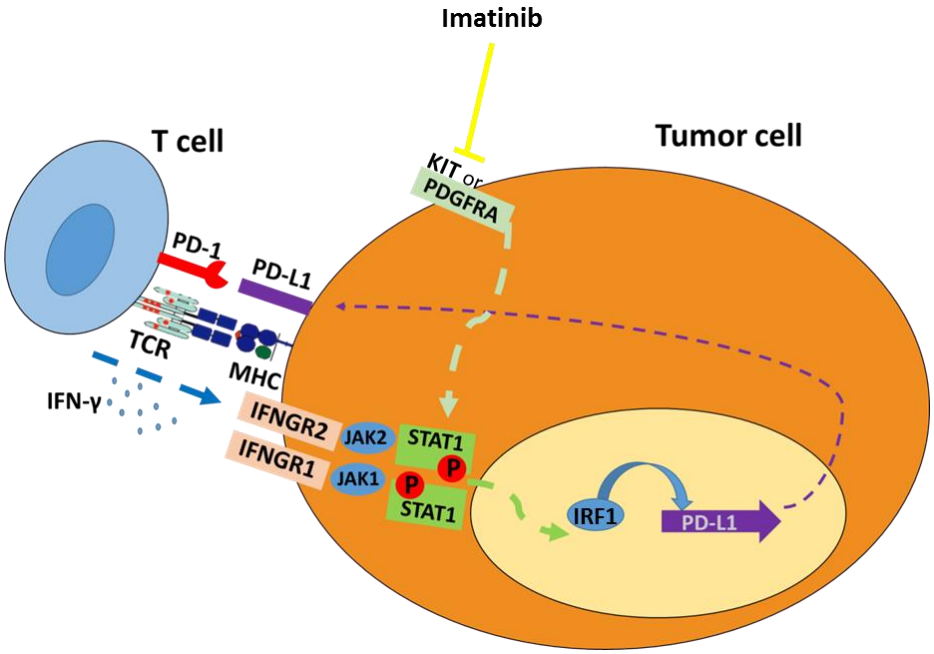


Figure 16*. Proposed model of PD-L1 modulation by imatinib. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

Analysis conducted on public available GIST data revealed that IRF1 and PD-L1 were significantly repressed in imatinib-treated with respect to pre-treatment samples (Figure 17). Together with the previous data reached by in vitro studies, these data from patient samples confirmed the hypothesis of a repressive role of imatinib on PD-L1 expression, supporting a view in which both TKI and checkpoint inhibitors cooperate in the stimulation of an adaptive immune response against GIST.

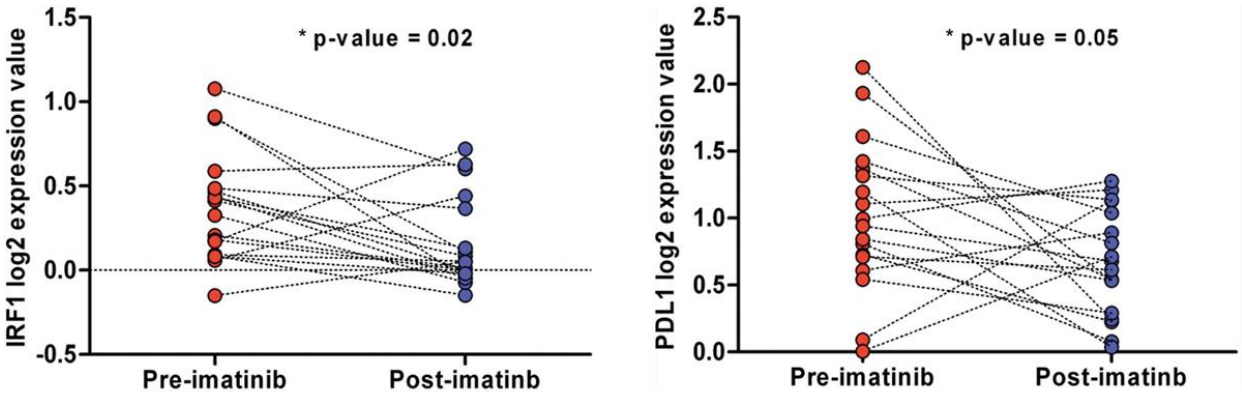


Figure 17. *IRF1 and PD-L1 normalized expression between pre- and post-imatinib-treated GIST samples of the GEO dataset GSE15966. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

Mutational landscape of GIST PDGFRA D842V

The study involved a total number of 19 tumor samples, obtained from 14 unique patients with PDGFRA D842V mutant GIST. To explore the mutational landscape of the D842V mutant GIST, we performed the WES analysis on these tumor samples and on the matched normal counterpart. Our customized bioinformatics pipeline lead to the identification of 316 high-confidence somatic rare variants, including coding single nucleotide variants (SNVs)(Supplementary table 2), frameshift and non-frameshift insertions and deletions (InDels), and variants at ±3 on the splice sites (average = 17, min–max = 3–26). Mutational burden was calculated and reported in Figure 18. The majority of the samples display a low degree of somatic mutations per Mb on captured coding exome ~0.5, which is comparable to the mutational load of chronic lymphocytic leukemia (CLL), neuroblastoma, and glioblastoma. Moreover, two samples (T06 and T14) showed a very low mutational load, similar to acute lymphocytic leukemia (ALL) and medulloblastoma.

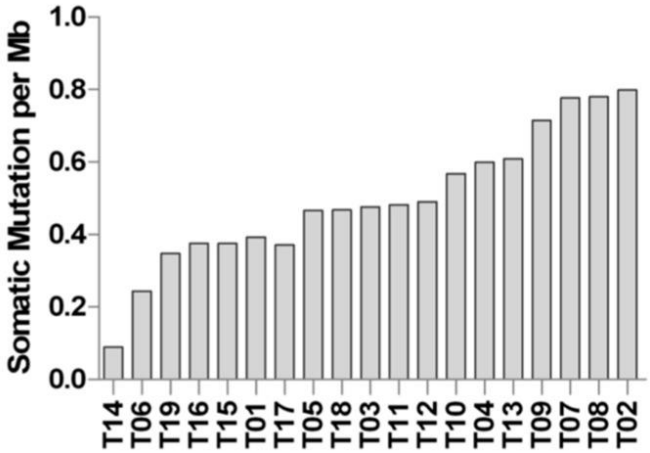


***Figure 18**. Mutational burden of D842V mutant GIST. The histogram bars indicate the number of somatic mutations per magabases (Mb) of coding regions. Figure presented in the work of Indio et al. On Int. J. Mol. Sci. 2018, 19(3), 732; https://doi.org/10.3390/ijms19030732*

The only recurrent somatic mutation was the one on the exon 18 of PDGFRA, the D842V mutant previously identified by molecular testing procedures. The results revealed also three other significantly mutated genes: the transmembrane protein TMEM140, the TEA Domain Transcription Factor 2 (TEAD2), and the Olfactory Receptor Family 1 Subfamily J Member 2 (OR1J2). All the three mutated genes were recurrent in five of the 19 samples. However the samples were five distinct geographically metastases of the same patient (samples T07–11 corresponding to patient P06). Although not significantly characterized, somatic mutations of GRIN2B (Glutamate Receptor, Ionotropic, N-Methyl D-Aspartate 2B) were found in two different samples (N622D on T06, and V42L on T07). For the patients with multiple samples of metastasis the T07–11 (five metastases) and T04–05 (two metastases) the tumor clonal evolution was also estimated (Figure 19).

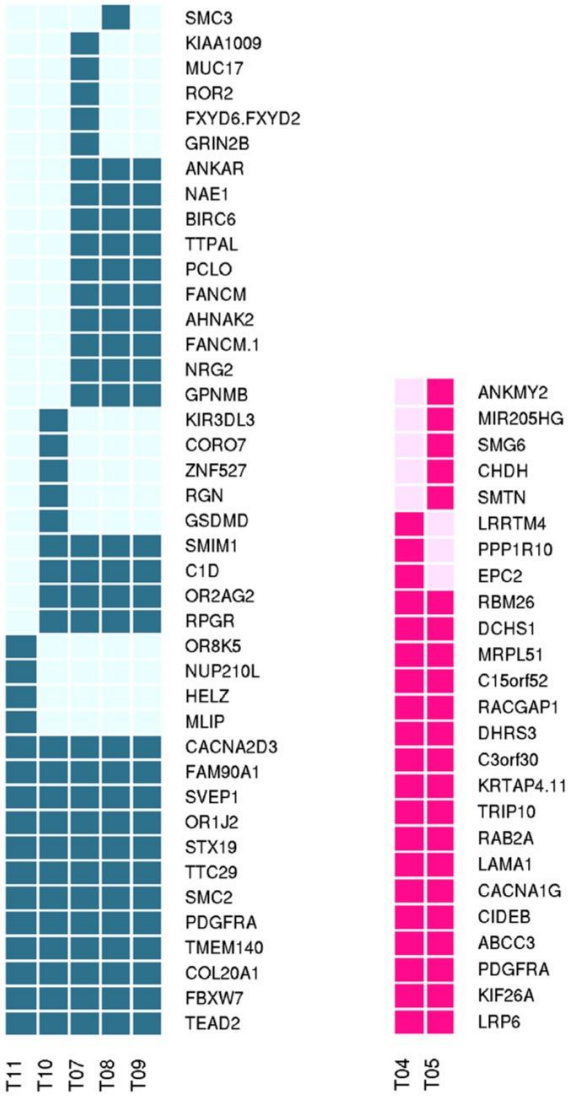


Figure 19. *Graphic representation of clonal evolution of metastatic tumors T07–11 and T04–05. The full filled squares indicate the somatic mutations in the corresponding gene (blue and pink respectively for patients P06 and P04). Figure presented in the work of Indio et al. On Int. J. Mol. Sci. 2018, 19(3), 732; https://doi.org/10.3390/ijms19030732*

All the other somatic events were private variants that we can consider as passenger mutations. Genes with somatic variants were mapped with the TARGET catalogue (tumor alterations relevant for genomics-driven therapy, http://archive.broadinstitute.org/cancer/cga/target). This lead to the identification of three relevant alterations: the Y272C on IDH1 (T02), the R465H on FBXW7 (T07–11), and two different mutations on TP53 (c.993 + 1G > A at exon 10 splice site and a C135G) on sample T13. Interestingly one patient presented a somatic mutation in the succinate dehydrogenase complex subunit B (SDHB), the mutation T60A, and in addition possess the

R38P germline variant in the subunit D of the same complex (SDHD). Both mutations were previously described in the archive of ClinVAr (https://www.ncbi.nlm.nih.gov/clinvar/) as germline variants in GIST, hereditary paraganglioma, and pheochromocytoma. In particular, the T60A in SDHB was annotated as uncertain significance, while the R38P in SDHD was extensively described as pathogenic. Copy number alterations were assessed from WES data (Figure 20). We confirmed the commonly known regions of losses in GIST located on chromosome 1, 14, and 22. We identify at least one of these aberrations in ~90% of tumor samples (corresponding to all patients except P05).

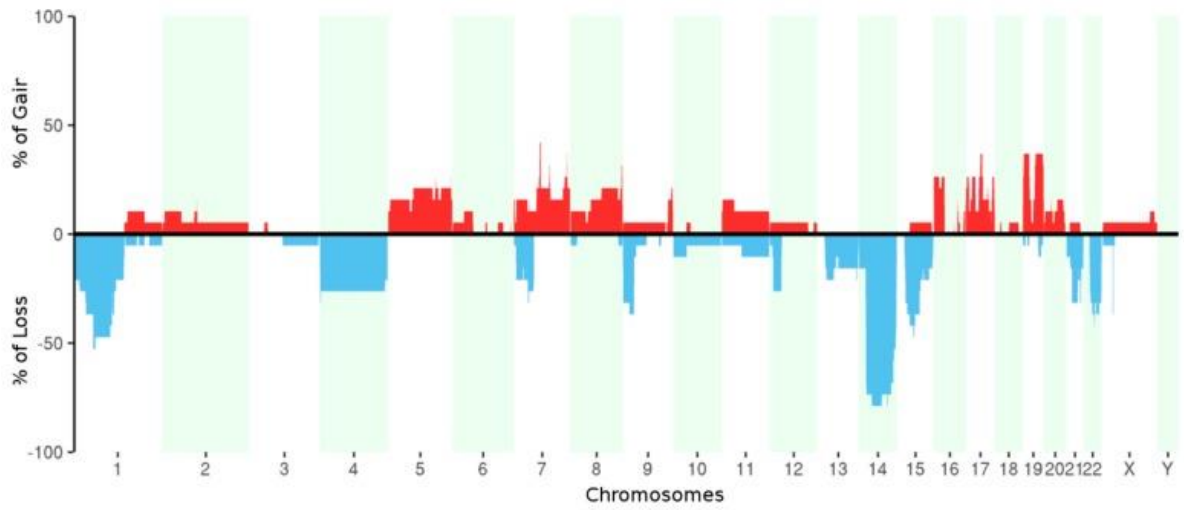


Figure 20. *Percentage of samples with copy number gains (red) and losses (blue) for each chromosome. Figure presented in the work of Indio et al. On Int. J. Mol. Sci. 2018, 19(3), 732; https://doi.org/10.3390/ijms19030732*

Furthermore we found a focal deletion of dystrophin gene (DMD) on chromosome X in the 42% of tumor samples (corresponding to four patients: P03, P06, P07, and P08). This particular aberration is known to be associated with the metastatic tumors [59]

Docking affinity results

Due to the lack of other driver molecular events in the GIST D842V subgroup, we examined the efficacy of the D842V substitution at the protein level, with the goal to define the role of this oncogenic modification within the peptide sequence. The single amino acid substitution of aspartic acid with valine at 842 position (D842V) is known to be related with the resistance to first-line and second-line tyrosine kinase inhibitors (such as imatinib and sunitinib). Since the crystallized structure of the tyrosine kinase domain of PDGFRA (Figure

21A) was recently released (PDB: 5K5X) [60], a protein structure model was created specifically for the D842V mutant and the structure of the complex c-Kit–imatinib (PDB: 1T46) was also used to extrapolate the imatinib 3D structure and the information of the KIT-binding pocket (structural alignment between PDGFRA and KIT with the Imatinib in the figure 21B). The mutation D842V is located in the activation loop (A-loop), known to cover a conserved motif Asp836-Phe837-Gly83 (DFG). This motif may adopt the "in" or "out" conformations corresponding respectively to the active and inactive form of the PDGFRA kinase domain. It was observed by several X-ray crystallographic studies of protein–ligand interactions that imatinib binds to the inactive form (DFG-out) of the Type III transmembrane receptor protein tyrosine kinase (RPTK) subfamily including c-Kit and PDGFRA [61]. Our predicted model also shows that the mutation D842V leads to the loss of polar interactions concerning the residues His845, Ile843, Met844, and Asp846 (Figure 21C,D) that are essential for the stability of the activation loop and for the stabilization of the DFG-out conformation. The consequence is the shift from the DFG-out to the DFG-in form triggering the kinase and making imatinib unable to bind the receptor. To comprehend the molecular basis for the efficacy of the crenolanib against the PDGFRA D842V mutant kinase, we simulated through molecular docking analysis the binding of crenolanib to the modeled PDGFRA in DFG-in conformation. The in silico analysis revealed eight different docking clusters (Table 5) of crenolanib at the ATP binding site with a RMSD > 2Å, and the top scoring docked model was selected (Run 54 , Table 6, Figure 21E).

Table 3. *List of Cluster rank in increasing order of binding energy.*

Cluster Rank	*Lowest Binding Energy*	*Run*	*Mean Binding Energy*	*Num in Clus*
1	-13,19	54	-12,25	37
2	-12,2	100	-10,93	28
3	-11,29	90	-10,34	20
4	-11,05	44	-11,01	2
5	-10,65	98	-10,39	5
6	-10,51	76	-10,3	4
7	-10,02	14	-9,82	3
8	-9,11	82	-9,11	1

Table 4. *List of the Run performed in increasing order of binding energy.*

Rank	*Subrank*	*Run*	*Binding*	*Cluster*	*Reference*
		Energy	*RMSD*	*RMSD*	*Pattern*
1	1	54	-13,19	0	0,92
1	2	26	-13,18	0,85	1
1	3	89	-13,12	1,16	1,23
1	4	62	-13,08	0,91	1,13
1	5	9	-13,06	0,93	1,23
1	6	96	-13,01	1,06	1,29
1	7	38	-13,01	0,95	1,22
1	8	66	-13,01	1,06	1,26
1	9	75	-12,99	0,96	0,97
1	10	5	-12,98	0,91	1,08
1	11	24	-12,91	0,65	1
1	12	25	-12,89	1,22	1,3
1	13	91	-12,68	1,2	1,36
1	14	57	-12,66	0,89	0,96
1	15	68	-12,63	1,2	1,29
1	16	64	-12,58	0,85	1
1	17	80	-12,43	1,05	1,31
1	18	29	-12,25	0,88	0,97
1	19	34	-12,24	1,36	1,32
1	20	6	-12,21	1,14	1,21
1	21	63	-12,2	1,28	1,47
1	22	4	-12,18	1,18	1,32
1	23	16	-12,17	1,2	1,43
1	24	10	-12,04	0,92	0,97
1	25	3	-12	0,95	1,19
1	26	81	-11,94	1,19	1,06
1	27	40	-11,93	1,17	1,18
1	28	20	-11,87	1,03	1,23
1	29	15	-11,82	1,44	1,32

1	30	65	-11,78	1,43	1,4
1	31	11	-11,75	1,31	1,28
1	32	46	-11,74	1,09	1,36
1	33	55	-11,7	1,32	1,23
1	34	88	-11,65	1,15	1,17
1	35	27	-10,75	1,28	1,31
1	36	17	-9,91	1,73	1,43
1	37	48	-9,69	1,8	2,06
2	1	100	-12,2	0	5,83
2	2	83	-11,53	1,64	5,54
2	3	13	-11,51	1,72	5,91
2	4	93	-11,4	1,61	5,79
2	5	36	-11,32	1,72	5,39
2	6	99	-11,28	1,67	5,89
2	7	18	-11,24	1,19	5,32
2	8	85	-11,17	2	5,84
2	9	52	-11,12	1,55	5,61
2	10	37	-11,1	0,94	5,67
2	11	35	-11,1	1,55	5,63
2	12	71	-11,04	1,97	5,76
2	13	79	-11,02	1,99	5,96
2	14	69	-10,99	1,95	5,2
2	15	47	-10,96	1,75	5,96
2	16	41	-10,93	1,33	5,82
2	17	61	-10,88	1,8	5,74
2	18	28	-10,82	1,98	5,14
2	19	87	-10,74	1,76	5,57
2	20	8	-10,73	1,65	5,85
2	21	56	-10,6	1,63	5,51
2	22	7	-10,6	1,68	5,54
2	23	30	-10,53	1,56	5,73
2	24	49	-10,5	1,91	5,08
2	25	72	-10,41	1,73	5,08
2	26	1	-10,32	1,91	5,25
2	27	58	-10,11	1,84	5,16

2	28	21	-9,93	1,95	5,26
3	1	90	-11,29	0	5,14
3	2	23	-10,99	1,54	5,34
3	3	45	-10,97	1,69	5,55
3	4	39	-10,67	1,23	5,15
3	5	84	-10,57	1,21	5,08
3	6	51	-10,52	1,95	5,52
3	7	97	-10,47	1,73	5,37
3	8	92	-10,47	1,45	5,3
3	9	43	-10,47	1,7	5,32
3	10	33	-10,35	1,04	5,36
3	11	77	-10,27	1,11	5,28
3	12	60	-10,2	1,33	5,52
3	13	2	-10,2	1,41	5,09
3	14	78	-10,16	1,51	5,27
3	15	31	-10,16	1,6	5,43
3	16	32	-10,13	1,13	5,31
3	17	86	-9,96	1,83	5,23
3	18	70	-9,92	1,91	5,2
3	19	50	-9,76	1,69	5,36
3	20	19	-9,37	1,77	5,55
4	1	44	-11,05	0	5,68
4	2	67	-10,97	1,24	5,65
5	1	98	-10,65	0	5,8
5	2	53	-10,47	1,85	5,77
5	3	42	-10,36	1,75	5,75
5	4	95	-10,25	1,59	5,56
5	5	73	-10,21	1,16	5,52
6	1	76	-10,51	0	4,45
6	2	94	-10,37	1,12	4,38
6	3	74	-10,23	1,02	4,08
6	4	59	-10,1	1,12	4,01
7	1	14	-10,02	0	5,58
7	2	12	-9,91	1,22	5,79
7	3	22	-9,54	1,41	5,73

8	1	82	-9,11	0	5,67

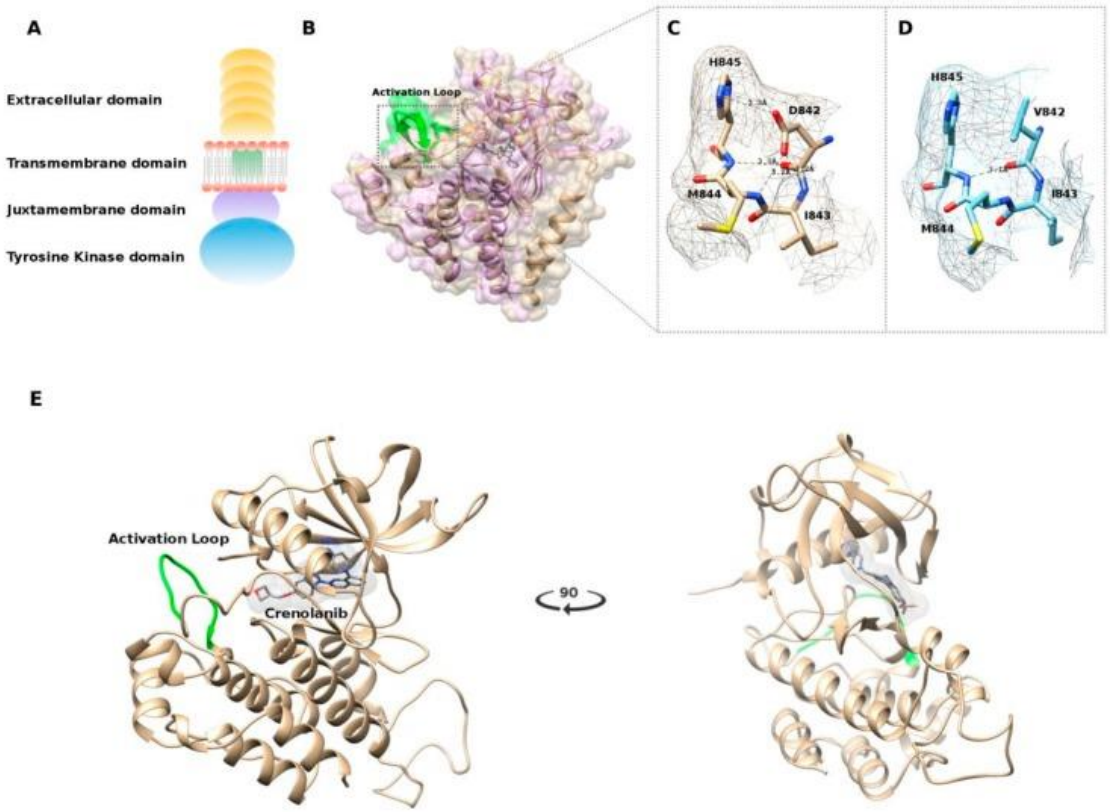


Figure 21. *Figure 5. (A) Schematic representation of the platelet-derived growth factor receptor alpha receptor (PDGFRA); (B) Structure alignment performed using the jCE algorithm between the crystallized structure of c-Kit-Imatinib complex (PDB ID 1T46) in purple and the structure of PDGFRA (PDB ID 5K5X) in gold. PDGFRA (PDB: 5K5X) in gold. Highlighted in green is the Activation loop; (C) Polar interactions of the 842 residue with wild-type Aspartic Acid; (D) Polar interactions of the 842 residue carrying the mutated Valine amino acid; (E) Representation of the best docked pose of crenolanib in the PDGFRA ATP binding site. Figure presented in the work of Indio et al. On Int. J. Mol. Sci. 2018, 19(3), 732; https://doi.org/10.3390/ijms19030732*

The docking analysis suggests that crenolanib, unlike imatinib and sunitinib, targets the active conformation of tyrosine kinase subunit of PDGFRA, in which the activation loop is phosphorylated, binding the ATP active residues without involving the adjacent allosteric site that is available only in the inactive conformation (suitable for the type II kinase inhibitors like imatinib and sunitinib).

DISCUSSION

In this study, PDGFRA and KIT mutant GIST were investigated by exploring serveral bioinformatics application using WES, RNA-seq and molecular modelling data to evaluate novel treatment strategy. In particular, the KIT/PDGFRA mutant GIST cohort analyzed with RNAseq and Microarray allowed us to characterized the tumor microenvironment and to evaluate the potential of the Immunotherapy treatment. Our evaluation showed a significative presence of immune infiltrates in GIST samples with a predominance of T cells, both $CD4^+$ and $CD8^+$, and M2 macrophages. Interestingly, even if our previous analysis identify a low mutational burden the immune profile emerges as closely similar to that of other solid tumor types in particular of melanoma, one of the most striking clinical responders to immunotherapy. Our results were confirmed by a recent work by Vitiello A et al. on the immune profiling of GIST that reported as main subpopulations the Macrophages M2 and the T cell $CD8^+$. They also reported a difference of immune cell infiltrate, immunological activity and expression of immune-related genes between PDGFRA- and KIT-mutant GIST. In our series, there was no significant correlation between TIL amount and clinical and tumor features, and the limited amount of samples did not allow to derive clinical correlations from the study results. Additionally, a significant data resulting from our work is the expression in GIST tumor samples of two gene signatures, the EIIS and TIS, that were recently identified as predictors of immune checkpoint inhibitors response in multiple tumor types. The EIIS was expressed in the majority of GIST samples and, interestingly, it positively correlated with PD-L1 expression. As I have described before this signature was considered a predictor of immunotherapy response in head and neck squamous cell carcinoma, melanoma, and gastric cancer. In our series, we found a positive correlation between the EIIS and PD-L1 expression and a positive correlation between PD-L1 expression and CD8A/CD8B genes supporting an adaptive mechanism of immune escape, as described in other oncological subtypes. In addition, even if the TMB was low, GIST showed high expression of the TIS signature and therefore high values of TIS score that were very close to tumor categories in which high rates of response to PD-1/PD-L1 inhibitors were seen. The TIS score was recently assessed together with the TMB to assess the predictive power in clinical response to anti-PD-1/PD-L1 therapy, stratifying human cancers into different clusters [62] and it was seen that these two predicting method

captures distinct tumor and microenvironment features. Finally, our analysis on public available data of GIST pre and post treatment with imatinib together with the ones obtained in vitro (GIST882 and GIST-T1 cell lines) and in vivo murine models by Saifert et al. have shown the Immune-modulatory effect of imatinib that can inhibit PD-L1 expression. Indeed, multiple reports have shown that imatinib may modulate the immune activity in several ways, ranging from the increase in the number and activation of CD8+ T cells and NK cells to the decrease of Tregs and of IDO expression in the tumor. Moreover, imatinib can synergize with immunotherapy since the treatment with anti-PD-1 or anti-PD-L1 in the Kit V558Δ/+ GIST murine model increases the effect of imatinib by enhancing CD8+ T-cell function. The results of our study on the presence of the immune microenvironment in GIST along with the expression of specific signatures, that are known to be predictive of response to checkpoint inhibitors, suggest that GIST may benefit from immunotherapy together with TKI. Each tumor in the D842V mutant cohort instead, was analyzed independently to highlight point mutations, and copy number variants and alteration at the protein level. Whole exome analysis revealed that no actionable molecular events occurs in this population and confirmed that the only recurrent somatic exomic mutation was the D842V. Many other genes variants were found but were determined to be private genetic events. Even the Copy number analysis lead to the identification of actionable variants. Indeed, a percentage of 42% of samples (corresponding to four unique patients) showed a focal deletion of dystrophin gene (on chromosome X). Deletions of dystrophin in KIT/PDGFRA mutant GIST have been previously reported and usually are associated with more advanced clinical stages of disease such as metastatic tumors. In this study, the deletion of DMD occurred mainly in tumors with a high mitotic index of the primary tumor and in metastatic lesions. Therefore, the main finding of the study on the GIST D842V dataset remains the key role of D842V mutation in this GIST subpopulation as the main and only relevant event of cancer development. These findings highlight the importance on the development of drugs such as the crenolanib that directly inhibit D842V kinase form. It was shown by Heinrich in vitro studies that crenolanib proved to act at least as a 100-fold more potent inhibitor than imatinib on PDGFRA D842V kinase demonstrating a good tolerability profile in phase I clinical studies on advanced solid tumors [63]. In addition a phase II study was recently completed to evaluate the antitumor efficacy and pharmacokinetics of crenolanib in patients with D842V mutant GIST, and a phase III trial of crenolanib versus placebo in combination with best supportive care in patients with D842V

mutant GIST is ongoing. Our analysis lead to the discovery by in silico modelling of eight different docking clusters of crenolanib at the ATP binding site. We identify that Crenolanib unlike imatinib and sunitinib binds the active residues without involving the adjacent allosteric site that is available only in the inactive conformation (suitable for the type II kinase inhibitors like imatinib and sunitinib). This suggests that crenolanib targets the active conformation of the tyrosine kinase domain of PDGFRA in which the activation loop is phosphorylated. Therefore, patients carrying this variant should be considered for treatment with a type I tyrosine kinase inhibitor that targets the ATP binding site when the kinase is in the active conformation such as the crenolanib or the blu-285 that is another type I TKI [64]. Today however, novel immunological approaches are emerging that exploit different immune-modulatory molecules to synergize with imatinib and future clinical trials should be encouraged. In fact, despite new TKIs, BLU-285 and crenolanib demonstrated good and interesting results in the prolongation of survival in early phase trials in the future it is expected that TKI resistance improves and thus more efforts are necessary in order to find a cure beyond or along with TKI for long-term advanced GIST. In conclusion, our study define a comprehensive picture of the GIST microenvironment, suggesting a potential susceptibility to respond to PD-1/PDL1 inhibitors and adds new data useful to build a robust basis for an immunotherapy approach in GIST treatment. In addition, we provides clues on how to target the GIST subgroup D842V mutant.

SUPPLEMENTARY DATA

Supplementary Table 1. *IHC quntification of CD8, TIA1, FOXP3, CD163 and PD-L1. Table presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*

intratumoral immune cells								
ID	**CD8+/mm2**	**TIA1/mm2**	**FOXP3+/mm2**	**CD163+/mm2**	**CD16+GNLY+/mm2**	**cells/mm2**	**PD-L1 tumoral cells(t)**	**PD-L1 t%**
GIST009	58.1	39.9	0.21	9.8	0	68.1	<1%	0%
GIST011	5.6	0.7	0	6.3	1.4	13.3	>1% <50%	30 %
GIST013	79.1	51.8	0	14	7	100.1	<1%	0%
GIST124	24.5	12.6	6.23	39.2	8.4	78.3	>1% <50%	10 %
GIST131	7	3.5	0	37.1	0	44.1	<1%	0%
GIST150	17.5	14.7	4.12	40.6	2.8	65.0	<1%	0%
GIST174	5.6	1.4	0	49	0	54.6	>1% <50%	5%
GIST178	88.9	23.1	2.01	26.6	6.3	123.8	>1% <50%	45 %
	immune cells at the invasive margine of the tumor							
ID	**CD8+/mm2**	**TIA1/mm2**	**FOXP3+/mm2**	**CD163+/mm2**	**CD16+GNLY+/mm2**			
GIST009	2.2	1	0	25.3	0			
GIST011	2.6	0.2	0	3.2	5.6			
GIST013	71.5	4.4	0	57.6	2.1			
GIST124	15.0	0.5	0	54.7	1.4			
GIST131	8.3	0.9	0	9.3	0			
GIST150	13.9	0.5	0.1	54.1	4.9			
GIST174	10.7	0.9	0.4	27.2	0			

GIST178	45.8	2.7	0	13.2	3.5	

Supplementary Table 2*. High confidence somatic SNV and Indel on 19 tumor samples. Table presented in the work of Indio et al. On Int. J. Mol. Sci. 2018, 19(3), 732; https://doi.org/10.3390/ijms19030732*

SAMPLE ID	**GENE**	**CDNA**	**PROTEIN**
T01	ADRA1A	c.G694C	p.V232L
T01	CDC20B	c.C670A	p.H224N
T01	FLNB	c.A1496T	p.E499V
T01	FLNB	c.G1495A	p.E499K
T01	HEATR5A	c.A6055G	p.S2019G
T01	LRIG1	c.T1475A	p.F492Y
T01	PAG1	c.C775A	p.P259T
T01	PDGFRA	c.A2525T	p.D842V
T01	POLE	c.C775T	p.R259C
T01	SLC19A3	c.C232T	p.R78C
T01	SZT2	c.8976delC	p.F2992fs
T01	TWISTNB	c.G40A	p.A14T
T01	USP54	c.C3656T	p.T1219I
T02	ADAMTS19	c.C1430T	p.A477V
T02	BCDIN3D	c.C128T	p.P43L
T02	CALY	c.C119T	p.P40L
T02	CDK20	c.G556C	p.D186H
T02	CNGA1	c.A992T	p.K331M
T02	CNOT1	c.4630_4631insTT	p.L1544fs
T02	COPA	c.G497A	p.G166D
T02	CYFIP1	c.822delG	p.K274fs
T02	DMD	c.G1114A	p.V372M
T02	DMXL1	c.C2684A	p.S895X
T02	EGR1	c.A1001G	p.H334R
T02	ELF4	c.C1346T	p.A449V
T02	FAM3B	c.T677G	p.I226R
T02	FKTN	c.T775C	p.F259L
T02	GPR98	c.C16780T	p.R5594C

T02	IDH1	c.A815G	p.Y272C
T02	IRF1	c.C847T	p.P283S
T02	LCA5L	c.1101_1102insT	p.L367fs
T02	MUC2	c.C6280T	p.L2094F
T02	NPR1	c.C2846T	p.A949V
T02	PCSK2	c.G1511A	p.R504H
T02	PDGFRA	c.A2525T	p.D842V
T02	PLXNA4	c.G841A	p.V281M
T02	SLC7A2	c.C1945T	p.H649Y
T02	TNRC18	c.A5029T	p.R1677W
T02	VCAN	c.A8737G	p.M2913V
T03	ABCB5	c.C566T	p.S189L
T03	CARD9	c.G292A	p.E98K
T03	CCSER1	c.C544T	p.P182S
T03	DAGLB	c.G436A	p.A146T
T03	DEAF1	c.T1127C	p.V376A
T03	HCN2	c.G1760C	p.G587A
T03	LRRC16A	c.G1393T	p.E465X
T03	MCUR1	c.T649G	p.F217V
T03	PDGFRA	c.A2525T	p.D842V
T03	PPP1R3A	c.C143T	p.S48F
T03	PTCHD4	c.G2061A	p.W687X
T03	QTRTD1	c.G121C	p.D41H
T03	ROPN1	c.C458T	p.S153L
T03	SLC17A4	c.G620A	p.S207N
T03	XYLT1	c.G2604A	p.M868I
T03	ZNF420	c.T1655A	p.I552N
T04	ABCC3	c.C3047T	p.A1016V
T04	C15orf52	c.C1520A	p.P507H
T04	C3orf30	c.1421delA	p.Q474fs
T04	CACNA1G	c.T5159C	p.F1720S
T04	CIDEB	c.337_338del	p.113_113del
T04	DCHS1	c.G551C	p.G184A
T04	DHRS3	c.C581G	p.S194X
T04	EPC2	c.C1219T	p.Q407X
T04	GRM4	c.C43G	p.R15G
T04	KRTAP4-11	c.G227A	p.R76H
T04	LAMA1	c.C3055A	p.H1019N

T04	LRRTM4	c.A998T	p.N333I
T04	MRPL51	c.C277T	p.R93X
T04	PDGFRA	c.A2525T	p.D842V
T04	PPP1R10	c.A226C	p.N76H
T04	RAB2A	c.A55G	p.K19E
T04	RACGAP1	c.A41G	p.Q14R
T04	RBM26	c.2188_2189insA	p.Q730fs
T04	TRIP10	c.G475A	p.D159N
T05	ABCC3	c.C3047T	p.A1016V
T05	ANKMY2	c.68-2->TT	nn
T05	C3orf30	c.1421delA	p.Q474fs
T05	CACNA1G	c.T5159C	p.F1720S
T05	CHDH	c.G229A	p.V77M
T05	CIDEB	c.337_338del	p.113_113del
T05	KIF26A	c.C2129T	p.A710V
T05	KRTAP4-11	c.G227A	p.R76H
T05	LAMA1	c.C3055A	p.H1019N
T05	LRP6	c.C4376G	p.P1459R
T05	MIR205HG	c.252_254del	p.84_85del
T05	PDGFRA	c.A2525T	p.D842V
T05	RAB2A	c.A55G	p.K19E
T05	SMG6	c.T1963G	p.F655V
T05	SMTN	c.C2765T	p.S922F
T05	TRIP10	c.G475A	p.D159N
T06	ASPN	c.156_157insTGA	p.E52delinsDE
T06	GALNT12	c.C1087G	p.H363D
T06	GRIN2B	c.A1864G	p.N622D
T06	HGC6.3	c.C239T	p.P80L
T06	OR4N4	c.A515T	p.N172I
T06	PDGFRA	c.A2525T	p.D842V
T06	PEG3	c.C1713G	p.S571R
T06	SDHB	c.A178G	p.T60A
T07	AHNAK2	c.14154_14155insT	p.K4718fs
T07	ANKAR	c.C1283A	p.P428Q
T07	BIRC6	c.A13362C	p.K4454N
T07	CACNA2D3	c.T1879C	p.Y627H
T07	COL20A1	c.T509A	p.F170Y
T07	FANCM	c.G745A	p.G249S

T07	FANCM	c.G1463C	p.S488T
T07	FBXW7	c.G1040A	p.R347H
T07	FXYD6-FXYD2	c.A398G	p.N133S
T07	GPNMB	c.G58C	p.D20H
T07	GRIN2B	c.G124T	p.V42L
T07	KIAA1009	c.A3236G	p.D1079G
T07	MUC17	c.A1793G	p.N598S
T07	NAE1	c.249+2T>A	nn
T07	NRG2	c.C1546A	p.H516N
T07	OR1J2	c.C113T	p.T38M
T07	OR2AG2	c.A520G	p.I174V
T07	PDGFRA	c.A2525T	p.D842V
T07	ROR2	c.C1498A	p.Q500K
T07	RPGR	c.C1163T	p.A388V
T07	SMC2	c.C746T	p.S249L
T07	STX19	c.A881C	p.K294T
T07	SVEP1	c.9565_9566del	p.3189_3189del
T07	TEAD2	c.C189A	p.C63X
T07	TMEM140	c.186_187del	p.62_63del
T07	TTC29	c.T920A	p.L307Q
T08	AHNAK2	c.14154_14155insT	p.K4718fs
T08	ANKAR	c.C1283A	p.P428Q
T08	BIRC6	c.A13362C	p.K4454N
T08	C1D	c.T28A	p.Y10N
T08	CACNA2D3	c.T1879C	p.Y627H
T08	COL20A1	c.T509A	p.F170Y
T08	FAM90A1	c.T83G	p.V28G
T08	FANCM	c.G745A	p.G249S
T08	FANCM	c.G1463C	p.S488T
T08	FBXW7	c.G1040A	p.R347H
T08	GPNMB	c.G58C	p.D20H
T08	NAE1	c.249+2T>A	nn
T08	NRG2	c.C1546A	p.H516N
T08	OR1J2	c.C113T	p.T38M
T08	OR2AG2	c.A520G	p.I174V
T08	PCLO	c.G13547T	p.R4516I
T08	PDGFRA	c.A2525T	p.D842V

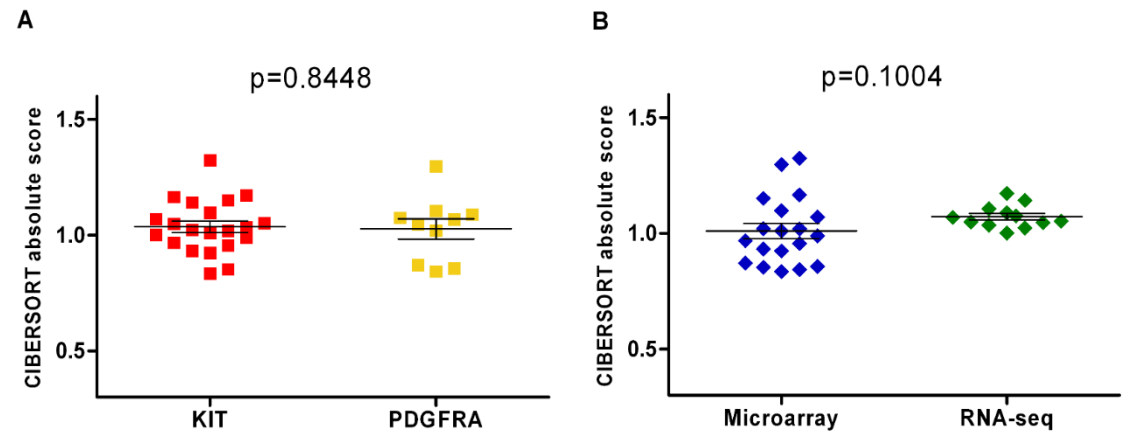
T08	RPGR	c.C1163T	p.A388V
T08	SMC2	c.C746T	p.S249L
T08	SMC3	c.A2135T	p.D712V
T08	STX19	c.A881C	p.K294T
T08	SVEP1	c.9565_9566del	p.3189_3189del
T08	TEAD2	c.C189A	p.C63X
T08	TMEM140	c.186_187del	p.62_63del
T08	TTC29	c.T920A	p.L307Q
T08	TTPAL	c.G334C	p.V112L
T09	AHNAK2	c.14154_14155insT	p.K4718fs
T09	ANKAR	c.C1283A	p.P428Q
T09	BIRC6	c.A13362C	p.K4454N
T09	C1D	c.T28A	p.Y10N
T09	CACNA2D3	c.T1879C	p.Y627H
T09	COL20A1	c.T509A	p.F170Y
T09	FAM90A1	c.T83G	p.V28G
T09	FANCM	c.G745A	p.G249S
T09	FANCM	c.G1463C	p.S488T
T09	FBXW7	c.G1040A	p.R347H
T09	GPNMB	c.G58C	p.D20H
T09	NAE1	c.249+2T>A	nn
T09	NRG2	c.C1546A	p.H516N
T09	OR1J2	c.C113T	p.T38M
T09	OR2AG2	c.A520G	p.I174V
T09	PCLO	c.G13547T	p.R4516I
T09	PDGFRA	c.A2525T	p.D842V
T09	RPGR	c.C1163T	p.A388V
T09	SMC2	c.C746T	p.S249L
T09	STX19	c.A881C	p.K294T
T09	TEAD2	c.C189A	p.C63X
T09	TMEM140	c.186_187del	p.62_63del
T09	TTC29	c.T920A	p.L307Q
T09	TTPAL	c.G334C	p.V112L
T10	CACNA2D3	c.T1879C	p.Y627H
T10	COL20A1	c.T509A	p.F170Y
T10	CORO7	c.G409A	p.A137T
T10	FAM90A1	c.T83G	p.V28G
T10	FBXW7	c.G1040A	p.R347H

T10	GSDMD	c.C1430A	p.S477X
T10	OR1J2	c.C113T	p.T38M
T10	OR2AG2	c.A520G	p.I174V
T10	PDGFRA	c.A2525T	p.D842V
T10	RGN	c.G250T	p.E84X
T10	RPGR	c.C1163T	p.A388V
T10	SMC2	c.C746T	p.S249L
T10	SMIM1	c.61_77del	p.21_26del
T10	STX19	c.A881C	p.K294T
T10	SVEP1	c.9565_9566del	p.3189_3189del
T10	TEAD2	c.C189A	p.C63X
T10	TMEM140	c.186_187del	p.62_63del
T10	TTC29	c.T920A	p.L307Q
T10	ZNF527	c.A43T	p.T15S
T11	CACNA2D3	c.T1879C	p.Y627H
T11	COL20A1	c.T509A	p.F170Y
T11	FAM90A1	c.T83G	p.V28G
T11	FBXW7	c.G1154A	p.R385H
T11	HELZ	c.A3949C	p.S1317R
T11	MLIP	c.G709T	p.E237X
T11	NUP210L	c.G3316T	p.G1106C
T11	OR1J2	c.C113T	p.T38M
T11	OR8K5	c.A617T	p.N206I
T11	PDGFRA	c.A2525T	p.D842V
T11	SMC2	c.C746T	p.S249L
T11	STX19	c.A881C	p.K294T
T11	SVEP1	c.9565_9566del	p.3189_3189del
T11	TEAD2	c.C189A	p.C63X
T11	TMEM140	c.186_187del	p.62_63del
T11	TTC29	c.T920A	p.L307Q
T12	AHRR	c.G1772C	p.R591T
T12	ARHGDIB	c.T321A	p.Y107X
T12	C15orf40	c.396_397insTT	p.L132fs
T12	CABIN1	c.C4408T	p.P1470S
T12	CELF1	c.G604C	p.G202R
T12	COL5A1	c.C1888T	p.R630W
T12	DTNB	c.G1448A	p.R483Q
T12	EPX	c.C1522T	p.R508W

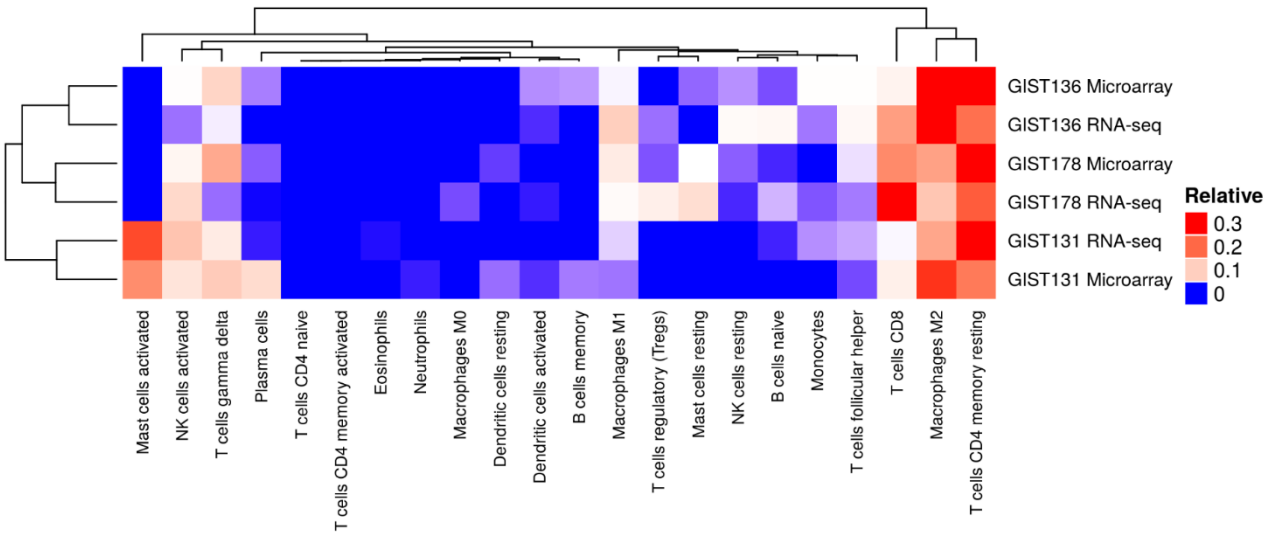
T12	GOLGA4	c.A5542T	p.T1848S
T12	LIMA1	c.G491A	p.C164Y
T12	NETO2	c.C148T	p.R50X
T12	NUBP2	c.G292A	p.V98M
T12	PDGFRA	c.A2525T	p.D842V
T12	SHISA9	c.G1136A	p.R379Q
T12	TNFRSF10B	c.G878A	p.R293K
T12	WDR36	c.T1982G	p.L661W
T12	ZNF679	c.G70A	p.E24K
T13	BCOR	c.C976A	p.P326T
T13	CTNND2	c.A3632G	p.N1211S
T13	CWF19L2	c.G2000T	p.S667I
T13	FAM122B	c.217delG	p.E73fs
T13	INO80D	c.C2473T	p.P825S
T13	INSR	c.T2947G	p.S983A
T13	KEL	c.C1262T	p.T421M
T13	KRTAP12-2	c.G340A	p.V114M
T13	MDN1	c.G1981A	p.E661K
T13	NLGN2	c.C1226T	p.T409I
T13	OR2A4	c.G719T	p.C240F
T13	OR4C12	c.G902T	p.R301I
T13	OR6B3	c.C766T	p.L256F
T13	PDGFRA	c.A2525T	p.D842V
T13	SBK1	c.C593T	p.T198M
T13	TACR3	c.G1304A	p.S435N
T13	TOPAZ1	c.A1703G	p.N568S
T13	TP53	c.T403G	p.C135G
T13	TP53	c.993+1G>A	nn
T13	ZNF197	c.T1145C	p.I382T
T14	EEF2	c.C1479G	p.N493K
T14	HNRNPA2B1	c.A106G	p.S36G
T14	PDGFRA	c.A2525T	p.D842V
T15	AGAP3	c.G2006C	p.R669P
T15	ARID1B	c.C1741T	p.P581S
T15	ASIC4	c.G1711A	p.D571N
T15	BICD2	c.C50A	p.A17E
T15	CACNG5	c.A773C	p.Y258S
T15	CCDC80	c.A1054C	p.T352P

T15	CT47B1	c.702_728del	p.234_243del
T15	DRD5	c.C825A	p.S275R
T15	ITGA3	c.C1582T	p.R528W
T15	MOV10	c.C1265A	p.P422H
T15	NUBP1	c.A620G	p.K207R
T15	PDGFRA	c.A2525T	p.D842V
T15	ZFHX3	c.7815_7816insCGGCGGCGG	p.G2605delinsGGGG
T16	ADAM19	c.A313T	p.T105S
T16	ADRBK2	c.G1976A	p.R659H
T16	ARID2	c.C2626T	p.Q876X
T16	BCAT1	c.A886G	p.I296V
T16	F8	c.5247delC	p.F1749fs
T16	GPR112	c.C6122T	p.T2041I
T16	IL2RG	c.G295A	p.V99I
T16	KCTD4	c.C98T	p.T33I
T16	KLK5	c.G280A	p.G94R
T16	PDGFRA	c.A2525T	p.D842V
T16	PIKFYVE	c.G3964T	p.V1322L
T16	RMND1	c.A676G	p.I226V
T16	SCUBE1	c.G1132A	p.V378I
T17	FAM73A	c.G68T	p.G23V
T17	FOXA3	c.A724C	p.T242P
T17	GALNT9	c.C661T	p.R221W
T17	GIMAP8	c.G1240A	p.E414K
T17	IQSEC3	c.G1399A	p.A467T
T17	MATR3	c.754delG	p.D252fs
T17	NFXL1	c.T2414C	p.I805T
T17	PDGFRA	c.A2525T	p.D842V
T17	PDP1	c.A944G	p.N315S
T17	PHIP	c.G2159T	p.S720I
T17	SCN11A	c.G1942T	p.V648F
T17	TMPRSS13	c.A233G	p.Q78R
T17	UCK2	c.C257T	p.P86L
T18	ARHGAP15	c.A653C	p.H218P
T18	CAPNS1	c.G554A	p.R185Q
T18	EDEM3	c.A1744G	p.M582V
T18	ENAH	c.651_668del	p.217_223del
T18	JAG2	c.C2720T	p.A907V

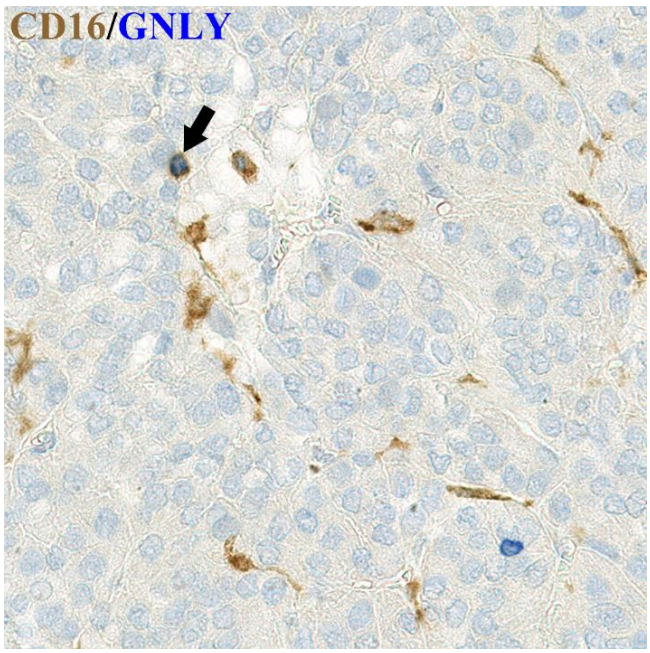
T18	KRTAP9-9	c.G422A	p.R141H
T18	LCE4A	c.129_130insAGCTCTGGGGGCTGCTGT	p.S43delinsSSSGGCC
T18	MGRN1	c.A761G	p.Y254C
T18	NIPA2	c.G1006C	p.E336Q
T18	NT5C2	c.C1334T	p.S445F
T18	NXF3	c.G1547T	p.S516I
T18	PDGFRA	c.A2525T	p.D842V
T18	PLCB1	c.C1031T	p.S344F
T18	PLXND1	c.4994-1G>C	nn
T18	RNF165	c.G336T	p.Q112H
T18	ZADH2	c.T1063G	p.Y355D
T19	C11orf49	c.C217A	p.H73N
T19	CACNG2	c.G484A	p.G162R
T19	CD163	c.G536A	p.R179Q
T19	EHBP1	c.G1829A	p.R610H
T19	FAM46A	c.102_131del	p.34_44del
T19	IMP3	c.A460G	p.M154V
T19	KCNN3	c.242_243insAGCAGC	p.P81delinsQQP
T19	KIAA1211	c.A2132T	p.K711M
T19	PDGFRA	c.A2525T	p.D842V
T19	PTPRU	c.C1897T	p.R633W
T19	TGFB1I1	c.C530T	p.S177F
T19	USP19	c.G1108T	p.A370S



***Supplementary Figure 1**. Cibersort absolute score stratified by GIST molecular subtype (**A**) and by platform type (**B**). Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*



***Supplementary Figure 2**. Clustering of Cibersort immune cell population deconvolution done on three samples analysed by both microarray and RNAseq techniques. Figure presented in the work of Pantaleo et al. OncoImmunology, 8:9, DOI: 10.1080/2162402X.2019.1617588*



***Supplementary Figure 3**. IHC. Picture of the double immunostaining performed by using the following monoclonal antibodies: anti-granulysin (blue) (GNLY) and anti-CD16 (brown). The GNLY+/CD16+ NK cell was indicated by the black arrow (x400).*

REFERENCES

1. Corless C.L., Barnett C.M., Heinrich M.C. Gastrointestinal stromal tumours: Origin and molecular oncology. Nat. Rev. Cancer. 2011;11:865–878. doi: 10.1038/nrc3143.
2. Biasco G., Velo D., Angriman I., Astorino M., Baldan A., Baseggio M., Basso U., Battaglia G., Bertin M., Bertorelle R., et al. Gastrointestinal stromal tumors: Report of an audit and review of the literature. Eur. J. Cancer Prev. 2009;18:106–116. doi: 10.1097/CEJ.0b013e32830c8da8.
3. Heinrich M.C., Corless C.L., Duensing A., McGreevey L., Chen C.J., Joseph N., Singer S., Griffith D.J., Haley A., Town A., et al. PDGFRA activating mutations in gastrointestinal stromal tumors. Science. 2003;299:708–710. doi: 10.1126/science.1079666.
4. Heinrich M.C., Corless C.L., Demetri G.D., Blanke C.D., von Mehren M., Joensuu H., McGreevey L.S., Chen C.J., Van den Abbeele A.D., Druker B.J., et al. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumors. J. Clin. Oncol. 2003;21:4342–4349. doi: 10.1200/JCO.2003.04.190.
5. Heinrich M.C., Owzar K., Corless C.L., Hollis D., Borden E.C., Fletcher C.D., Ryan C.W., Von Mehren M., Blanke C.D., Rankin C., et al. Correlation of kinase genotype and clinical outcome in the North American Intergroup Phase III Trial of imatinib mesylate for treatment of advanced gastrointestinal stromal tumor: CALGB 150105 Study by Cancer and Leukemia Group B and Southwest Oncology Group. J. Clin. Oncol. 2008;26:5360–5367.
6. Debiec-Rychter M., Dumez H., Judson I., Wasag B., Verweij J., Brown M., Dimitrijevic S., Sciot R., Stul M., Vranck H., et al. Use of c-KIT/PDGFRA mutational analysis to predict the clinical response to imatinib in patients with advanced gastrointestinal stromal tumours entered on phase I and II studies of the EORTC Soft Tissue and Bone Sarcoma Group. Eur. J. Cancer. 2004;40:689–695. doi: 10.1016/j.ejca.2003.11.025.
7. Gastrointestinal Stromal Tumor Meta-Analysis Group (MetaGIST) Comparison of two doses of imatinib for the treatment of unresectable or metastatic gastrointestinal stromal tumors: A meta-analysis of 1640 patients. J. Clin. Oncol. 2010;28:1247–1253.
8. Heinrich M.C., Maki R.G., Corless C.L., Antonescu C.R., Harlow A., Griffith D., Town A., McKinley A., Ou W.B., Fletcher J.A., et al. Primary and secondary kinase genotypes correlate with the biological and clinical activity of sunitinib in imatinib-resistant gastrointestinal stromal tumor. J. Clin. Oncol. 2008;26:5352–5359. doi: 10.1200/JCO.2007.15.7461.
9. Debiec-Rychter M., Sciot R., Le C.A., Schlemmer M., Hohenberger P., Van Oosterom A.T., Blay J.Y., Leyvraz S., Stul M., Casali P.G., et al. KIT mutations and dose selection for

imatinib in patients with advanced gastrointestinal stromal tumours. Eur. J. Cancer. 2006;42:1093–1103. doi: 10.1016/j.ejca.2006.01.030.

10. Heinrich M, Jones R, Von Mehren M, Schöffski P, Mir O, Cassier PA, Eskens F, Shi H, Alvarez Diez T, Schmidt-Kittler O, et al. Clinical activity of BLU-285, a highly potent and selective KIT/PDGFR α inhibitor designed to treat gastrointestinal stromal tumor (GIST). J Clin Oncol. 2017;35:11011. doi:10.1200/JCO.2017.35.15_suppl.11011
11. Janku F, Razak ARA, Gordon MS, Flynn D, Kaufman M, Pitman J, Smith B, Somaiah N, Jennings J, Salah S, et al. Encouraging activity of novel pan-KIT and PDGFRα inhibitor DCC-2618 in patients (pts) with Gastrointestinal Stromal Tumor (GIST). Ann Oncol. 2017;28:v521–v538. doi:10.1093/annonc/mdx387
12. Debiec-Rychter M, Wasag B, Stul M, et al. Gastrointestinal stromal tumours (GISTs) negative for KIT (CD117 antigen) immunoreactivity. J Pathol 2004;202:430-8.
13. Miettinen M, Wang ZF, Lasota J. DOG1 antibody in the differential diagnosis of gastrointestinal stromal tumors: a study of 1840 cases. Am J Surg Pathol 2009;33:1401-8.
14. Rubin BP, Heinrich MC, Corless CL. Gastrointestinal stromal tumour. Lancet. 2007 May 19;369(9574):1731-41.
15. Huang E, Nocka K, Beier DR, et al. The hematopoietic growth factor KL is encoded by the Sl locus and is the ligand of the c-kit receptor, the gene product of the W locus. Cell 1990;63:225-33.
16. Duensing A, Medeiros F, McConarty B, et al. Mechanisms of oncogenic KIT signal transduction in primary gastrointestinal stromal tumors (GISTs). Oncogene 2004;23:3999-4006.
17. Ernst SI, Hubbs AE, Przygodzki RM, Emory TS, Sobin LH, O'Leary TJ. KIT mutation portends poor prognosis in gastrointestinal stromal/smooth muscle tumors. Lab Invest 1998;78:1633-6.
18. Singer S, Rubin BP, Lux ML, et al. Prognostic value of KIT mutation type, mitotic activity, and histologic subtype in gastrointestinal stromal tumors. J Clin Oncol 2002;20:3898-905.
19. Taniguchi M, Nishida T, Hirota S, et al. Effect of c-kit mutation on prognosis of gastrointestinal stromal tumors. Cancer Res 1999;59:4297-300.
20. Andersson J, Bümming P, Meis-Kindblom JM, et al. Gastrointestinal stromal tumors with KIT exon 11 deletions are associated with poor prognosis. Gastroenterology 2006;130:1573-81.
21. Martín J, Poveda A, Llombart-Bosch A, , et al. Deletions affecting codons 557–558 of the c-KIT gene indicate a poor prognosis in patients with completely resected gastrointestinal stromal tumors: a study by the Spanish Group for Sarcoma Research (GEIS). J Clin Oncol 2005;23:6190-8.

22. Wasag B, Debiec-Rychter M, Pauwels P, et al Differential expression of KIT/PDGFRA mutant isoforms in epithelioid and mixed variants of gastrointestinal stromal tumors depends predominantly on the tumor site. Mod Pathol 2004;17:889-94.
23. Lasota J, Dansonka-Mieszkowska A, Sobin LH, Miettinen M. A great majority of GISTs with PDGFRA mutations represent gastric tumors of low or no malignant potential. Lab Invest 2004;84:874-83.
24. Manley PW, Cowan-Jacob SW, Buchdunger E, et al. imatinib: a selective tyrosine kinase inhibitor. Eur J Cancer 2002;38(Suppl 5):S19-27.
25. Corless CL, Barnett CM, Heinrich MC. 2011. Gastrointestinal stromal tumours: origin and molecular oncology. Nat Rev Cancer 11(12):865-78.
26. Heinrich MC, Corless CL, Demetri GD, et al. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumors. J Clin Oncol 2003;21:4342-4349.
27. Gramza AW, Corless CL, Heinrich MC. Resistance to Tyrosine Kinase Inhibitors in Gastrointestinal Stromal Tumors. Clin Cancer Res 2009;15:7510-7518
28. Wang WL, Conley A, Reynoso D, Nolden L, Lazar AJ, George S, Trent JC. 2011b. Mechanisms of resistance to imatinib and sunitinib in gastrointestinal stromal tumor. Cancer Chemother Pharmacol 67 Suppl 1:S15-24
29. Demetri GD, Reichardt P, Kang YK, Blay JY, Rutkowski P, Gelderblom H, Hohenberger P, Leahy M, von Mehren M, Joensuu H, Badalamenti G, Blackstein M, Le Cesne A, Schöffski P, Maki RG, Bauer S, Nguyen BB, Xu J, Nishida T, Chung J, Kappeler C, Kuss I, Laurent D, Casali PG; GRID study investigators. Efficacy and safety of regorafenib for advanced gastrointestinal stromal tumours after failure of imatinib and sunitinib (GRID): an international, multicentre, randomised, placebo-controlled, phase 3 trial. Lancet. 2013 Jan 26;381(9863):295-302.
30. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010 Jan;11(1):31-46.
31. Fischer, E. (1894), Einfluss der Configuration auf die Wirkung der Enzyme. Ber. Dtsch. Chem. Ges., 27: 2985-2993. doi:10.1002/cber.18940270364
32. Koshland DE Jr. Correlation of Structure and Function in Enzyme Action. Science. 1963; 142:1533–1541
33. Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des. 2011 Jun;7(2):146-57.
34. van Dongen M, Savage NDL, Jordanova ES, Briaire-de Bruijn IH, Walburg KV, Ottenhoff THM, Hogendoorn PCW, van der Burg SH, Gelderblom H, van Hall T. Anti-inflammatory M2 type macrophages characterize metastasized and tyrosine kinase inhibitor-treated gastrointestinal stromal tumors. Int J Cancer. 2010;127:NA–NA. doi:10.1002/ijc.25113.

35. Tan Y, Trent JC, Wilky BA, Kerr DA, Rosenberg AE. Current status of immunotherapy for gastrointestinal stromal tumor. Cancer Gene Ther. 2017;24:130–133. doi:10.1038/cgt.2016.58.
36. Rusakiewicz S, Semeraro M, Sarabi M, Desbois M, Locher C, Mendez R, Vimond N, Concha A, Garrido F, Isambert N, et al. Immune infiltrates are prognostic factors in localized gastrointestinal stromal tumors. Cancer Res. 2013;73:3499–3510. doi:10.1158/0008-5472.CAN-13-0371.
37. D'Angelo SP, Shoushtari AN, Agaram NP, Kuk D, Qin L-X, Carvajal RD, Dickson MA, Gounder M, Keohan ML, Schwartz GK, et al. Prevalence of tumor-infiltrating lymphocytes and PD-L1 expression in the soft tissue sarcoma microenvironment. Hum Pathol. 2015;46:357–365. doi:10.1016/j.humpath.2014.11.001.
38. Balachandran VP, Cavnar MJ, Zeng S, Bamboat ZM, Ocuin LM, Obaid H, Sorenson EC, Popow R, Ariyan C, Rossi F, et al. Imatinib potentiates antitumor T cell responses in gastrointestinal stromal tumor through the inhibition of Ido. Nat Med. 2011;17:1094–1100. doi:10.1038/nm.2438.
39. Seifert AM, Zeng S, Zhang JQ, Kim TS, Cohen NA, Beckman MJ, Medina BD, Maltbaek JH, Loo JK, Crawley MH, et al. PD-1/PD-L1 blockade enhances T-cell activity and antitumor efficacy of imatinib in gastrointestinal stromal tumors. Clin Cancer Res. 2017;23:454–465. doi:10.1158/1078-0432.CCR-16-1163.
40. D'Angelo SP, Shoushtari AN, Keohan ML, Dickson MA, Gounder MM, Chi P, Loo JK, Gaffney L, Schneider L, Patel Z, et al. Combined KIT and CTLA-4 blockade in patients with refractory GIST and other advanced sarcomas: A phase Ib study of dasatinib plus ipilimumab. Clin Cancer Res. 2017;23:2972–2980. doi:10.1158/1078-0432.CCR-16-2349.
41. Lindgreen, S. AdapterRemoval: Easy cleaning of next-generation sequencing reads. BMC Res. Notes 2012, 5, 337.
42. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics 2009, 25, 2078–2079.
43. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010, 20, 1297–1303.
44. Cibulskis, K.; Lawrence, M.S.; Carter, S.L.; Sivachenko, A.; Jaffe, D.; Sougnez, C.; Gabriel, S.; Meyerson, M.; Lander, E.S.; Getz, G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. 2013, 31, 213–219.

45. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010, 38, e164.
46. Cingolani, P.; Platts, A.; Wang le, L.; Coon, M.; Nguyen, T.; Wang, L.; Land, S.J.; Lu, X.; Ruden, D.M. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012, 6, 80–92.
47. Boeva, V.; Popova, T.; Bleakley, K.; Chiche, P.; Cappo, J.; Schleiermacher, G.; Janoueix-Lerosey, I.; Delattre, O.; Barillot, E. Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics 2012, 28, 423–425.
48. Amarasinghe, K.C.; Li, J.; Hunter, S.M.; Ryland, G.L.; Cowin, P.A.; Campbell, I.G.; Halgamuge, S.K. Inferring copy number and genotype in tumour exome data. BMC Genom. 2014, 15, 732.
49. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. Curr. Protocol Bioinform. 2016.
50. Morris, G.M.; Huey, R.; Lindstrom,W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDockTools4: Automated docking with selective receptorflexibility. J. Comput. Chem. 2009, 30, 2785–2791.
51. Schüttelkopf, A.W.; van Aalten, D.M. PRODRG: A tool for high-throughputcrystallography of protein-ligand complexes. Acta Crystallogr. D Biol. Crystallogr. 2004, 60 Pt 8, 1355–1363.
52. Nannini M, Astolfi A, Urbini M, Indio V, Santini D, Heinrich MC, Corless CL, Ceccarelli C, Saponara M, Mandrioli A, et al. Integrated genomic study of quadruple-WT GIST (KIT/PDGFRA/SDH/RAS pathway wild-type GIST). BMC Cancer. 2014;14:685. doi:10.1186/1471-2407-14-685.
53. Newman AM, Gentles AJ, Liu CL, Diehn M, Alizadeh AA. Data normalization considerations for digital tumor dissection. Genome Biol. 2017;18:128. doi:10.1186/s13059-017-1257-4.
54. Gentles AJ, Newman AM, Liu CL, S V B, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015;21:938–945. doi:10.1038/nm.3909.
55. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, Albright A, Cheng JD, Kang SP, Shankaran V, et al. IFN-γ–related mRNA profile predicts clinical response to PD-1 blockade. J Clin Invest. 2017;127:2930–2940. doi:10.1172/JCI91190.
56. Danaher P, Warren S, Lu R, Samayoa J, Sullivan A, Pekker I, Wallden B, Marincola FM, Cesano A. Pan-cancer adaptive immune resistance as defined by the Tumor Inflammation Signature (TIS): results from The Cancer Genome Atlas (TCGA). J Immunother Cancer. 2018;6:63. doi:10.1186/s40425-018-0367-1.

57. Rink L, Skorobogatko Y, Kossenkov AV, Belinsky MG, Pajak T, Heinrich MC, Blanke CD, von Mehren M, Ochs MF, Eisenberg B, et al. Gene expression signatures and response to imatinib mesylate in gastrointestinal stromal tumor. Mol Cancer Ther. 2009;8:2172–2182. doi:10.1158/1535-7163.MCT-09-0193.
58. Garcia-Diaz A, Shin DS, Moreno BH, Saco J, Escuin-Ordinas H, Rodriguez GA, Zaretsky JM, Sun L, Hugo W, Wang X, et al. Interferon receptor signaling pathways regulating PD-L1 and PD-L2 Expression. Cell Rep. 2017;19:1189–1201. doi:10.1016/j.celrep.2017.04.031.
59. Wang, Y.; Marino-Enriquez, A.; Bennett, R.R.; Zhu, M.; Shen, Y.; Eilers, G.; Lee, J.C.; Henze, J.; Fletcher, B.S.; Gu, Z.; et al. Dystrophin is a tumor suppressor in human cancers with myogenic programs. Nat. Genet. 2014, 46, 601–606.
60. Liang, L.; Yan, X.E.; Yin, Y.; Yun, C.H. Structural and biochemical studies of the PDGFRA kinase domain. Biochem. Biophys. Res. Commun. 2016, 477, 667–672.
61. Roskoski, R., Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. Pharmacol. Res. 2016, 103, 26–48.
62. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, Sher X, Liu XQ, Lu H, Nebozhyn M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science (80-). 2018;362:eaar3593. doi:10.1126/science.aar3593.
63. Heinrich, M.C.; Griffith, D.; McKinley, A.; Patterson, J.; Presnell, A.; Ramachandran, A.; Debiec-Rychter, M. Crenolanib inhibits the drug-resistant PDGFRA D842V mutation associated with imatinib–resistant gastrointestinal stromal tumors. Clin. Cancer Res. 2012, 18, 4375–4384.
64. Heinrich, M.C.; Jones, R.L.; von Mehren, M.; Schoffski, P.; Bauer, S.; Mir, O.; Cassier, P.A.; Eskens, F.; Shi, H.; Alvarez-Diez, T.; et al. Clinical activity of BLU-285 in advanced gastrointestinal stromal tumor (GIST). J. Clin. Oncol. 2017, 35, 11011.

Acknowledgements

First and foremost I would like to thank my supervisor, the professor Maria Abbondanza Pantaleo. Getting a PhD is a rather long process and I have appreciated her guidance throughout this endeavor. I am especially grateful for the way in which she has always gone through my work. Not every graduate student receives such careful appraisal and guidance during this extensive voyage. Thanks to prof. Guido Biasco for having encouraged me to make my contribution in the field of genomics and transcriptomics of tumors and for introducing me to Professor Pantaleo. Essentially all of the research contained in this thesis was carried out at the Interdipartimental centre for cancer research "Giorgio Prodi" (CIRC) in Bologna. It has been a pleasure to work with such a talented group of people over the last 4 years. In particular, I would like to thanks Prof. Andrea Pession head of the CIRC, for always supporting my ideas. Thanks to Valentina Indio and Annalisa Astolfi that spent much of their time training me.

Special thanks also to my parents Adele Di Prima and Santi Tarantino and to my brother and sisters for allowing me study outside of Palermo.

My research has benefited from interaction with many other colleagues, both at CIRC and other research centers, in particular Milena Urbini, Gloria ravegnini, Daria Messelodi, Margherita Nannini & Milly Bwengue. I would also like to thank the administrative staff of the CIRC, Fiorella Gherenzè and Ciro Danzo, for their amazing ability to find solutions to all manner of issues. They made coming to work at CIRC and living in Bologna a breeze.

Thanks to my friends: Edoardo De Caro, Fabrizio Glorioso, Riccardo Privolizzi, Guglielmo Puccio, Amin Boroomand, Emanuela Garbo, Silvia Compagno, Stefano Pascarelli, Maria Luisa Mauro, Emanuele Feleppa, Cinzia Barcellona, Martina Zambito, Francesca Aglialoro, Giorgio Cucuzza, Claudio Mannino, Giuliana Mannino and all my teammates from GRANAMICA calcio, for helping me to live far from Palermo and for always making me feel loved.

Finally, I must express my profound gratitude to my other half Anna Maria Giudice, who always kept my confidence up even when I was under pressure and who continues to support me in all my endeavors, no matter how grand, silly or nerdy they may be.

Part of this research was funded by the grant "Corri con Gio" from the Giovanna Fatato association that I want to thank.

Bologna, Itaia, Febbraio 2020