UNIVERSITÀ DEGLI STUDI DI GENOVA

DIPARTIMENTO DI SCIENZE DELLA TERRA,

DELL'AMBIENTE E DELLA VITA (DISTAV)

Classe delle Lauree Magistrali in Biologia LM/6

Corso di Laurea Magistrale in Biologia Applicata e Sperimentale:

curriculum biosanitario

Glycine-tRNA synthetase 1 (GARS1) gene expression-based reporter system as a new platform for drug repositioning in non-small cell lung cancer (NSCLC) cells

Laureando:

Matteo Raineri

Relatore:

Prof.ssa Tiziana Bachetti

Prof.ssa Elena Grasselli

Correlatore:

Prof.ssa Chiara Cervetto

ANNO ACCADEMICO 2024-2025

RIASSUNTO

In questo lavoro è stato iniziato un progetto di riposizionamento di farmaci ("drug repositioning") volto all'identificazione di composti già validati e confermati avere determinati bersagli molecolari, che possano aver un effetto su sull'espressione del gene *GARS1*; *GARS1* che codifica per la proteina GlyRS1 (glicil-tRNA sintetasi) che nelle cellule ha il compito di caricare lo specifico tRNA con l'amminoacido glicina per svolgere il suo ruolo nella sintesi proteica.

Il razionale del lavoro si basa su numerose osservazioni sul ruolo della famiglia degli ARS in differenti patologie, tra cui il contesto tumorale e più nel dettaglio del ruolo dell'over-espressione del trascritto (mRNA) di *GARS1* nella progressione della malattia, nella maggiore evasione dal tessuto d'origine del tumore primario ed infine nella minore aspettativa di vita nei pazienti con alti livelli di espressione del gene *GARS1* a livello tumorale. Ciò sposta l'attenzione su tutte quelle funzioni, ancora parzialmente in fase di studio, che può svolgere *GARS1* all'interno dei differenti tipi cellulari. In particolar modo è stato osservato che all'interno di specifiche linee tumorali *GARS1* potrebbe essere un promettente bersaglio molecolare per il trattamento della malattia, come ad esempio nell'adenocarcinoma polmonare non a piccole cellule (NSCLC), dove GARS1 è over-espresso e prognosticamene sfavorevole; a conferma dell'effetto patogenetico di GARS1 in NSCLC, nel nostro laboratorio è stato analizzato il nuovo farmaco sperimentale Fraisinib, che fa parte della classe dei calixpirroli, il quale si è dimostrato avere un'azione diretta sulla proteina codifica da *GARS1*, andandola a bloccare stericamente. Analisi di proteomica differenziale a seguito del trattamento nella linea A549 di NSCLC ha mostrato molteplici proteine differentemente espresse (DEPs), la cui analisi ha evidenziato il loro coinvolgimento nella riduzione della proliferazione tumorale, la diminuita capacità di invasione ed infine una maggiore propensione all'apoptosi delle cellule tumorale, dimostrando come GARS1 sia coinvolto nella progressione tumorale e nella resistenza alla morte cellulare programmata.

Utilizzando la linea cellulare umana di NSCLC A549, per svolgere il "drug repositioning" sono stati effettuati due approcci:

1. Generazione in vitro di una piattaforma utilizzabile per High Throughput Screening (HTS) basata sulla regolazione della trascrizione di *GARS1* rilevabile con un saggio enzimatico luminescente;

2. Analisi di "drug repositioning" *in silico* basato sull'analisi delle DEPs derivate dal trattamento con Fraisinib

Nell'approccio sperimentale è stato creato un sistema di rilevazione ad alta efficienza capace di testare un gran numero di farmaci in un breve lasso di tempo, ma per prima cosa è stato necessario creare dei cloni stabili, usando le cellule A549 come modello di NSCLC, trasfettati con un plasmide reporter in cui il promotore del gene *GARS1* regola l'espressione del gene che codifica per l'enzima Luciferasi di lucciola. Così facendo è stato creato un sistema di rilevazione dell'espressione del gene bersaglio basato sulla differente emissione di luce da parte della luciferasi, indice dell'attività trascrizionale di *GARS1* e quindi della sua espressione.

Dopo aver selezionato il clone stabile migliore in base ad omogeneità di espressione del gene *luciferasi*, è stato possibile analizzare su di esso la libreria di 43 farmaci noti per avere attività epigenetica. Per normalizzare l'effetto dei differenti farmaci è stato necessario rilevare sia l'emissione della Luciferasi, grazie ad un luminometro, sia l'emissione dell'AFC, un composto che permette la quantificazione della vitalità cellulare; questo perché era necessario escludere chela diminuzione di emissione di luciferasi nel campione fosse determinata da diminuzione del numero di cellule a seguito del trattamento e quindi una finale riduzione nell'emissione. Dopo aver confrontato i dati ottenuti, a 24 ore dal trattamento, sono stati osservati in particolare 10 farmaci in grado di diminuire la trascrizione di GARS1: butirolactone 3, B2, CTPB, acido anacardico, 6-cloro-2,3,4,9-tetraidro-1H-carbazole-1-carboxamide, sirtinolo, salermide, MC-1293 e sodio fenilbutirrato. L'analisi statistica del triplicato biologico ha mostrato significativamente efficaci butirolactone 3, CTPB, l'acido anacardico MC-1293, B2; a quest'analisi sono stati aggiunti anche farmaci statisticamente quasi significativi, in particolare: sodio fenilbutirrato, acido valproico, 6-cloro-2,3,4,9-tetraidro-1H-carbazole-1-carboxamide, salermide; la scelta della loro inclusione è stata basata sul fatto che nei tre esperimenti avevano mostrato un effetto marcato sulla trascrizione di *GARS1* e la variabilità dei risultati, che non aveva permesso di raggiungere la significatività statistica, molto probabilmente si sarebbe attenuata aumentando la numerosità dei replicati. Successivamente è stato valutato se tutti questi farmaci potessero indurre una diminuzione della vitalità cellulare. L'analisi di vitalità ha mostrato una sostanziale riduzione principalmente alle 72 ore per la maggior parte dei farmaci, mentre alle 48 ore i dati si sono rilevati più variabili. Questi risultati aprono comunque le porte a studi futuri di validazione e maggior comprensione sui farmaci analizzati; inoltre, questi farmaci efficaci nel diminuire la trascrizione di GARS1 saranno utilizzati come controlli positivi nel calcolo dello Z-factor in futuri HTS in cui saranno utilizzate librerie di farmaci più ampie.

Parallelamente all'approccio sperimentale è stata effettuata un'analisi informatica di "drug repositioning", sfruttando i dati di espressione genica in termini di lista di proteine differentemente

espresse (DEPs), risultanti dal trattamento delle A549 con il Fraisinib, utilizzando iLINCS, una piattaforma integrativa online per l'analisi di dati omici. La piattaforma facilita l'esplorazione e la rianalisi di una vasta collezione di dataset omici, firme molecolari già anlazzate e delle loro connessioni, oltre all'analisi di dati omici caricati dagli utenti stessi. Le interfacce di iLINCS consentono l'esecuzione di analisi complesse delle firme omiche, l'analisi dei meccanismi d'azione e il drug repositioning basato sulle firme molecolari. L'utilizzo di questa piattaforma era necessario per identificare nei dataset di iLINCS altri composti, già validati, che avessero un effetto simile sulla cellula. Da questa analisi sono stati osservati migliaia di composti con bersagli molecolari differenti da GARS1 ma che risultavano in una simile alterazione dell'espressione genica. Tra questi sono stati selezionati solo quelli che erano stati testati sulle A549, alla concentrazione di 10 μM, che fossero significativi in base al p-value<0.05.

Inoltre, è stata effettuata la ricerca di putativi fattori di trascrizione che potrebbero regolare *GARS1*; a questo scopo è stato utilizzato Promo3, un programma che permette di predire i possibili fattori di trascrizione che possono avere un effetto su una determinata sequenza nucleotidica, nel caso specifico GARS1. I fattori di trascrizione risultanti sono stati usati per ricercare tra i dati prodotti da iLINCS dei composti che presentavano come bersagli confermati uno o più di queste molecole. Da questa analisi incrociata sono emersi 22 composti. Due di questi hanno suscitato un maggiore interesse per le loro caratteristiche: Treitonina, il cui principio attivo è l'acido retinoico, noto regolatore dell'espressione genica, e Quinacrina che ha come bersaglio molecolare p53, un gene ben noto in ambito tumorale grazie alla sua azione protettiva verso i danni al DNA ed in generale alla progressione tumorale, ma la cui azione diretta su GARS1 non era ancora nota. L'analisi dell'effetto dell'acido retinoico sulla trascrizione di *GARS1* mediante il saggio luciferasico non ha portato a dati significativi, probabilmente a causa di una minor azione e maggiore tossicità del principio attivo rispetto alla formulazione Trietonina.

La validazione di un possibile effetto della Quinacrina mediante p53 è stata analizzata effettuando nelle A549 una co-trasfezione del promotore di *GARS1* e di un plasmide di espressione per p53, che ha mostrato una diminuzione della transattivazione del promotore di *GARS1* di circa il 30% con significatività statistica.

L'intero lavoro è stato svolto presso IRCCS ospedale policlinico San Martino, nella UOS Proteomica e Spettrometria di Massa del Centro Biotecnologie Avanzate (CBA).

ABSTRACT

This project initiated a drug repositioning study aimed at identifying already approved drugs, with validated molecular targets, that could affect the expression of the gene *GARS1*, which encodes the GlyRS1 protein (glycyl-tRNA synthetase). This enzyme is responsible for charging the specific tRNA with glycine to support its role in protein synthesis in cells.

The rationale of the study is based on several observations regarding the role of the ARS family (aminoacyl-tRNA synthetases) in various pathological conditions, including cancer. In particular, the overexpression of *GARS1* has been associated with disease progression, increased tumor evasion from the primary tissue of origin, and poor overall survival in patients with high *GARS1* expression in tumor tissues. This shifts the focus to the multiple, partially unexplored, functions that GARS1 may perform in various cell types. Specifically, it has been observed that *GARS1* may represent a promising molecular target in certain cancer types, such as non-small cell lung cancer (NSCLC), where it is overexpressed and associated with poor prognosis. Supporting the pathogenic role of GARS1 in NSCLC, our laboratory analysed the novel experimental drug Fraisinib, a member of the calixpyrrole family, which demonstrated direct binding to the *GARS1*-encoded protein, acting via steric inhibition. Differential proteomic analysis after Fraisinib treatment in the A549 NSCLC cell line revealed multiple differentially expressed proteins (DEPs), whose functional enrichment analysis indicated roles in reduced tumor proliferation, decreased invasive capacity, and increased propensity for programmed cell death (apoptosis), thus confirming the involvement of *GARS1* in cancer progression and resistance to apoptosis.

To perform drug repositioning, two complementary approaches were undertaken:

1. In vitro generation of a high-throughput screening (HTS) platform based on luminescent enzymatic assay linked to *GARS1* expression.
2. In silico drug repositioning analysis based on the DEPs resulting from Fraisinib treatment.

For the experimental approach, a high-efficiency detection system was developed to allow rapid screening of many drugs at the same time. This required the generation of stable A549 clones, NSCLC model, transfected with a reporter plasmid in which the *GARS1* promoter drives expression of the firefly luciferase gene. This setup enables quantification of *GARS1* transcriptional activity through luminescence emitted by luciferase. After selecting the most stable clone

based on uniform luciferase expression, the platform was used to test a library of 43 drugs with known epigenetic activity. To normalize the effects of different treatments, both luciferase luminescence, via luminometer, and cell viability, via AFC fluorescence, were measured. This dual readout ensured that any observed reduction in luciferase signal was not due to cytotoxicity-related cell loss.

After 24 hours of treatment, 10 compounds were identified as effective in reducing *GARS1* transcription: butyrolactone 3, B2, CTPB, anacardic acid, 6-chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, sirtinol, salermide, MC-1293, and sodium phenylbutyrate. Statistical analysis highlighted significant efficacy for butyrolactone 3, CTPB, anacardic acid, MC-1293, and B2; this analysis also included drugs that were statistically nearly significant, specifically: sodium phenylbutyrate, valproic acid, 6-chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, and salermide. Their inclusion was based on the fact that in all three experiments they showed a marked effect on *GARS1* transcription, and the variability in the results, which prevented statistical significance from being reached, would very likely have decreased with a higher number of replicates. Subsequent analysis evaluated whether these compounds could also reduce cell viability. Additional drugs that showed near-significant transcriptional inhibition were included, such as sodium phenylbutyrate, valproic acid, 6-chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, and salermide. The viability assay showed a substantial reduction in cell survival primarily after 72 hours, while 48-hour data were more variable. These findings pave the way for further validation and mechanistic studies of the compounds identified. Moreover, drugs showing strong *GARS1* transcriptional inhibition will serve as positive controls in future HTS campaigns using broader drug libraries, with Z-factor calculations to assess assay quality.

Parallel to the experimental approach, a computational drug repositioning strategy was conducted using gene expression data (DEPs) from A549 cells treated with Fraisinib, processed through iLINCS, an integrative online platform for omics data analysis. iLINCS allows exploration and re-analysis of large omics datasets, precomputed molecular signatures, and their interconnections. The platform also supports user-uploaded datasets and facilitates complex signature analysis, mechanism of action investigation, and drug repositioning based on molecular profiles.

This analysis aimed to identify additional validated compounds from iLINCS datasets that elicited gene expression changes like those induced by Fraisinib. Thousands of such compounds

were identified, with different molecular targets from *GARS1*, yet producing similar transcriptomic alterations. Among these, only compounds tested in A549 cells at 10 μM and statistically significant (based on p-value) were retained for further evaluation.

Additionally, putative transcription factors regulating *GARS1* were investigated using Promo3, a bioinformatic tool predicting transcription factor binding sites for specific nucleotide sequences. The predicted transcription factors were used to query iLINCS datasets for compounds known to target these factors. This cross-analysis yielded 22 candidate compounds.

Two compounds of particular interest emerged: Tretinoin (active ingredient: retinoic acid) and Quinacrine, which targets p53, a well-characterized tumor suppressor involved in DNA damage response and tumor progression, though not previously associated with direct regulation of *GARS1*. Subsequently, retinoic acid was tested directly in place of Tretinoin, but at equivalent concentrations it showed lower efficacy and higher cytotoxicity, leading to non-significant results. Conversely, co-transfection of *GARS1* and p53 in A549 cells produced promising results, leading to a ~30% reduction in cell viability with statistical significance.

This entire project was conducted at IRCCS Ospedale Policlinico San Martino, within the Proteomics and Mass Spectrometry Laboratory of the Advanced Biotechnology Center (CBA).

1 INTRODUCTION

1.1 Glycyl-tRNA synthetase (GlyRS)

Glycyl-tRNA synthetase (GlyRS) (**Figure 1**) belongs to the ARS family, a universally conserved enzyme family essential for protein synthesis in all extant organisms. They catalyse the attachment of a specific amino acid to its cognate tRNA, thereby generating aminoacyl-tRNA that participates in translation, by delivering the correct amino acid to the ribosome according to the mRNA codon-anticodon pairing.

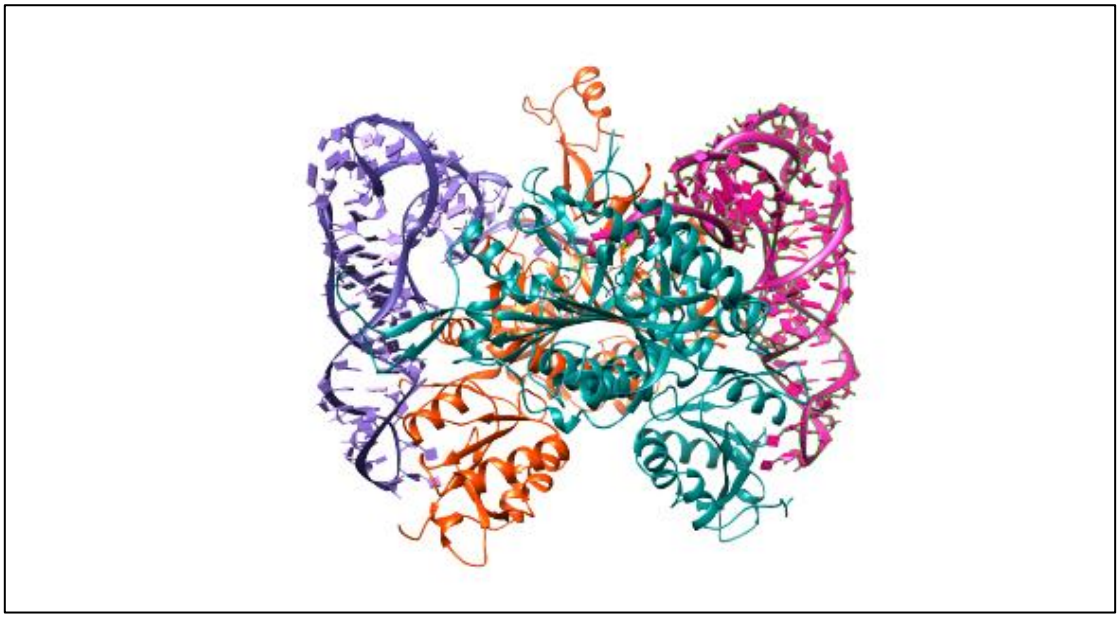

Figure 1. The GARS1 homodimer, the mitochondrial isoform, tridimensional structure.

ARS mediates a two-step enzymatic reaction where at the beginning there is the activation of the amino acid, forming aminoacyl-AMP through the reaction of the amino acid with ATP, and then the transfer of the activated amino acid to the 3' end of its cognate tRNA, producing aminoacyl-tRNA and releasing AMP. Among them, the glycyl-tRNA synthetase (GlyRS), encoded by the gene *GARS1*, plays a critical role in protein biosynthesis by catalysing the attachment of the aminoacyl glycine to its specific tRNA. (**Figure 2**).

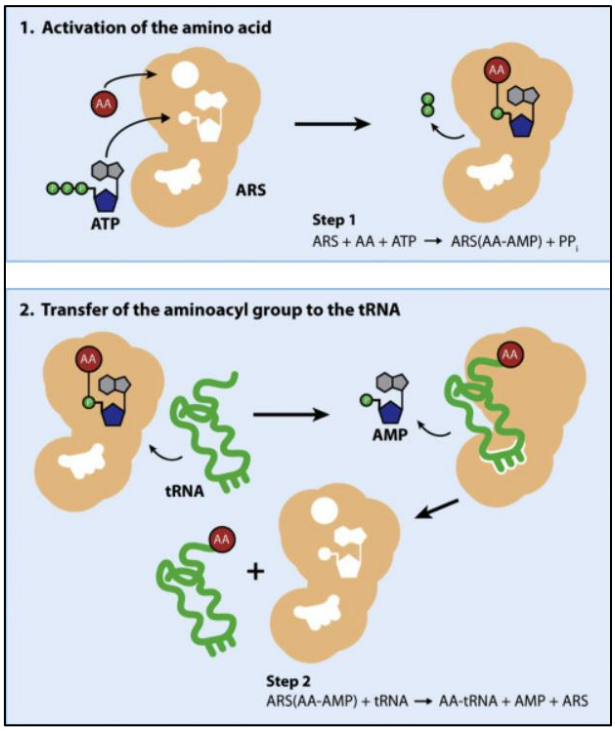

Figure 2.1. ***The activation of the amino acid*** *through the formation of aminoacyl-AMP with the consume of ATP; 2.2* ***The transfer of the aminoacyl group to the tRNA.*** *The amino acid is transferred to the 3' end of the tRNA, producing aminoacyl-tRNA, with AMP releasing, that could be used from the ribosome in the protein synthesis.*

This class of enzyme exhibits high substrate specificity for both the amino and the tRNA which guarantee an accurate decoding of the genetic code and to enhance translation fidelity certain ARS also possess an editing domain, that hydrolyses incorrectly charged tRNA, preventing incorporation of incorrect amino acids into the growing polypeptide chain. The number of ARS identified in contemporary organisms are 24, classified into two groups, each having three subclasses, from a to c, based on similiraty in sequences and structures. The classification is: class 1a (MetRS, ValRS, LeuRS, IleRS, CysRS, and ArgRS), class 1b (GluRS, GlnRS and LysRS-class I), class 1c (TyrRS and TrpRS), class 2a (SerRS, ThrRS, AlaRS, GlyRS-1, ProRS, and HisRS), class 2b (AspRS, AsnRS, and LysRS-class II), class 2c (PheRS, GlyRS-2, SepRS, and PylRS). The structural general constitution of ARS is based on a catalytic domain, with class-specific characteristic motif and topology, an anticodon-binding domain and often an editing domain (R. Furukawa et al., 2022).

In humans aminoacylation occurs in three different locations: in the nucleus, to charge the tRNA produced by nucleus with proper amino acid, cytoplasm where the charged tRNA-amino acid complex encounters the ribosome to protein synthesis, in addition the protein synthesis could occur in the mitochondria so to facilitate protein synthesis, the nuclear-encoded AARS is transported to mitochondria for translation utilizing specific localization signal of the AARS proteins.

Based on this location ARSs can be classified into cytoplasmatic ARS, indicated with the suffix -1, and mitochondrial ARS, with the suffix -2.

Based on high resolution crystal structure of ARSs, was observed that the class 1 family is characterized by HIGH and KMSKS motif in the active site of the Rossmann fold, a tertiary fold found in proteins that binds nucleotides, which is made up of alternate α-helices and β-sheet which participates in ATP binding in the extended conformation. This class of ARS aminoacylates at the 2'-OH of an adenosine nucleotide, but a few members are capable of aminoacylation of both 2' and 3' -OH. Instead, class 2 of ARS, like GlyRS, shares three degenerate motifs made of antiparallel β-sheet, where the first contain a proline residue important for homo-dimerization, while the second and the third contain arginine residues which have a role in ATP binding. This class of ARS aminoacylates only at the 3'-OH of an adenosine nucleotide.

1.1.1The GARS1 gene

The *GARS1* gene is located on human chromosome 7 (figure 3A) and encodes both the cytosolic and mitochondrial forms of the GARS protein via alternative transcription start sites, unlike other member of ARS family for which the two isoforms are produced by two separate gene, with only other exception for LARS (lysil-trna-synthetasis) (Alexandrova J. et al., 2015). The two isoforms of the GARS proteins are translated by different mRNAs with the same 3'-UTR sequence (166 nucleotides) but with a different length of their 5'-UTR, that presenting two alternative initiation codons (Mudge SJ. Et al.,1998). The longest *GARS1*-mRNA, with 2580 bp, which codifies for the shortest protein isoform, while the shortest GARS-mRNA with 2437 bp codifies mainly for the longest mitochondrial ones (Alexandrova J. et al., 2015) (**Figure 3B**). The promoter sequence, which drives the transcription of both the two isoforms, is composed of 2019 bp.

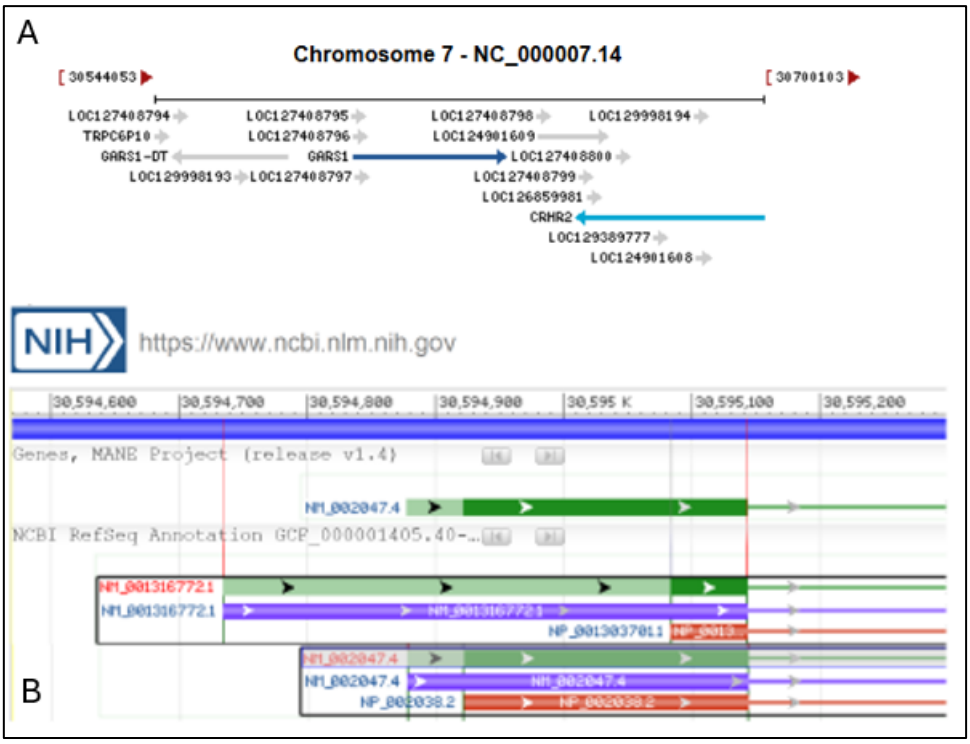

Figure 3. ***A.*** *The visualization of GARS1gene in human chromosome 7;* ***B.*** *The two isoform encoded by GARS1 gene, the longest transcript NM002047.4, the shorter mRNA variant, which codify for the longest protein, and NM 001316772.1, the longest mRNA variant, which codify for the shortest one.*

1.2 GARS1 implications in diseases

1.2.1Neurological diseases

Some ARSs mutations has been discovered driver for some pathological events, specifically in neurological disease, like Charcot-Marie-Tooth disease (CMT) were GlyRS plays a crucial role in the manifestation of the disease. CMT is one of the most common inherited neurological disorders that affect peripheral nerves and specific mutation of GlyRS lead to a sub-group called CMT disease type 2D which affect motor and sensory peripheral neuropathy, or to a distal spinal muscular atrophy type V. These mutations altered the normal structure of the protein, which gain a new function, different from the original role in protein synthesis (W. He, 2011).

1.2.2Cancer

In the last few years new evidence has shown that members of ARS family have a critical role in different pathological events, participating in sustaining proliferative signal and specifically in tumor-cells enhancing cancer cell inflammation pattern, cell aggressivity and metastasis. ARSs pathological role has been demonstrated through alteration in both transcriptional and translational processes, with different importance in many types of cancer (Sung et al.,2022).

Focusing on GARS1, it has been shown having opposite effects depending on the cellular context. The correct expression plays a protective role against cancer growth, throw its secretion by macrophages in the tumor microenvironment which was demonstrated having a pro-apoptotic effect on cancer cell. GlyRS protein is released by macrophages after a DNA damage signal or starvation signal, through recognition of Fas receptor on cancer cell surface, and activation of signal cascade resulting in ERK-activated cell arrest or apoptosis (Park MC. et al.,2022). In contrast in different cancer cell *GARS1* has been shown having a higher expression and an opposite role, becoming a safeguard in different cancer cell, for example in prostate cancer, which is the most common cancer in males and its unpredictable outcomes have made intervention and treatment a difficult mission, high expression levels of it have shown having a strict correlation with cell migration and invasion, and apoptosis resistance; this was confirmed by knockdown of *GARS1* gene in PC3 cells, a cellular model for metastatic prostate cancer, where after the gene manipulation, the PC3 cell acquired an attenuated phenotype (Khosh Kish E. et al.,2023). Moreover, in hepatocellular cancer (HCC) which is the most common primary liver cancer, accounting for 80-90% of all liver cancer cases and the sixth most common cancer worldwide, GARS1 high expressional level has been demonstrated to have a direct correlation with the HCC grade, focusing on its role in shifting the cell to a proliferative stage (Wang J. et al., 2022). Furthermore, in breast cancer, one of the five big killers in cancer, an AI-based molecular scoring system, developed to predict 10-year survival rates, has identified *GARS* as the third most impactful gene out of 23 coding genes when overexpressed, pointing eve more its importance in the poor overall survival (Shimizu H. et al., 2019).

1.2.3Specific role of GARS1 in Non-Small Cell Lung Cancer (NSCLC)

Non-Small Cell Lung Cancer (NSCLC) accounts for 85-90% of lung cancer, while small cells lung cancer (SCLC) is decreasing rapidly in many countries in the last few decades (Reck et al.,2014). Smoking is the main cause of lung cancer worldwide, responsible for more than 80% of cases, with fluctuation of rate which reflects the different consume of tobacco between countries. There are other known risk factors like asbestos, arsenic and other chemicals compound which had demonstrated a strong correlation with the insurgence of NSCLC. Usually, stage I and II NSCLC are treated with surgery while IV NSCLC are treated with conventional chemotherapeutics, in particular cisplatin (CDDP) has been the most effective for the last few decades, but NSCLC cases has shown frequently resistant to CDDP (Seve et al., 2005), driving scientific community to the research of new chemical perturbagens. cBio-portal is a web-based platform, which consists of different cancer genomic dataset, and provides the users the visualization,

analysis and download of the data. Observing the data, in cBio-portal, of *GARS1* in NSCLC can be seen that is predominantly dysregulated in its expression levels rather than through genetic mutation (**Figure 4A-4B**). Moreover, the higher expression levels seem to be correlated with a poor overall survival in NSCLC patients (**Figure 4C**).

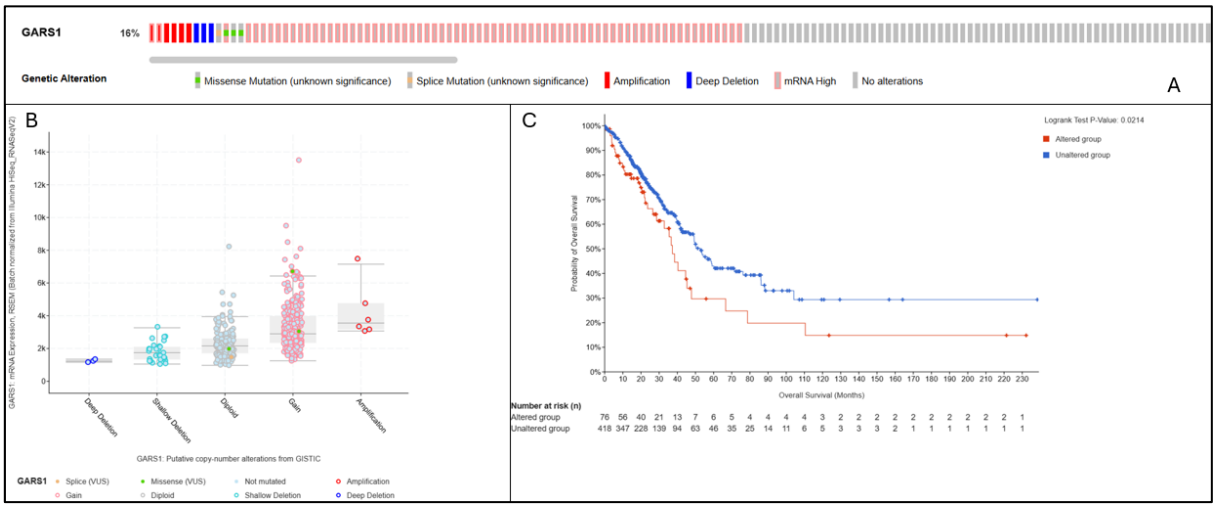

*Figure 4. **A**. The visualization of the GARS1 gene with the distribution of possible amplification, deep delection, mRNA high levels, missens mutation and splicing mutation occurring in NSCLC. **B.** The Kaplan-Meier curve describing the altered group (red) where GARS1is overexpressed, and the unaltered group (blue) where it is not; **C**.The distribution of all the GARS1 putative copy number alteration from cBio-portal.*

The product of *GARS1* has been demonstrated to be the target of a novel compound Fraisinib, able to inhibit GARS1 enzymatic activity. Fraisinib treatment of A549, a cellular model for NSCLC, resulted to have beneficial effect through decreasing cell proliferation and promoting apoptosis.

Macrocyclic compounds meso-(p-acetamidophenyl)-calix [4] pyrrole and meso-(m-acetam-idophenyl)-calix[4]pyrrole have been reported to have a cytotoxic action on lung cancer cells and prostate cancer cells in vitro and in vivo studies. The cytotoxic action of specific compounds against selected cell lines, was carried out acting as ionophores capable of altering chloride transport across cell membranes. In particular two derivates has been reported to have a different effect in specific cell line samples, the calixpyrrole derivates #107 and #563 (**Figure 5**).

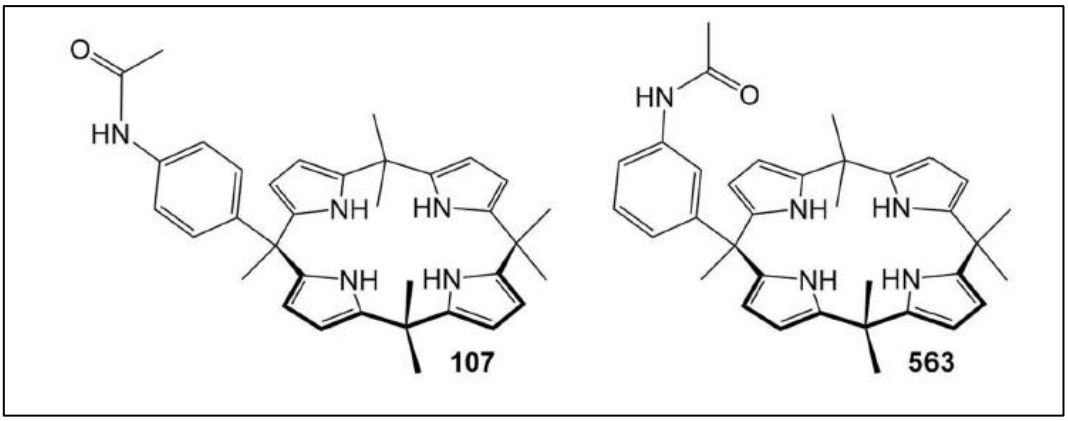

Figure 5. The calixpyrrole derivates 107 and 563 molecular composition

Means of *in vivo* and *ex vivo* experiments have shown a high cytotoxicity on prostate cancer cells (PC3, DU145) and non-small cell lung cancer cells (A549) demonstrating a potential role in different cancer cells treatment. Furthermore, following in silico prediction studies, through a functional test the GARS1 protein has been identified as the target of the meso-(p-acetamidophenyl)-calix pyrrole, named Fraisinib (Ben Toumia et al., 2024). Specifically, the study demonstrated that Fraisinib binds GARS1 protein in the same pocket of glycine plus ATP, inhibiting the catalysing action of the enzyme (**Figure 6**). These considerations take calixpyrrole derivates, in particular Fraisinib, as potential drug for treatment of NSCLC in the future and further investigations onto different cancer lines.

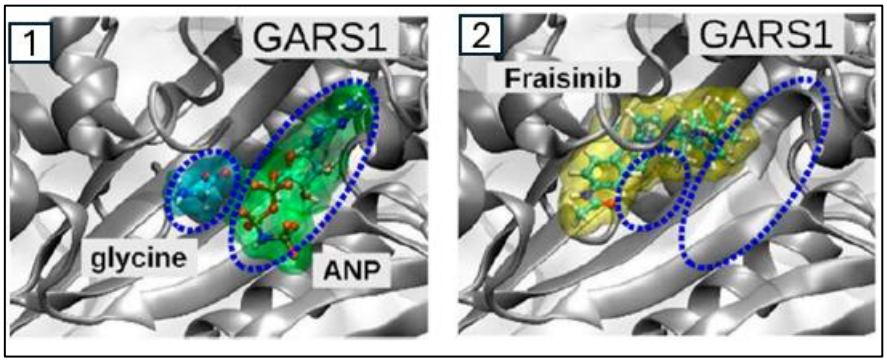

Figure 6.1 The GARS1 protein enzymatic tusk with glycine and ANP, which mimics the ATP, inside; 4.2. Fraisinib in the enzymatic tusk of GARS1 protein, preventing glycine activation.

1.3 Drug repositioning

1.3.1Definition

The development of new drugs or therapeutic molecules experiences multiple difficulties like approval by health agencies, human safety, marketing and commercial availability for medical system, which take up to 15 years to reach the market as mean (Parvathaneni et al., 2019). These conditions make drugs development a highly risky and expensive, in addition to the difficulty

of discovering new molecules targeting specific biological pathways. Furthermore, most of the novel drugs studied fail to demonstrate a good safety or efficacy when arrived in the clinical trials, failing to enter the market (Juárez-López et al., 2021).

Drug repositioning is defined as the exploration of a molecule's ability to treat diseases or condition beyond the original intended purpose to promote novel uses for known drugs, based on the principle of polypharmacology, i.e., the ability of a drug to have several targets and both off- and on-target effects, linking the genetic findings of rare disease and drug repositioning into the same framework to accelerate drug development (Delavan, et al., 2017) (**Figure 7**).

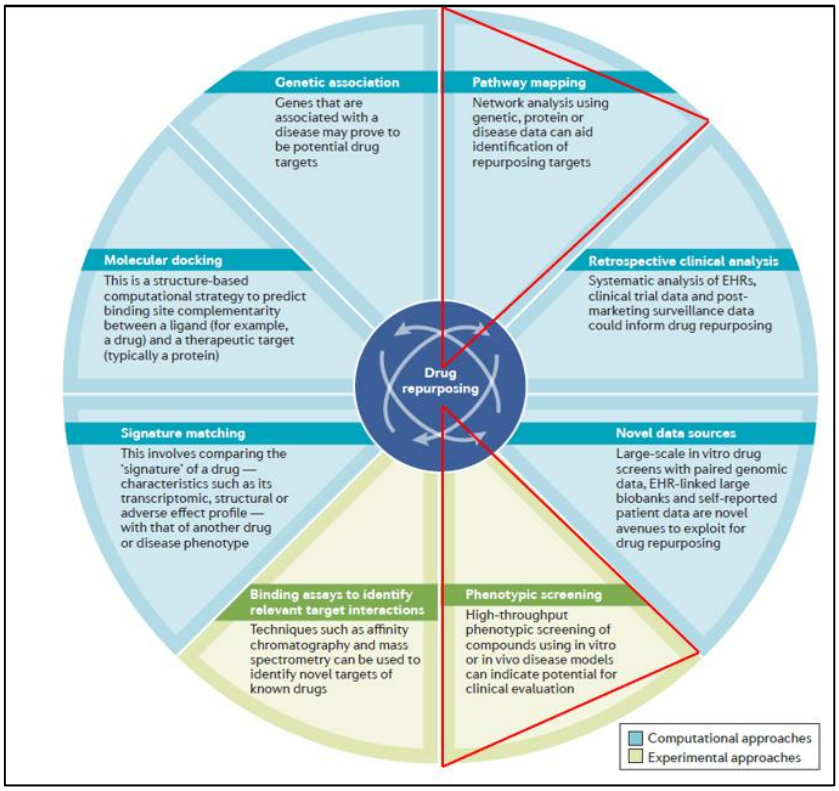

Figure 7. All the possible approaches to setting up a drug repositiong study, both computational and experimental. Highlighted in red there are the two methods used in this study. As computational method was used the pathway mapping, comparing Fraisinib gene signature onto A549 cells with dataset in ILINCS; while as experimental approach was used the phenotypic screening, setting up a high throughput screening onto A549 cells.

1.3.2Bioinformatic tools for drug repositioning

Gene signatures perturbations consequent to a disease or a treatment can be used to find drugs able to revert the pathological phenotype or, alternatively, to induce perturbation connected to

the treatment. This is the rationale of the development of sereal tools suitable to perform in silico drug repositioning.

iLINCS (http://ilincs.org) is a web-based platform that integrate multiple omics data, combined with bioinformatic tools for integrative analyses and visualization that do not require programming skills. This platform facilitates the mining of large collection of omics data, more than 34000, or pre-computed signature, more than 200000, and their connection. Moreover, iLINCS allows to submit the user omics signatures, of diseases or cellular perturbations, to compare with the omics data in the database, enabling the user to drive a drug repositioning study. So, a signature can be selected by querying the iLINCS database, submitted by the user or derived from analyzing an iLINCS omics dataset and the connected signature can be examined to identify cellular perturbation or to elucidate mechanisms of action of different compound which could lead to insights and hypotheses about potential therapeutic targets and therapeutic agents (**Figure 8**).

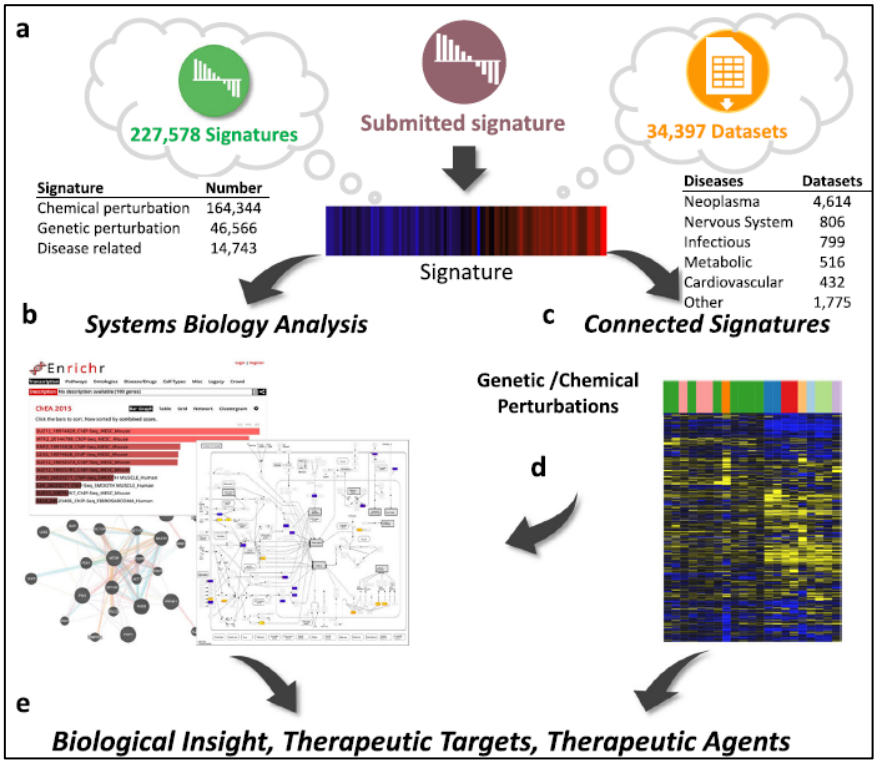

Figure 8. ***a****. Signatures can be selected by querying the iLINCS database, submitted by the user, or constructed by analysing an iLINCS omics dataset. Signatures in the database include chemical and genetic perturbation, and disease-related signatures.* ***b****. The signature can be analysed using a range of systems biology methods.****c*** *Signature "connectivity" analyses can be applied to identify cellular perturbations and biological states of similar signatures. d. The analysis of connected signatures, as well as the identity of the perturbed genes and proteins leading to the connected signatures can be used for further examination.****e.*** *Ultimately, the results of the analyses lead to insights and hypotheses about potential therapeutic targets or therapeutic agents. Picture from Pilarczyk et al., 2022.*

1.3.3In silico prediction of transcription factors

Transcription factors (TFs) recognize sequences on DNA; affinity of binding sites may change in presence of nucleotide changes and the consequence is variable strength of TFs-DNA interactions and thus different level of gene transactivation. Based on this evidence, from DNA sequence it can be predicted which TF could bind it. Promo3 is a bioinformatic tool which allows the prediction of possible transcriptional factor binding sites that could bind a specific promoter sequence target, in this case *GARS1*, from any species or group of species (D. Farre, 2003). Considering the intrinsic variability of protein recognition signals is used a representation of the binding sites that are positional weight matrices, which store information of relative frequency of different nucleotides in the recognition site, constructed on data from TRANSFAC, a database of eukaryotes transcriptional factors.

1.4 p53

The p53 protein, encoded by the *TP53* gene, was discovered in the late seventies, during research on cells transformed by simian virus 40 (SV40) and from that moment was object of different studies pointed on its better understanding. Initially was considered an oncogene involved in cell transformation (A. J. Levine and M. Oren, 2009). Then more studies on this protein have understand its role in the opposite manner, identifying p53 as an onco-suppressor.

The p53 protein (FLp53) in human consists of 393 amino acids, organized into five different domains: the two N-terminal transactivation domains (TAD1/2), the proline-rich domain (PRD), the central DNA-binding domain (DBD), the tetramerization domain (TD), and the C-terminal regulatory domain (CTD) (A. C. Joerger and A. R. Fersht, 2016). Its primarily role is a transcriptional factor and in response to stressful signal rapidly the majority of p53 protein assemble into a functional tetramer via its TD and thank to the DBD recognize the p53 binding domain located in promoter or enhancers of target genes to modulate transcription, and the TAD and CTD of p53 facilitate their interaction with cofactors for transcription mediation. The activation of p53 pathway upon cellular stress through multiple phosphorylation events and depending on the type of stress, this activation results in upregulation or repression of genes involved in cell-cycle arrest, DNA repair, apoptosis, and senescence. In addition to these canonical p53 functions, p53 has shown novel target implicated in autophagy, anti-angiogenesis, protection against oxidative stress, the regulation of metabolic homeostasis, and stem cell maintenance (**Figure 9**).

The p53 protein has many roles into the cell, from the regulation of the cell cycle, the apoptosis signalling and senescence, which are the earliest function discovered, that leads to the genomic

stability of the cell, by activating the *p21* gene which drive the cell with DNA damage to apoptosis; to modulation of the metabolism of glucose, lipids, amino acids, nucleotides, iron, redox processes, and autophagy signalling (**Figure 9**); and in the end has a great role in stem cells gene regulation, driving to differentiation, which function is similar in cancer cells, in particular metastasis cancer cells, that point the attention on its role in preventing cancer spreading.

The great variability of function that p53 performers in physiology and physio-pathologic contests explain why its loss of function, due to mutation, deletion or repression in a lot of cancer cells phenotypes is so damaging for the cell (Hanahan and R. A. Weinberg, 2011).

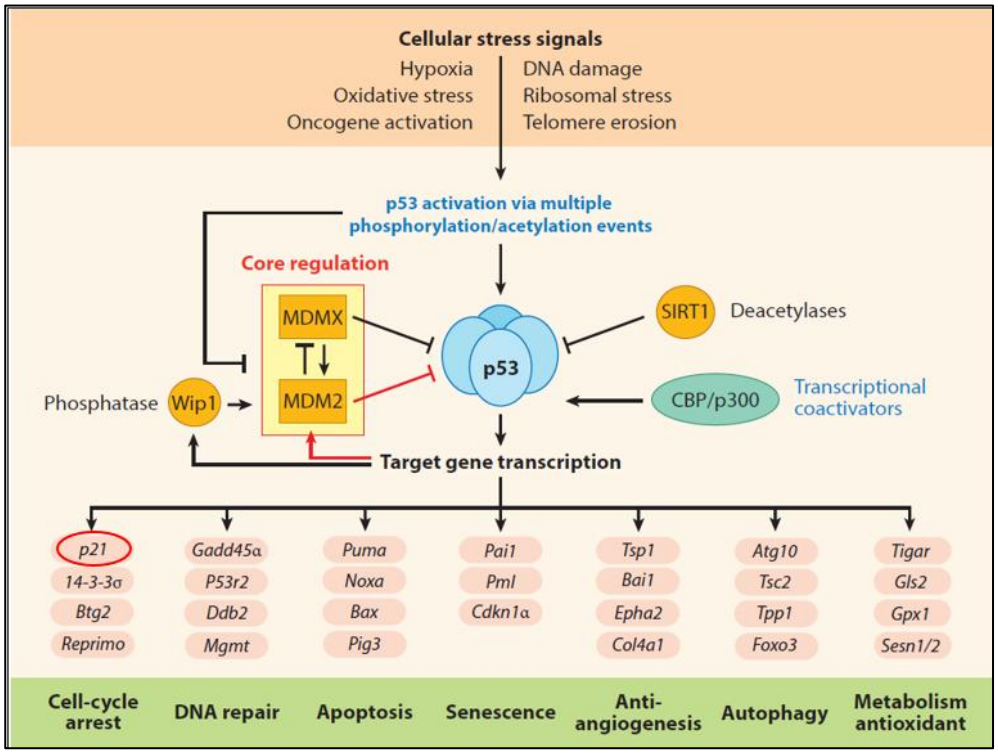

Figure 9. The p53 known pathways in cellular stress response, and circled in red the p21 factor, involved in cellular programmed death.

2 PURPOSE

The purpose of his work is to use of GARS1 expression as molecular target in non-small cell lung cancer (NSCLC). To this aim we have applied two approaches, one in vitro and one in silico, both converging in the search for drugs effective in reducing GARS1 function, that, due to the GARS1 gene overexpression, plays a pathogenetic role in NSCLC. To reach this purpose, we have generated stable cell cultures, based on the reporter luciferase, suitable to undergo automated drug screening, and we have used bioinformatics tools to search drugs acting as the

GARS1 inhibitor. Finally, as GARS1 transcription in cancer has not been investigated yet, we have searched for transcription factors as possible regulators of GARS1 expression in NSCLC.

3 MATERIALS AND METHODS

3.1 A549 NSCLC CELL LINE

A549 cells were maintained in DMEM enriched with 10% of FBS, 1% of antibiotics (penicillin/streptomycin), 1% of glutamine, incubated in a humidified atmosphere of 95% air/5% CO2 at 37°C, divided every two or three days, depending on the confluency. The freezing of the cell line was done using freezing medium composed in 10% DMSO, 50% of FBS and 40% of complete DMEM medium.

3.2 Plasmid pGL4.17[luc2/Neo]-GARS1 promoter

The pGL4.17[luc2/Neo] Vector (Promega) is suitable for the study of gene promoters as it lacks its own promoter; therefore, the expression of the luciferase reporter gene luc2 (Photinus pyralis) depends on the transcriptional activity of the regulatory region cloned in the polylinker upstream. This vector also contains a mammalian selectable marker for neomycin resistance, that allows to generate stable clones following integration on the plasmid in the mammalian genome and G418 selection. This vector is also engineered with fewer consensus regulatory sequences for reduced background and a decreased risk of anomalous transcription The pGL4.17[luc2/Neo] Vector is a basic vector with no promoter. However, the vector contains a multiple cloning region to allow cloning of a promoter of choice (**Figure 10**).

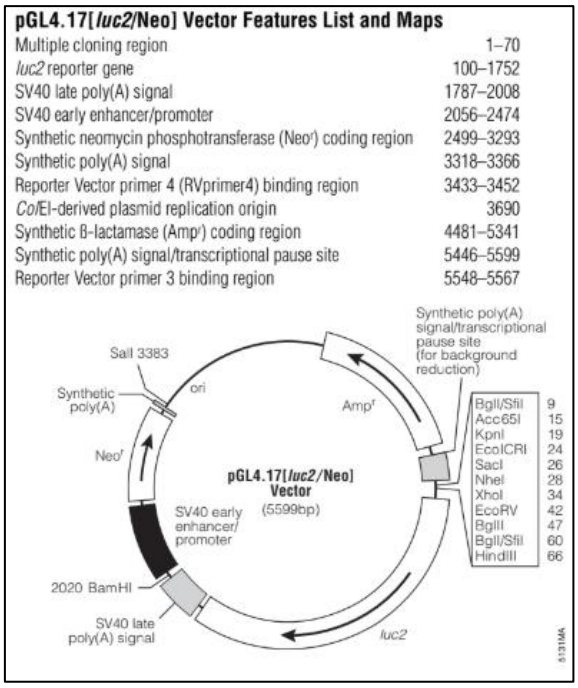

Figure 10. The pGL4.17 [luc2/Neo] vector with indicated: cloning sites, the luciferase gene luc2, ampicillin resistance, the origin of replication, and other essential components of the pGL4.17[luc2/Neo] plasmid.

3.3 Plasmids preparation

3.3.1Transformation in bacteria

The plasmid pGl4.17 was diluted to obtain a final concentration of 200 ng/ml. Then 100 uL of the bacteria line was cultured for 1 hour in double snap tubes during incubation at 37°C. After that 2 uL of plasmid was added to the solution with bacteria and mixed, and maintained in ice for 30 minutes, and after 90 seconds at 42 °C in termoblock. Under the laminal cap were added 900 uL of LB medium and, after another hour of incubation, the bacteria were plated in Agar gel overnight suspended in only 3 ml of LB added with antibiotics for the selection. The day after the bacteria were suspended in 100 ml LB and incubated overnight to obtain a higher concentration.

3.3.2Plasmids purification

For the midiprep preparation has been used the Nucleobond Xtra Midi Plus EF (Macherey-Nagel) protocol which employs a silica-based ion-exchange resin in a column to extract the plasmid DNA. At the beginning the sample was centrifugated in and suspended in 8 ml of resuspension buffer with lysozyme and RNase A and then was added the lysis buffer all in same proportion.

The ion-exchange column was prepared, equilibrated and neutralized, then the cell lysate was introduced to eliminate the others cell's components. At this point the plasmids were in the column, which was washed with a buffer, and were incorporated in a solution for the elimination of endotoxins and washed again. At the end an elution buffer was introduced to shift the pH by 7.0 to 9.0 for the DNA precipitation and the solution centrifuged for on and half hour. After that it was added ethanol, and after centrifugation the pellet was air-dried and then resuspended in sterile water, and analysed at the Nanodrop to quantify the plasmid concentration.

3.3.3Plasmid quality control

To check the correct amplification of the plasmid with the insert the plasmid was digested with restriction enzymes. In particular, the following 20µl reaction mix was prepared and incubated at 37°C for 2 hours:

1. 5 µl plasmid (100ng/ul)
2. 0.5 µl Hind III for pGl4.17 *GARS1* promoter-*luciferase* plasmid, and XhoI+NotI for pcNeo-*p53*
3. 2 µl Buffer 10X
4. 12,5 µl H2O for

At the end, the products were run on a 1% agarose gel electrophoresis with TAE and bromophenol blue to check the different bands sizes of:

1. 5562 nt (vector) and 2263 nt (*GARS1* promoter).
2. 5472 nt (vector) and 1065 nt (*p53*)

3.4 Generation of A549-*GARS1* promoter stable clones

3.4.1Choice of G418 concentration

The rational of the use of the antibiotic G418 is based on resistance given to cells when the plasmid is integrated; as integration is effective around 10-15 days after transfection, to ensure the elimination of cells without integration, the G418 dose able to kill un-transfected cells at that time without plasmid has to be chosen. For scale, G418 500 ng/ul, 600 ng/ul, 800 ng/ul were compared with an untreated plate.

3.4.2A549 transfection for stable clones generation

For the transfection, in suspension, of one million A549 cells with *GARS1* promoter- *luciferase* plasmid, was prepared a mix composed of each:

1. 5 µg of the *GARS1* promoter-*luciferase* plasmid
2. Lipofectamine 2000 in a rate of 3 µL :1µg plasmid ratio
3. 250 µl of fresh DMEM medium

The three different 10 mm plates and two 60 mm plates plates were incubated for 30 minutes at room temperature and then incubated at 37°C and checked every two days. The two 60 mm plates, one transfected with the plasmid the other used as negative control plates, were used for the luciferase assay single tube.

3.4.3Luciferase assay: single tube

The two 60 mm plates, one transfected with the GARS1 promoter-luciferase plasmid, the other used as negative control (no-transfected) were analysed.

Figure 11.The GloMax-Multi Jr Promega E6070, a single tube luminometer useful for the analysis of one sample at a time in a short period.

For the test of transfected cells, each plate was added with lysis buffer 1x and plates were placed in agitation for 15 minutes; after that cell were set in ice, incubated briefly with *Firefly* Luciferase substrate and then added the Luciferase buffer. The light produced by the reactions (**Figure 12**) was read by a single tube luminometer (**Figure 11**).

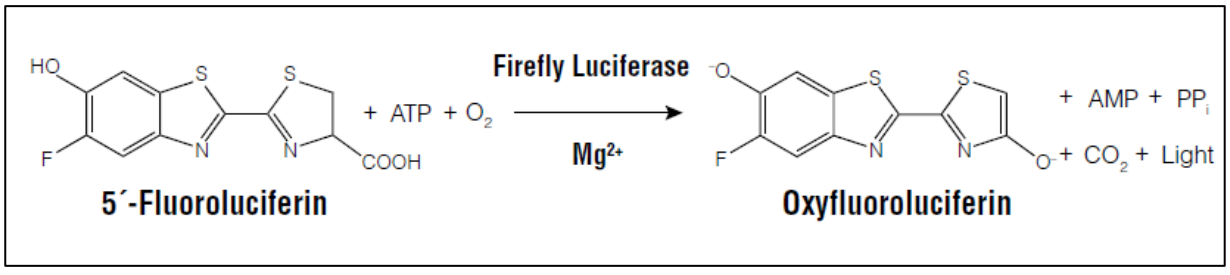

Figure 12. The reaction of 5'-Fluoroluciferin with the Luciferase enzyme which produces the Oxyfluoroluciferin plus light that can be measured with luminometer.

3.4.4Clones' expansion

The clones which survived for 15 days under geneticin selection, in the 10 mm plates, observable under a light microscope were selected and splitted in a 96 wells plate and other white ones for the luciferase control. To allow the grow of clones, cells were transferred in progressively larger areas, such as48 wells plate- 24 wells plate-6 wells plate. At the end the clones were

moved to 60 mm plate and after luciferase test to assess the maintenance of plasmid integration, cells were frozen and stocked.

3.4.5Selection of clones

The luciferase test was made with OneGlo luciferase assay; clones were plated in 96 wells; the day after, culture medium was replaced with 60 ul of fresh medium, added with 60 µl of one-Glo buffer (1:1), waited for ten minutes, and then read on the Glomax Discover.

3.5 One-Glo + Tox luciferase reporter assay

This protocol was necessary to evaluate luciferase emission and as a viability reporter on the same sample in two consecutive reactions. The luciferase emission was evaluated through the addition of luciferase substrate and the quantification of the resulting signal using a luminometer. The Ac-DEVD-AFC (C30H34F3N5O13, PM 729,61)) is a fluorogenic substrate which allows to detect the caspase-3 activity, an apoptotic signal, because when it is active it hydrolyses Ac-DEVD-AFC that releases the fluorophore AFC (**Figure 13**), so was used as cell viability control.

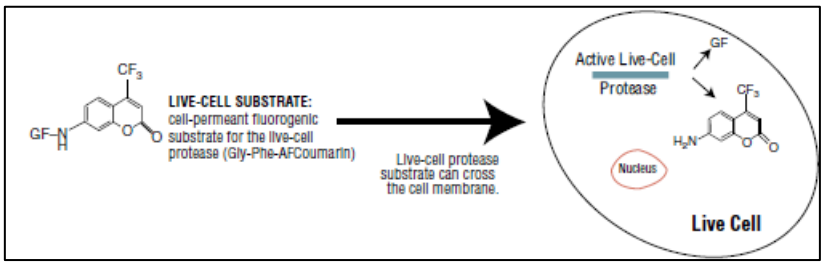

Figure 13. The AC-DEVD-AFC which can cross the cell membrane and react with the living cells protease, producing AFC.

As first the medium was changed in each well with fresh one in the white 96 wells plate and added a 5X solution of GF-AFC substrate in its buffer for every well, incubated for 30 minutes at 37°C and then measured the resulting fluorescence using the Glomax Discoverer (Promega) automated plate reader, at the length of 495-505 nm. After that, the luciferase substrate was added in proportion 1:1 with the initial volume of medium, incubated for 3 minutes at room temperature, and then the luminescence was measured with Glomax Discover (**Figure 14**).

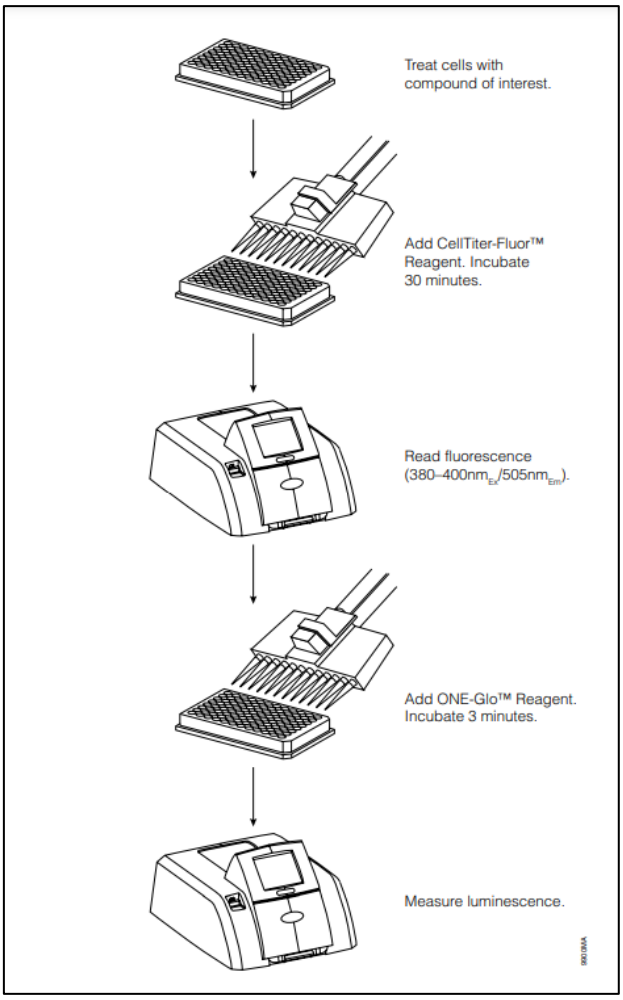

Figure 14. This figure represents the workflow of the one-glow+tox assay. After 24 hours of treatment were first added the cellTitler-Fluor, which contains Ac-DEVD-AFC and incubated for 30 minutes. After that was measured the fluorescence emission at 495-505nm with Glomax Discoverer. Thereafter was added the One-Glo reagent, which was the luciferase substrate, a measured the luciferase emission at the Glomax Discoverer.

3.6 Immunofluorescence assay

The #27 and #31 clones were subjected to fixation with methanol/acetone for twenty minutes at -20°C. Then were permeabilized with Triton-x-100 0.3% for ten minutes and incubated 5 minutes with FBS 5%/ triton-x-100 0.3% in PBS for blocking the reaction. After that the primary antibody goat-anti-luciferase (Novus Bio- nb100-1677) was prepared in a 1:50 proportion with dilution buffer and maintained in agitation at 4°C overnight. The day after, the secondary antibody anti-goat (donkey- Alexa 488) was diluted at 1:500 and added to the plate for 45 minutes at room temperature. At the end the two clones were stained with DAPI and analysed at fluorescence microscope. Alexa 488 excitation and emission spectra are shown in **Figure 15** and fluorescence is visible as green light at fluorescence microscope.

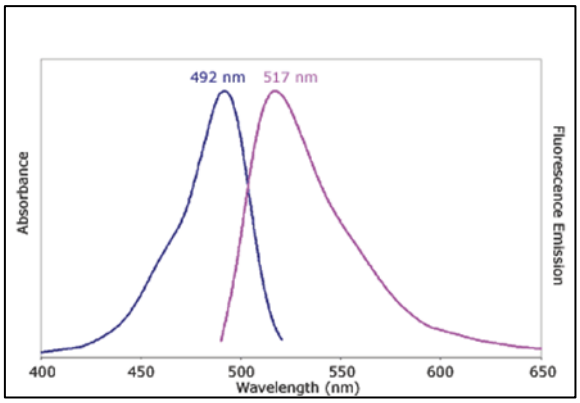

Figure 15. Alexa 488 wave lenght excitation, the blue line, and emission spectra, purple line.

3.7 Automated drug screening

The Epigenetics Library (Enzo BML-2830 v1.0), contains forty-three epigenetics drugs of different classes at 100mM in DMSO, was tested at the final concentration of 5 µM for every compound. The dilution has followed the scheme in **Figure 16**.

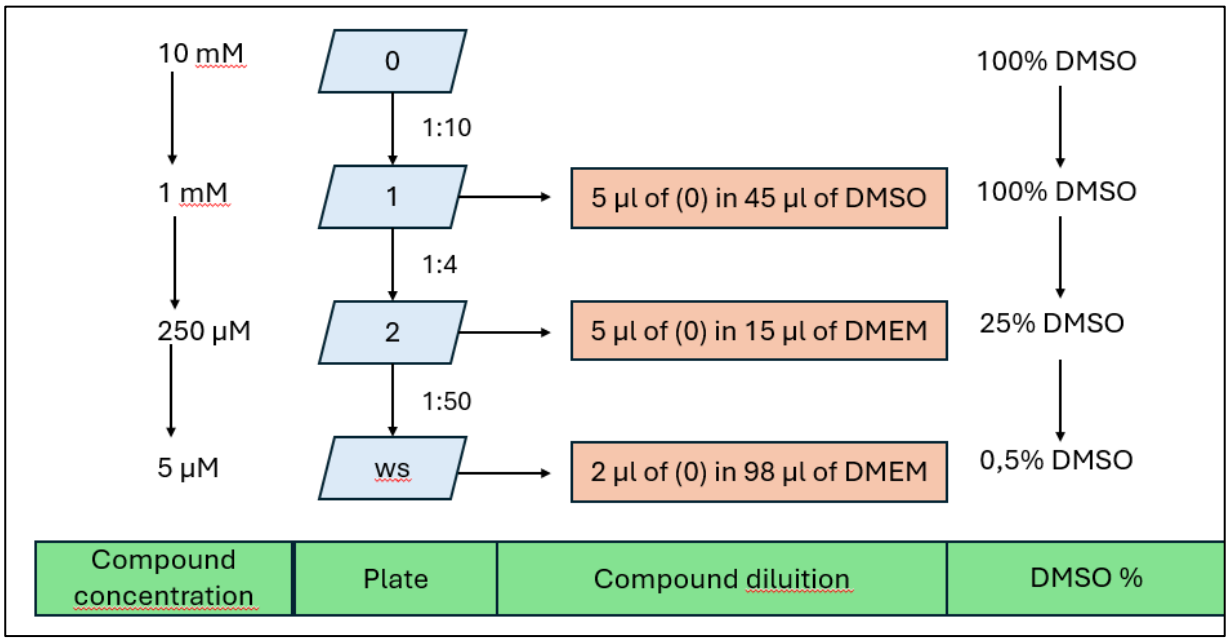

Figure 16. The scheme of dilution followed in the experiment. From the initial plate at the concentration of 10mM was diluted in a 1:10 rate at the concentration of every compound of 1 mM, and a percentage of DMSO of 100%. After that a second dilution in a 1:4 rate, to obtain every drug at the concentration of 250 µM in 25% of DMSO. At the end the last dilution in a 1:50 rate in the working solution plate to have a final concentration of every compound of 5 µM in 0,5% of DMSO.

3.8 Dual Luciferase Reporter assay

To check the co-transfection of different combination of plasmids and the observation of the luciferase data we performed a Dual-GLO Luciferase report assay, with *Renilla Reniformis* luciferase as internal control of transfection.

In a 24 well plate was added the passive lysis buffer 1x and after 15 minutes of incubation at room temperature was taken from wells the cell lysate and transferred in a 96 wells plate. At this moment the renilla luciferase substrate, at 1:50 ratio with Stop and Glo buffer, was added to the samples and measured the luciferase emission, after that luciferase substrate was added and checked this too.

3.9 Co-transfection of GARS1 promoter and p53

The co-transfection was performed in triplicate as follows:

1. Triplicate with 750 ng pcNeo-*p53* expression vector (gently gift by dr. Paola Monti) + 250ng pGl4.17 plasmid with *GARS1*-promoter+50ng renilla CMV plasmid as internal control of transfection.
2. Triplicate with equimolar pcNeo plasmid empty+ 250ng pGl4.17 plasmid with *GARS1*-promoter+50ng renilla CMV plasmid as internal control of transfection.

The same conditions were assays by using the promoter of the *p21* gene, cloned in pGL3basic (Promega) as positive control because it is known to be transactivated by *p53*.

For transfection Lipofectamine 2000 was used at 3:1 ratio with DNA in DMEM medium and incubated at room temperature for twenty minutes. The transfection was made in a 24 wells plate, where 70.000 A549 cells were incubated with transfection mix for 48 hours.

3.10 MTT assay

For the evaluation of the cell viability in A549, an assay based on the reduction of MTT 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide),was used; the principle relies on active mitochondria, that reduce the yellow dye MTT to an insoluble crystal purple formazan product, which amount produced is directly proportional to the viable cells (**Figure 17**). The reaction can be evaluated, after dissolving the formazan crystal. In detail, 50.000 A549 cells were plated in 96 well plate to have 6 wells for each drug; in addition, 6 wells for the negative control DMSO, diluent of drug, were prepared. The day after cells were added with each drug at 5 µM. The vitality check was made at 48 and 72 hours, adding the MTT solution to a final concentration of 135 ng/µl and incubated for four hours at 37°C. The solution was then removed, replaced with DMSO and incubated at room temperature for 20 minutes. At the end the absorbance at 560nm was valuated with Glomax Discover (Promega).

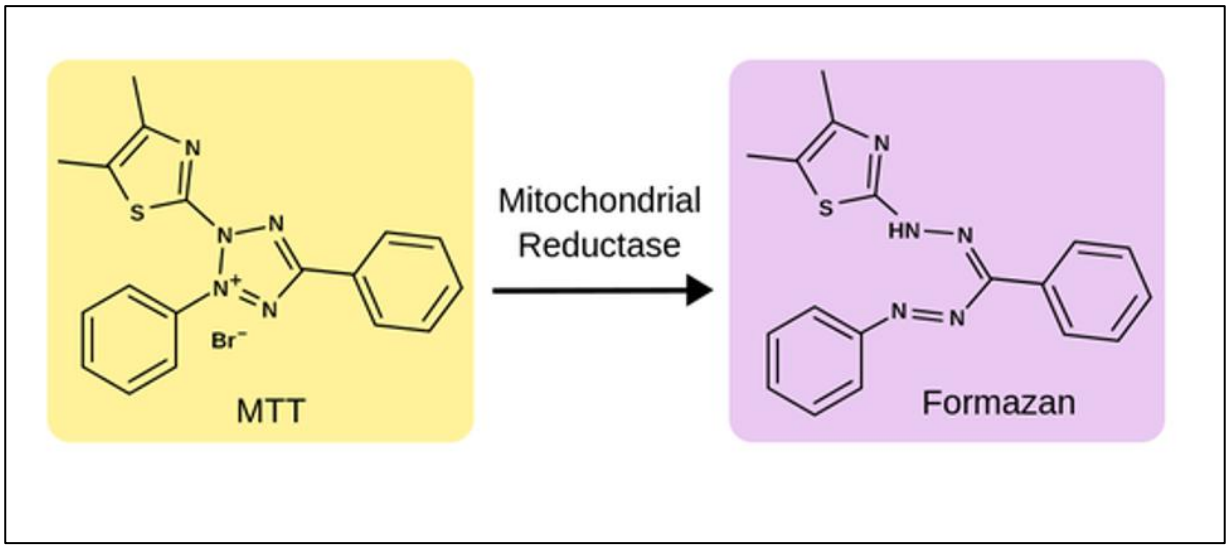

Figure 17. This figure shows the reaction from MTT, a yellow dye, to formazan, purple dye, catalysed by mitochondrial reductase of active cells.

3.11 *In silico* drug repositioning

3.11.1iLINCS

Gene signature-based drug repositioning aims to identify new therapeutic uses for already used drugs based on the idea that drugs can reverse the gene expression patterns associated with a disease or mimic the gene expression perturbation induced by another effective drug. In this work we have used the second approach. iLINCS (http://ilincs2018.ilincs.org/ilincs/) is an integrative web-based platform for analysis of omics data and signatures of cellular perturbations. This platform includes more than 34000 omics data, precomputed signatures, with their connection, and the analysis of omics signatures of disease and cellular perturbation submitted by users. This vast amount of dataset can be used to make sophisticated analyses of omics signatures, mechanism of action analysis and signature-driven drug repositioning. Among the various functions of the program there is one in particular, very useful for the comparison of the up and down regulated sequences, obtained for example from the treatment of specific cell with a target drug, which allow the research in the various dataset available in the program. Selecting the dataset of interest inside the program it allows the identification of chemical perturbagens with similar effect to the gene signature in possess.

3.11.2*In silico* prediction of transcription factors

Promo3 is a bioinformatic tool suitable for the identification of putative transcription factor binding sites (TFBS) in DNA sequences from a species or groups of species target. TFBS defined in the TRANSFAC database are used to construct specific binding site weight matrices for

TFBS prediction, with different grade of concordance. Having the sequence of the promoter of interest, in this case GARS1, the software predicts the transcription factors which could bind it.

3.12 Statistics tools

3.12.1T-test for paired comparison

Student's t-test is a statistical test used to test whether the difference between the response of two groups is statistically significant or not. If it is any statistical hypothesis test in which the test statistic follows a normal distribution, i.e. a Gaussian distribution, under the null hypothesis, i.e. there is no relation between the two group under study.

One of the possible uses for t-test analysis is the paired samples t-test which consist of a sample of matched pairs of similar set of data that has been tested twice, typically prior and after a specific treatment. A paired samples t-test based on a "matched-pairs sample" results from an unpaired sample that is subsequently used to form a paired group, by using additional variables that were measured along with the variable of interest. So, observing one time the two similar samples but with different conditions allows to reduce or eliminate the effect of confounding factors.

3.12.2p-value

The p-value is a strong statistical tool, used as the first filter in the analysis of the results of an experiment. In null-hypothesis significance testing, the p-value is the probability of obtaining test results at least extreme as the result observed, assuming that the null hypothesis is correct. Considering an observed test statistic (t) from unknown distribution (T), the p-value represents the prior probability to observe a test-statistic value as extreme as t if null hypothesis were true. Furthermore, during a study, the error more important to avoid is the "error of the first kind", i.e. rejecting a null hypothesis that was actually true. To reject the null hypothesis the p-value must have a value below 0,05 or 0,01 for avoiding the error of the first kind and reject correctly the null hypothesis.

4 RESULTS AND DISCUSSION

4.1 pGL4.17 - *GARS1* promoter vector-quality check

The pgl4.17 vector in which the GARS1 promoter driving the expression of the Luciferase gene was cloned in the polylinker in KpnI-NheI sites (**Figure 18A**) was first transformed with competent bacteria, E. Coli, to obtain a high concentration suitable for transfection. Before doing the transfection in A549 cell was performed an enzymatic digestion with Hind III which cuts at the 3' extreme of the polylinker (**Figure 18A**) and in the *GARS1* promoter gene sequence in position 150 (**Figure 18B**). The *GARS1* promoter-*luciferase* gene plasmid was digested in two fragments of 5562 pb and 2263 pb **Figure 18C-lane B**, while the not-digested resulted in a single band of DNA supercoiled thus running faster than expected in the electrophoresis (**Figure 18C-lane A**).

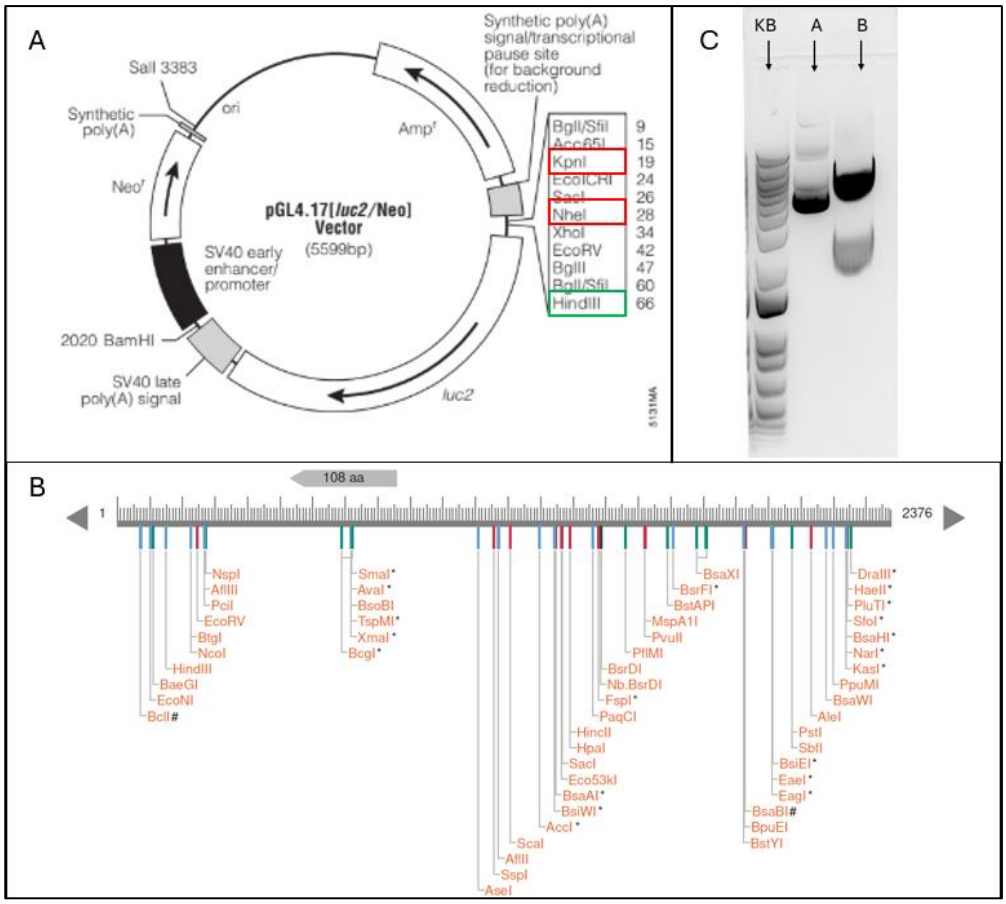

*Figure 18.**A**. In this figure are marked in red the cloning sites Kpnl and Nhel of the pGL4.17 plasmid where was cloned the GARS1 promoter. In green is marked HindIII site, in the 3' extreme of the polylinker region, where was performed the enzymatic digestion.* ***B.*** *The promoter sequence and all its restriction enzyme sites, with HindIII in position 150.* ***C.*** *The electrophoresis of the GARS1 promoter-luciferase gene plasmid digested by HindIII (pointed with B), and no-digested (pointed with A), with the marker KB by side to visualize the fragments size.*

4.2 Generation of GARS1 promoter-luciferase gene plasmid stable clones

4.2.1G418 scale

The G418 or Geneticin was used to select le stable clones, because the plasmid contains a resistance for this antibiotic which allows the selection against non-transfected cells. After fourteen days treatments of with increasing doses, 500 ng/µl was chosen as concentration for the generation of stable clones as it was able to kill all un-transfected cells.

4.2.2Transfection in A549 cells

One million A549 cells were transfected on 10mm diameter dishes with 5 µg of plasmid by Lipofectamine 2000 in 3:1 rapport. Then the cells were cultured under G418 selection to select clones with DNA plasmid integration. The isolated clones growing into petri plate were observed at the brightfield light microscope after 14 days selection, picked and transferred two in 96 wells plate, one for growing and one white for the check of the luciferase activity, allowing to identify the positive clones with plasmid integration. We checked the Luciferase activity for 77 picked clones (Figure 19) and we were able to obtain 23 stable clones positive for the Luciferase activity. The positive clones were selected and maintained to final plate of 60 mm and then collected.

	First experiment											
N.C.	250,00	290,00	350,00	360,00	380,00	320,00	230,00	340,00	240,00	260,00	400,00	450,00
CLONES	310,00	9524,00	57620,00	990,00	990,00	510,00	280,00	19830,00	3230,00	1040,00	40470,00	137300,00
	300,00	4261,00	1050,00	497400,00	830,00	350,00	420,00	1980,00	76000,00	300,00	480,00	720,00
	490,00	690,00	3391,00	760,00	580,00	340,00	270,00	570,00	24200,00	360,00	530,00	3200,00
	1730,00	330,00	680,00	660,00	840,00	5371,00	2330,00	260,00	38420,00	1180,00	240,00	390,00
	3351,00											
	Second experiment											
	260,00	310,00	250,00	320,00	330,00	270,00	310,00	390,00	300,00	270,00	270,00	230,00
	2290,00	220,00	1690,00	28060,00	5171,00	1530,00	9795,00	1880,00	19300,00	760,00	270,00	360,00
	190,00	220,00	990,00	510,00								

Figure 19. The visualization obtained by the analysis with the Glomax Discover of the Luciferase activity onto alle the 77 potential clones. In the first line there are the negative control and below all the picked clones. Were considered selectable all the clones with a value above 990, for a total of 23 stable clones. The colour pattern is produced by the Glomax Discoverer, and indicates to the lowest value analysed (blue) to the higher (red) for each detection.

4.3 Qualitative validation of clones

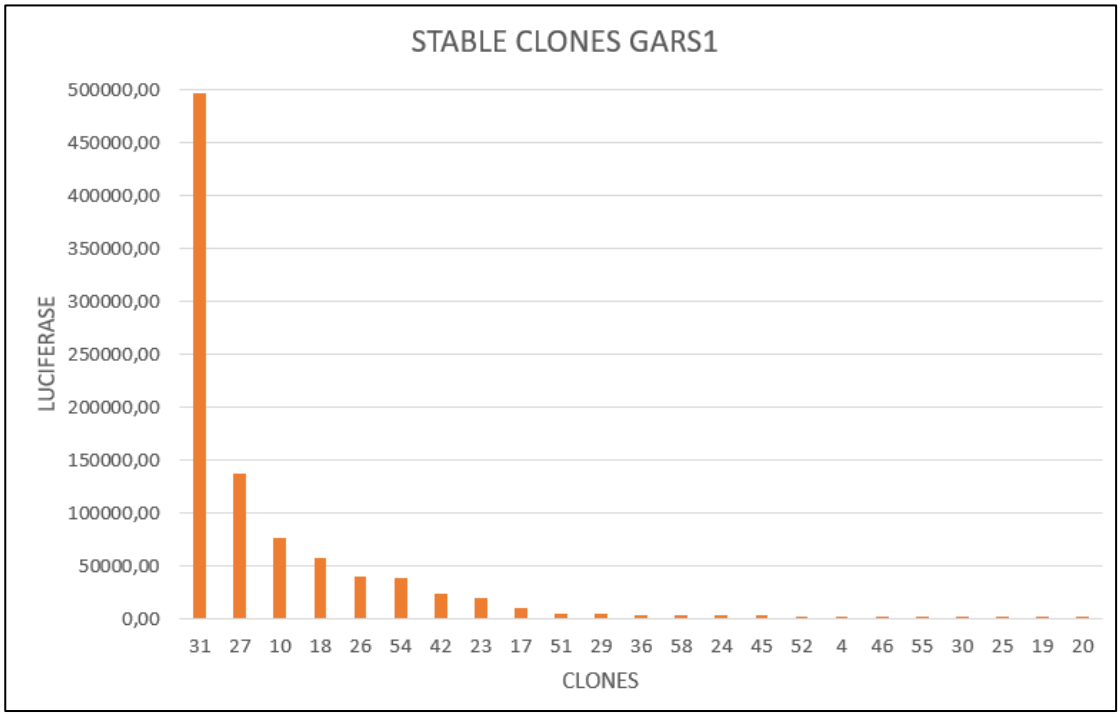

Figure 20. The visualization of all 23 stable clone for GARS1-promoter-luciferase gene reporter, where the clone #31 has the top value of all, and the clone #27 as the second most high clone.

The luciferase activity associated with clones #27 and #31 were the highest and we performed further validations to choose the best to use in automated drug screening. First, six wells for each clone were re-analysed to assess the homogeneity of the signal; after confirming such stability (**Figure 21**), the homogeneity of the clone was evaluated by fluorescence microscopy.

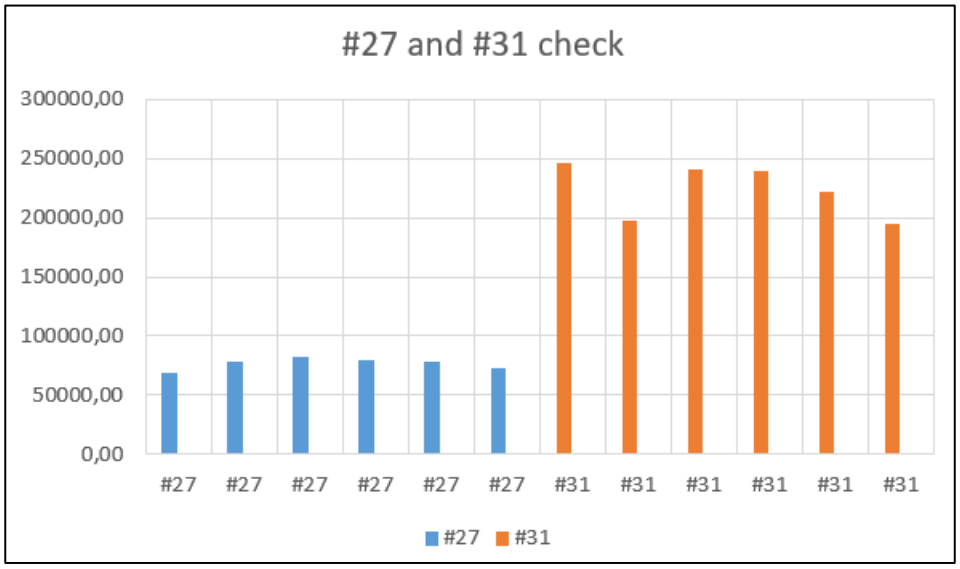

Figure 21. The clone #27 and #31, in six wells for each clone, were re-analysed to assess the homogeneity of the signal.

The expression of the Luciferase protein was evaluated by using a primary goat anti-Luciferase antibody followed by anti-goat Alexa 488 secondary antibody. To visualize the difference into the sample a primary antibody anti-luciferase and a secondary antibody Alexa 488-α-goat were used. Fluorescence microscope analysis showed that clone #31 was heterogenous, with some cells with a low intensity of Luciferase and some very high that justify the high Luciferase emission but was not suitable for the experiment, while the clone #27 was homogeneous, with all cells characterized by a similar *luciferase* expression (**Figure 22**); therefore, clone #27 was chosen for the screening with the epigenetic drugs library.

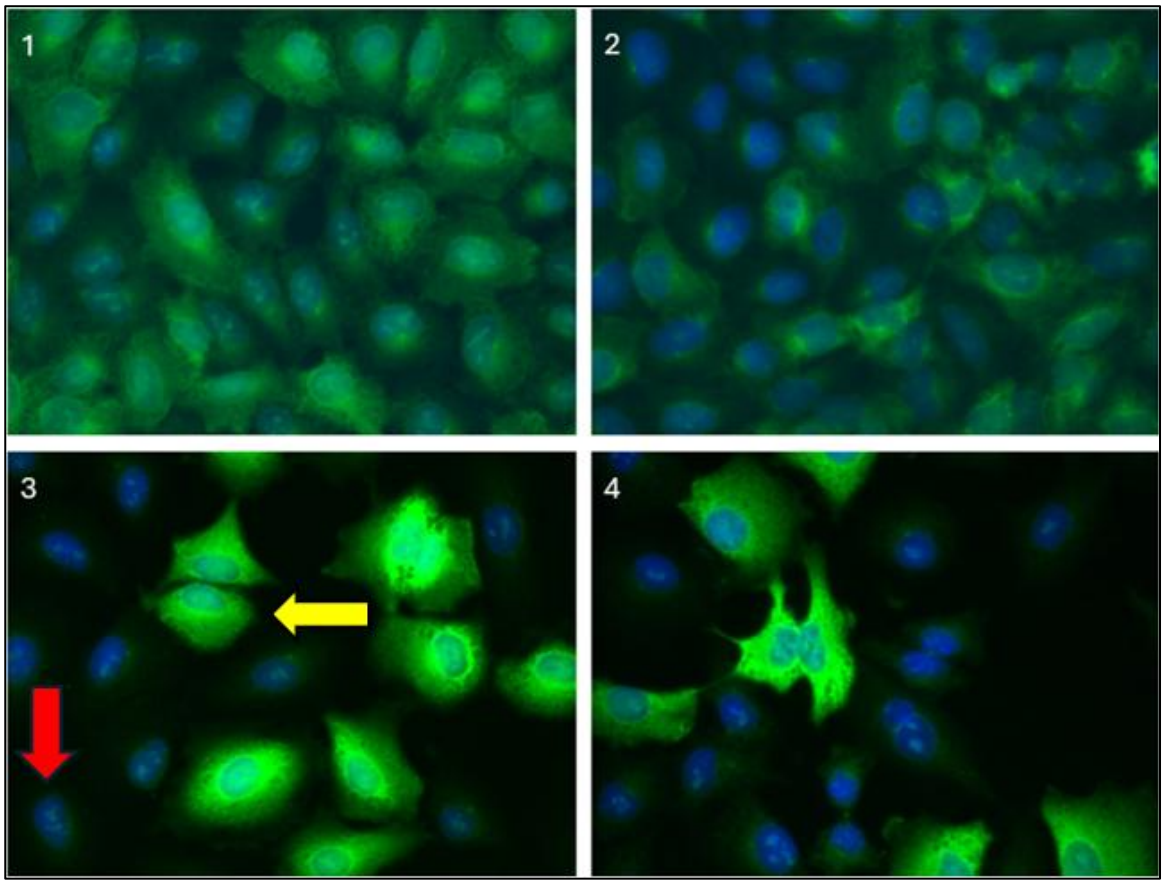

*Figure 22. **1-2**. In these two imagines is visualized the clone #27 where the antibody anti-Luciferase with emission in green, and marked with DAPI, which have emission in blue, the cellular nucleuses. The different cells have a homogenous emission of Luciferase. **3-4**. In these two imagines is visualized the clone #31 and as can be seen some cells are positive(yellow) others are negative(red), so the clone is heterogenous.*

4.4 Epigenetic library screening

The clone #27 was used to perform the screening of all 43 drugs contained in the ENZO's epigenetic library, compared with negative control samples, added only with DMSO. The same treatment was replicated three times to allow statistical analysis. The evaluation was made 24 hours after the treatment, using a OneGLO+TOX protocol, which allows to evaluate the cells vitality using AFC emission at 495-505 nm and in the same sample the Luciferase activity at

550-570 nm. The ENZO epigenetic library is composed of 43 compounds with different effects on DNA transcription and translation, such as: DNA metil-transferase inhibitors, Histone Acetyltransferase (HAT) inhibitors, Histone deacetylases (HDAC) inhibitors, SIRT activator/inhibitors, and others as descripted in the **Figure 23**.

EPIGENETICS LIBRARY
Enzo BML-2830 v1.0

Plate Location	Name	MW	Compound description
A-01	Trichostatin A	302,4	HDAC inhibitor
A-02	2,4-Pyridinedicarboxylic Acid	167,1	Histone demethylase inhibitor
A-03	Garcinol	602,8	HAT inhibitor
A-04	Splitomicin	198,2	SIRT-2 inhibitor
A-05	BML-210	339,4	HDAC inhibitor
A-06	Apicidin	623,8	HDAC inhibitor
A-07	Suberoyl bis-hydroxamic acid	204,2	HDAC inhibitor
A-08	Scriptaid	326,4	HDAC inhibitor
A-09	Nullscript	298,3	Scriptaid Neg control
A-10	Decitabine	228,2	DNA Me transferase inhibitor
A-11	Zebularine	228,2	DNA Me transferase inhibitor
A-12	Vorinostat (SAHA)	264,3	HDAC inhibitor
B-01	Isonicotinamide	122,1	nicotinamide antagonist
B-02	ITSA-1	292,1	Inhibitor of TSA activity
B-03	Sodium Phenylbutyrate	186,2	HDAC inhibitor
B-04	Tranylcypromine hemisulfate	182,2	Lysine demethylase inhibitor
B-05	Valproic acid	144,2	HDAC inhibitor
B-06	6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide	248,7	SIRT1 inhibitor
B-07	Resveratrol	228,2	SIRT1 activator
B-08	M-344	307,4	HDAC inhibitor
B-09	Nicotinamide	122,1	SIRT inhibitor
B-10	BML-266	478,5	SIRT2 inhibitor
B-11	Piceatannol	244,2	SIRT activator
B-12	Fluoro-SAHA	282,3	HDAC inhibitor
C-01	Valproic acid hydroxamate	159,2	HDAC inhibitor
C-02	AGK2	434,3	SIRT2 inhibitor
C-03	Salermide	394,5	SIRT inhibitor
C-04	MC-1293	284,3	HDAC inhibitor
C-05	Anacardic acid	348,5	HAT inhibitor
C-06	B2	396,8	SIRT2 inhibitor
C-07	BIX-01294 . Trichloride. Hydrate	600,0	Histone methyl transferase inhibitor
C-08	Butyrolactone 3	184,2	HAT inhibitor
C-09	CTPB	554,1	HAT inhibitor
C-10	Oxamflatin	342,4	HDAC inhibitor
C-11	Sirtinol	394,5	SIRT inhibitor
C-12	Suramin . hexasodium salt	1429,2	SIRT1 inhibitor
D-01	BML-278	391,5	SIRT1 actvator
D-02	NCH-51	390,6	HDAC inhibitor
D-03	CI-994	269,3	HDAC inhibitor
D-04	NSC-3852	174,2	HDAC inhibitor
D-05	Aminoresveratrol sulfate	325,3	SIRT1 activator
D-06	BML-281	446,5	HDAC-6 inhibitor
D-07	Triacetylresveratrol	354,4	SIRT1 activator

Figure 23. The epigenetic library compounds (ENZO Life Sciences), with the description of the epigenetic effect onto cells.

The experimental goal was to search all the possible drugs acting by decreasing *GARS1* promoter transcription, observable through the Luciferase activity; in addition, we have excluded drugs too toxic reducing cell viability too much, not allowing a correct evaluation of *GARS1* transcriptional activity. For each sample, Luciferase's activity values were normalized on via-

bility values, to select drugs really reducing *GARS1* transcription despite viability. This approach was necessary because the decrease in Luciferase activity evaluated without normalization may result from a direct effect of the drug on the *GARS1* promoter but also from a reduction in cell number, which in turn may be due to drug-induced cytotoxicity.

So, the normalized data was compared with the analysis on the negative controls, treated with DMSO, and then expressed in percentages to visualization. From the three experiments has emerged that from the 43 initial compounds, 10 had a potential activity on *GARS1* promoter (**Figure 24**), visualized by the decrease of luciferase emission with the luminometer, in particular: Butyrolactone 3, CTPB, B2, Anacardic acid, Valproic acid, 6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, Sirtinol, Salermide, MC-1293,SodiumPhenylbutyrate.

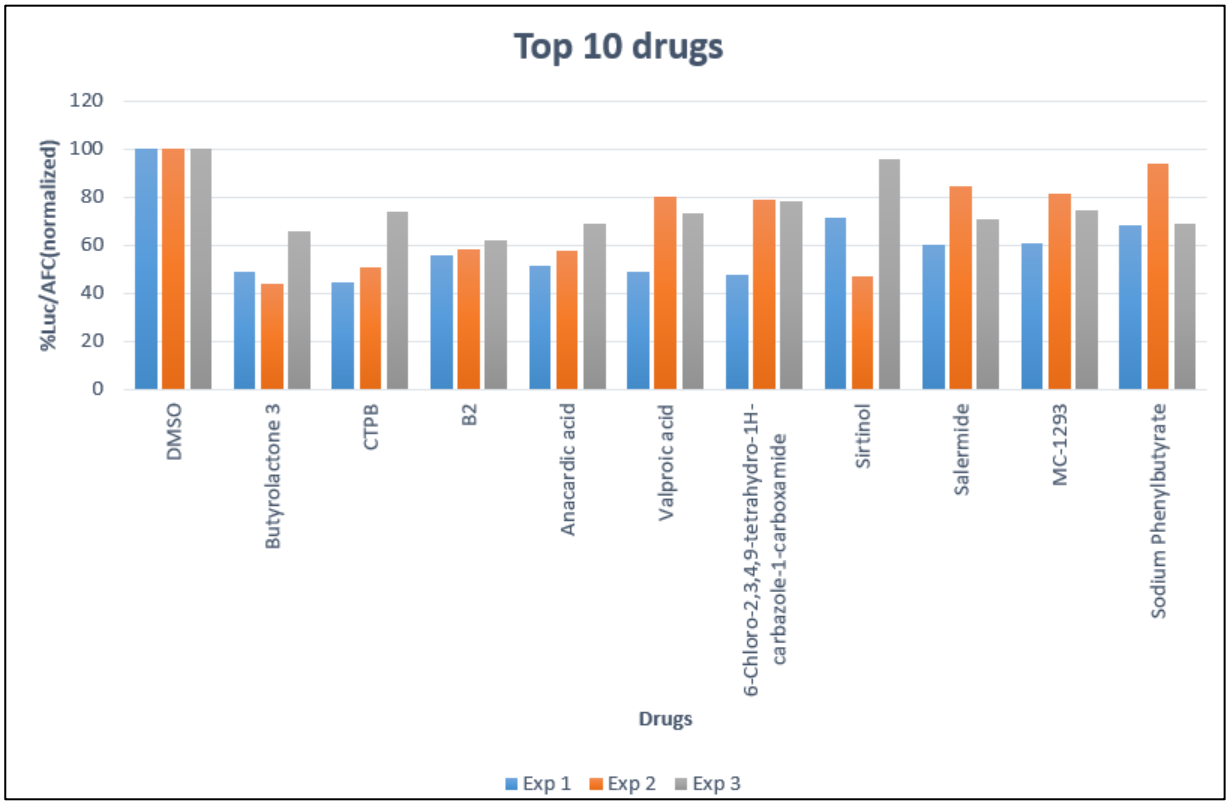

Figure 24. This figure represents the graphic of all three experiments of the top 10 drugs which had a decreasing action on the percentage of Luciferase/AFC value, used to normalize the experimental results.

For each of this 10 compounds the p-value was calculated to finally select with statistically significant; in the end, butyrolactone 3, CTPB, B2, anacardic acid and MC-1293, were confirmed to have a statistic effect on the decrease of normalized Luciferase activity, and thus *GARS1* promoter activity (**Figure 25**).

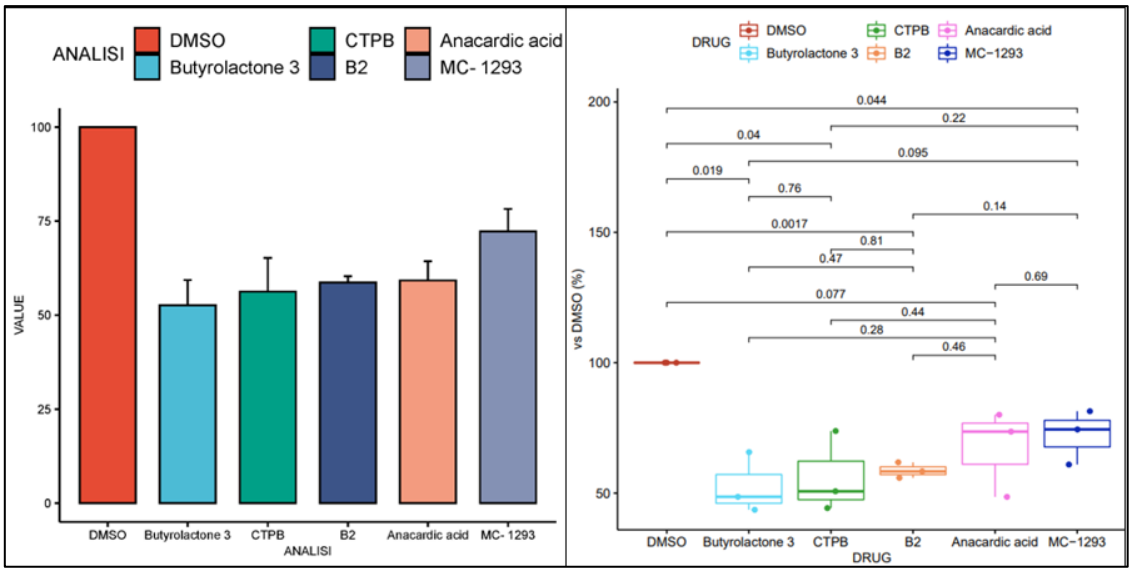

Figure 25.In the left of the figure there is the representation of the decreasing action of butyrolactone 3, CTPB, B2, anacardic acid and MC-1293 compared to DMSO. In the right are indicated the p-value resulted from the three experiments performed on the drugs, calculated using SRPlot.

4.5 Time-course of drugs effect on cell viability

To investigate whether the effect of treatment beyond 24 hours was able to reduce cancer cells viability, a MTT assay was performed A549 wildtype cells after 48 and72 hours treatment. In particular, the following drugs were tested: Butyrolactone 3, CTPB, B2, Anacardic acid, Valproic acid, 6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, Salermide, MC-1293, Sodium Phenylbutyrate. Every compound was examinated in sextuplicate at 48 and 72 hours after treatment and the absorbance was determinate with the Glomax Discover. This work is still in progress and data showed here for some drugs are still preliminary as lacking statistical analysis (Anacardic acid, CTPB and Butyrolactone 3). For Anacardic acid, CTPB and Butyrolactone 3 was made only one experiment (group A) while for Sodium Phenylbutyrate, Valproic acid, 6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide, Salermide, MC-1293 and B2 an experimental triplicate was made (group B).

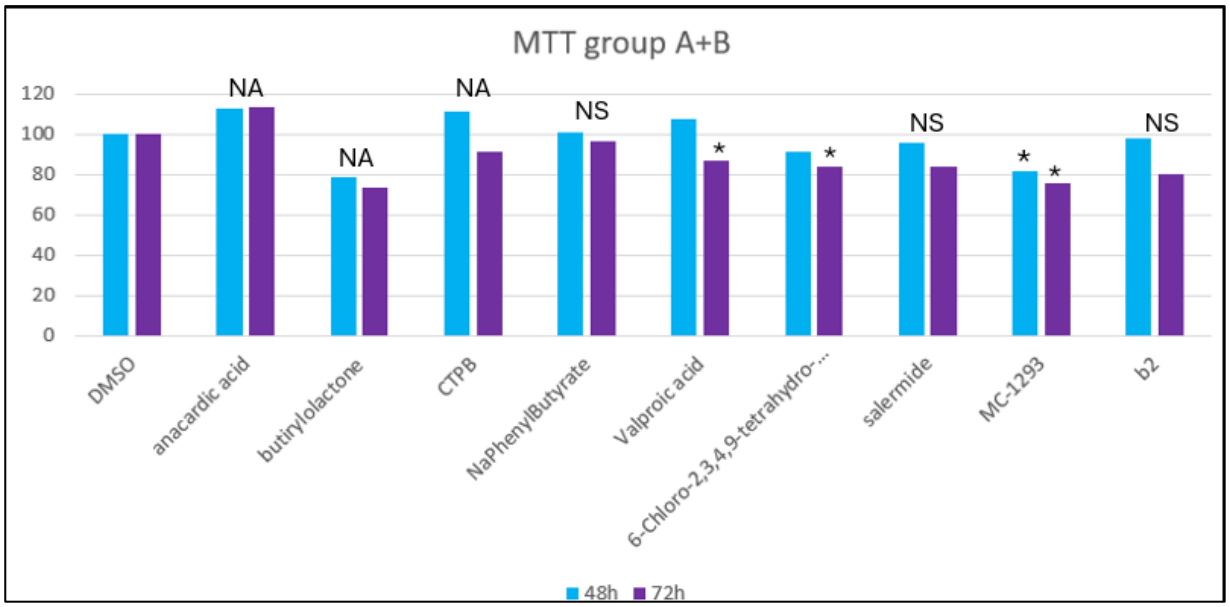

Figure 26.The graphical representation of all the MTT assay data, where "NS" means non-significant, "NA" for the compound with only one experimental data and "" for the significant data.*

From these experiments had emerged that: in the group A the most interesting was butirylolactone, which could be trustful of further investigation; in the group B the only statistical significant at 48 hours was MC-1293), while at 72 hours valproic acid, 6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide and MC-1293 were statistical significant (figure 23). At the end from an initial library composed of 43 compounds, we discovered butyrolactone 3, CTPB, B2, anacardic acid and MC-1293 having a statistical effect in decreasing the GARS1-promoter activity at 24 hours and we observed a similar effect, through cell viability, at 48 hours for MC1293, and 72 hours for 6-Chloro-2,3,4,9-tetrahydro-1H-carbazole-1-carboxamide and MC-1293. Drugs seems to have effect in counteracting NSCLC viability and further experiments should be set up to evaluate whether the decreased in cell viability can be ascribed to apoptosis or reduced proliferation

4.6 Gene signatures based-*in silico* drug repositioning

For this experiment was used the data of gene signature of the treatment on A549 wildtype cells with the Fraisinib, where GlyRS1 was sterically hindered by the drug, resulting in an overall inhibition of GARS1 protein and thus of all pathways regulated by its function. The rationale of this approach is based on observation that inhibition of GARS1 by drug (Fraisinib) resulted in beneficial effects against cancer cells fitness. Therefore, the search for drugs inducing gene signatures mimicking Fraisinib effects could unravel novel molecules acting on *GARS1* expression. The different gene signatures were examined at three different time of 24/48/72 hours. Specif-

ically, were used the top 150 up-regulated and the top 150 down-regulated differentially expressed proteins derived from proteomics analysis in A549 after Fraisinib treatment ranked for Fold change and with p-value<0,05.

iLINCS allows to visualize both the drugs inducing a gene signature similar to the ranked data used for the analysis and the chemical perturbagens that have an opposite results; in addition it allows to filter results based on time of treatment, p-value, cell line used for the specific compound, tissue of origin, concentration used in the experiment and to visualize the top 10 ranked for concordance or discordance with the data used (**Figure 27**).

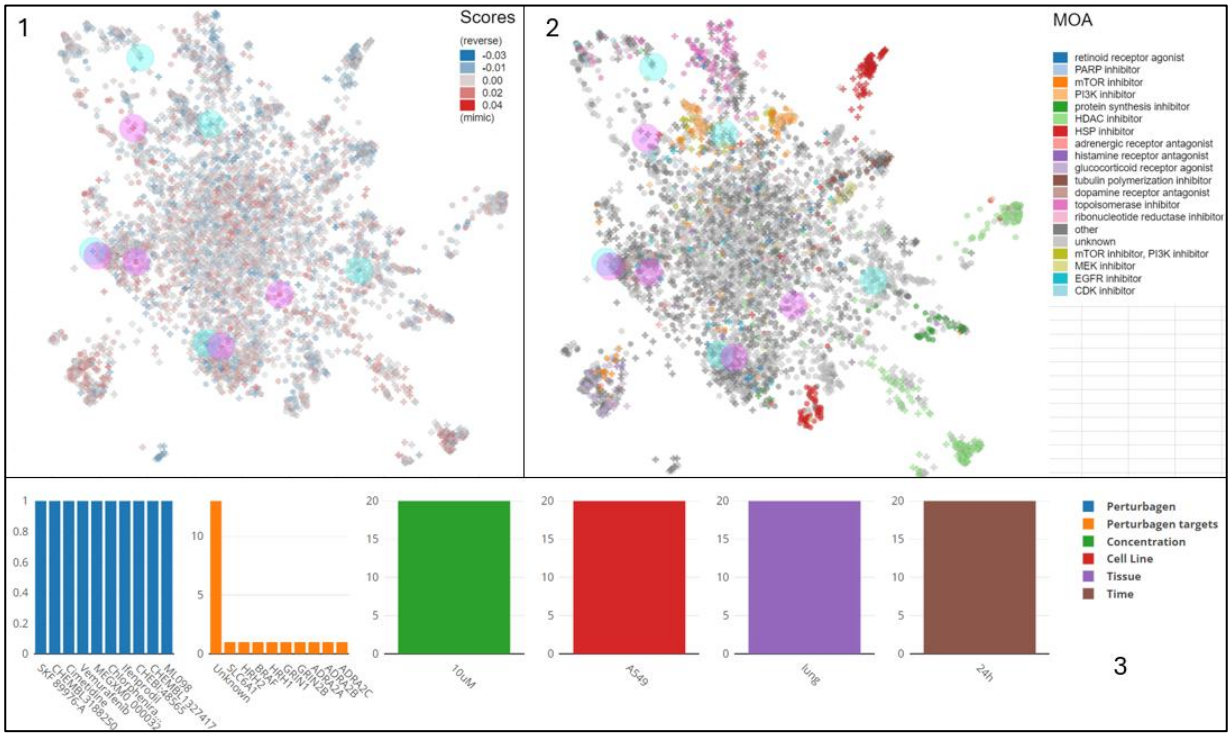

Figure 27. This is one of the visualizations produced using Ilincs tools, were in 1. This is the visualization of the compounds in the dataset with similar(red) or opposite(blue) molecular signatures compared to Fraisinib one, with a score of similarity;2. There are indicated all the classes of compound resulted from the analysis; 3.the representation of the top 10 ranked compound with similarity with the Fraisinib DEPs.

From the preliminary analysis thousands of perturbagens with action on gene expression similar to Fraisinib had emerged for each time; results were filtered for concentration of compound (10µM), tissue (lung), cell line (A549), time (24 hours was the maximum selectable for all the results), concordance (settled >0) and p-value <0,05. These data were further filtered based on the identification or no of compounds' gene targets. At the end approximately 100 chemical perturbagens for each time, with their validated gene target, were identified.

4.7 Search for transcription factors among iLINCS drugs targets

Promo3 is a useful informatic tool that allowed, after the input of the nucleotides sequence of GARS1 promoter, to investigate the possible transcriptional factor (TF) which could physically bind the DNA sequence; 3 different grade of dissimilarity, 0%5%/15 % are selectable, where in the higher grade of dissimilarity are included the TF of the lower class, as indicated in TAB.2. The TF list obtained was used to filter the data obtained by ILINCS, searching all the chemical perturbagens which had as target one or more TF that could interfere with GARS1 gene expression, as described in TAB.1.

Chemical perturbagens resulted from the in silico-based approach		
TAB.1		
SIGNATURE ID	CHEMICAL PERTURBAGENS 24H	TF TARGET
LINCSCP_5740	Ethylestrenol	AR
LINCSCP_7746	Hydroxyprogesterone Caproate	AR|ESR1|ESR2|NR1H4|NR1I2|NR3C1|PGR
LINCSCP_7309	Danazol	AR|ESR1|GNRHR|PGR
LINCSCP_6398	Arundine	AR|HIF1A|IFNG|PI3
LINCSCP_7354	Dienestrol	ESR1
LINCSCP_190602	ZINC27646218	ESR1
LINCSCP_7251	979-32-8	ESR1
LINCSCP_10102	Raloxifene	ESR1|ESR2
LINCSCP_189475	Toremifene	ESR1|ESR2
LINCSCP_10106	Raloxifene	ESR1|ESR2
LINCSCP_10104	Raloxifene	ESR1|ESR2
LINCSCP_7046	Clocortolone Pivalate	NR3C1
LINCSCP_7323	Mometasone Furoate	NR3C1
LINCSCP_6877	Dexamethasone 21-acetate	NR3C1
LINCSCP_6767	SPECTRUM_001522	NR3C1
LINCSCP_9147	Quinacrine	TP53
LINCSCP_6473	25-Hydroxycholecalciferol	VDR
SIGNATURE ID	CHEMICAL PERTURBAGENS 48H	TF TARGET
LINCSCP_7602	Flufenamic acid	AKR1C3|AR|PTGS1|PTGS2
LINCSCP_190602	ZINC27646218	ESR1
LINCSCP_193174	INCA-6	NFATC1
LINCSCP_7304	Isoflupredone Acetate	NR3C1
LINCSCP_191493	Tretinoin	RARA|RARB|RARG|RORB
LINCSCP_6437	Acitretin	RARA|RARB|RARG|RXRA|RXRB|RXRG
SIGNATURE ID	CHEMICAL PERTURBAGENS 72H	TF TARGET

LINCSCP_10123	2-(4-(4-Chloro-1,2-diphenylbut-1-en-1-yl)phenoxy)-n,n-dimethylethanamine	ESR1
LINCSCP_7305	Zuclomiphene	ESR1

The TF predicted with Promo3		
TAB.2		
TF 0% dissimilarity	**TF 5% dissimilarity**	**TF 15% dissimilarity**
C/EBPbeta	C/EBPbeta	C/EBPbeta
GR-beta	GR-beta	GR-beta
STAT4	TFIID	XBP-1
TFIID	TFII-I	TFIID
ER-alpha	STAT4	HNF-3alpha
YY1	STAT1beta	GR
XBP-1	GR	NF-AT2
HNF-3alpha	PXR-1:RXR-alpha	TFII-I
RXR-alpha	Pax-5	STAT4
TFII-I	GR-alpha	c-Ets-1
HOXD9	p53	STAT1beta
HOXD10	PR B	SRY
GR-alpha	PR A	TCF-4E
HNF-1C	ER-alpha	STAT5A
GATA-1	YY1	VDR
Pax-5	GATA-1	FOXP3
p53	HNF-3alpha	PXR-1:RXR-alpha
GR	NF-Y	c-Myb
IRF-2	FOXP3	MEF-2A
FOXP3	C/EBPalpha	GR-alpha
NF-1	NF-AT1	Pax-5
	NFI/CTF	p53
	AR	NFI/CTF
	XBP-1	PR B
	PITX2	PR A
	Ik-1	ETF
	AP-2alphaA	ENKTF-1
	HOXD9	ER-alpha
	HOXD10	YY1
	VDR	GATA-1
	RXR-alpha	HNF-1C
	c-Ets-2	HNF-1B
	c-Jun	C/EBPalpha
	NF-AT1	NF-Y
	c-Ets-1	NF-1
	Elk-1	IRF-1
	GATA-2	NF-AT1
	HNF-1C	Elk-1
	HNF-1A	AR
	HNF-1B	AP-1

	GCF	c-Jun
	IRF-1	POU2F1
	NF-1	USF2
	Sp1	Ik-1
	TBP	HNF-4alpha
	IRF-2	LEF-1
	NF-kappaB1	AP-2alphaA
	HIF-1	COUP-TF1
	TCF-4	HOXD9
	PEA3	HOXD10
	NF-AT2	AhR:Arnt
	RelA	RAR-beta:RXR-alpha
	T3R-beta1	RXR-alpha
		RAR-beta
		AhR
		ATF3
		c-Ets-2
		POU2F2 (Oct-2.1)
		EBF
		PPAR-alpha:RXR-alpha
		HNF-1A
		NF-AT1
		E2F-1
		GATA-2
		GCF
		GATA-3
		NF-kappaB
		NF-kappaB1
		RelA
		SRF
		c-Fos
		T3R-beta1
		Sp1
		PU.1
		TBP
		IRF-2
		RAR-alpha1
		HIF-1
		TCF-4
		PEA3
		MAZ

After the filtration of ILINCS analysis with the Promo3 data has emerged a total of 25 chemical perturbagens that could have a possible effect on A549 cells like Fraisinib treatment, but with a direct effect on gene expression of the *GARS1* gene. From this data emerged that Treitoin, whose active pharmaceutical ingredient (API) is the retinoic acid (RA), was a possible drug trustful of

further examination, given its high availability. Furthermore, p53 was a target emerged from this analysis as regulated by the second drugs Quinacrine.

4.8 Retinoic acid treatment

To analyse the effect of RA on GARS1 promoter, the clone #27 cells were treated with RA 10 µM, in biological triplicate; the Luciferase activity and the AFC was measured after 24/48/72 hours. After collecting the data, the normalization between the luciferase and AFC emission was performed but the results were not significant. The reason could rely on the low efficacy of RA with respect to Treitoin or with effects of the drug on cell viability but independently from *GARS1* transcription.

4.9 Evaluation of p53 transactivation of the *GARS1* promoter

The pcNeo-p53 and pcNeo-empty were purified and amplified as descripted in materials and methods, to prepare an optimal quantitative of plasmids, and was checked with an elettrophoresis as shown in **Figure 28**.

Figure 28. The electrophoresis results, where A is pcNeo-empty no-digested, B the pcNeo-empty digested, C is pcNeo-p53 no-digested, D is pcNeo-p53 plasmid digested. The expected fragments size were: 5472 nt for pcNeo plasmid and 1065 for p53 insert.

To investigate the interaction between GARS1-promoter and P53, the experiment was projected to have:

1. a triplicate transfected with pcNeo-*p53*, pGl4.17-*GARS1*-prom, renilla CMV,
2. a triplicate transfected with pcNeo-empty, pGl4.17-*GARS1*-prom, renilla CMV,
3. pcNeo-*p53*, pG3b-*p21* activator, renilla CMV,
4. pcNeo-empty, pG3b-*p21* activator, renilla CMV.

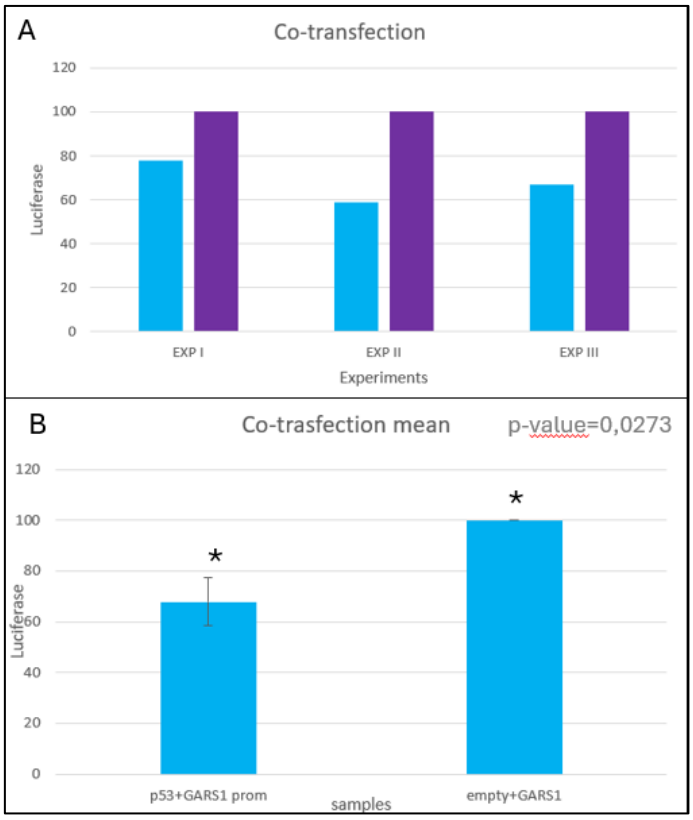

*Figure 29.**A**. is the visualization of the three experiments, where the blue columns represent the samples where is co-transfected GARS1-promoter and p53, while the purple columns represent the samples transfected with the empty-vector and GARS1. **B**. It is the co-transfection mean, with the p-value indicated. In both the figures the (*) indicates the statistical significance of the data.*

In all three experiments the activity of GARS1-promoter decreased in presence of p53 as observable in the **Figure 29**; statistical analysis by paired Student t test was performed.

5 CONCLUSIONS

In conclusion in this work, it has been possible to create a gene reporter system, based on the *GARS1*-promoter sequence, driving the luciferase gene, suitable future high-throughput screening of larger libraries of drugs.

The experimental approach of drug repositioning has discovered five compounds, butyrolactone 3, CTPB, B2, anacardic acid and MC-1293, still unknown to have a specific inhibition activity on the *GARS1* promoter and in NSCLC.

All the drugs resulted from the experimental approach were investigated in literature for having mutagenesis and toxicity effects already tested. Only anacardic acid was already tested and has showed an acute lethal dose in BALB/c mice which is higher tan 2000 mg/kg and has produced no mutagenesis effects on mice (Carvalho AL. et al.,2011).

From further investigation on the specific drugs has emerged that: Butyrolactone 3, CTPB and anacardic acid belongs to the class of histone acetyltransferase (HAT) inhibitors which inhibits the normal activity of histone acetylation of this class of enzyme that promotes chromatin relaxion and a higher gene expression, which in cancer could drive to a higher proliferation and progression to a more aggressive tumor stage.

B2 is a SIRT2 inhibitor where SIRT2 is a protein involved in different cellular pathways like DNA repairing, programmed cell death and cellular aging process, has been shown that SIRT2 inhibitors has a potential target in the fight of Alzheimer disease and the formation of amyloid plaque formation, and its hypothetical role versus cancer progression is furthermore trustful of future investigations.

MC-1293 is a histone deacetylase (HDAC) inhibitor that which can alter DNA accessibility and, consequently, gene expression, exerting an effect HATs like. Therefore, its role in the regulation of GARS1 remains even more intriguing and deserving of further investigation.

Furthermore the data collected from the epigenetic library screening will be useful for a more complex work based on the calculation of the z-factor, a statistical value that might be used for the evaluation of the High Throughput Screening (HTS), to determinate if this method is suitable to distinguish true positive and true negatives, to validate the experimental method, and use this assay for other bigger library of drugs.

Moreover, the in-silico approach, based on the gene signatures perturbation induced by Fraisinib in A549 cells, has produced at least twenty drugs which could have an interaction with GARS1 promoter and inhibits its expression similarly to Fraisinib in A549 cells as model of NSCLC. In this work only one of the chemicals perturbagens discovered with the iLINCS- Promo3 research, the active principle of Treitoin has been tested but failed to have effects on GARS1 transcription. Reasons could be effects on DNA regulatory regions upstream the sequence cloned, no effects

or low activity with respect to the Treitoin emerged by iLINCS, which has been shown a different cell toxicity and activity, but future studies to better understand its possible therapeutic role on A549 cell through the interaction with *GARS1*-promoter will be done.

Furthermore, thanks to the ILINCS-promo analysis has emerged a potential role of *p53* in the decreasing *GARS1* gene expression. This result is in accordance with p53 down-regulation in many cancers and in particular NSCLC. In this hypothesis, p53 downregulation could be a driver of *GARS1* overexpression in many cancers among which NSCLC (**Figure 30**).

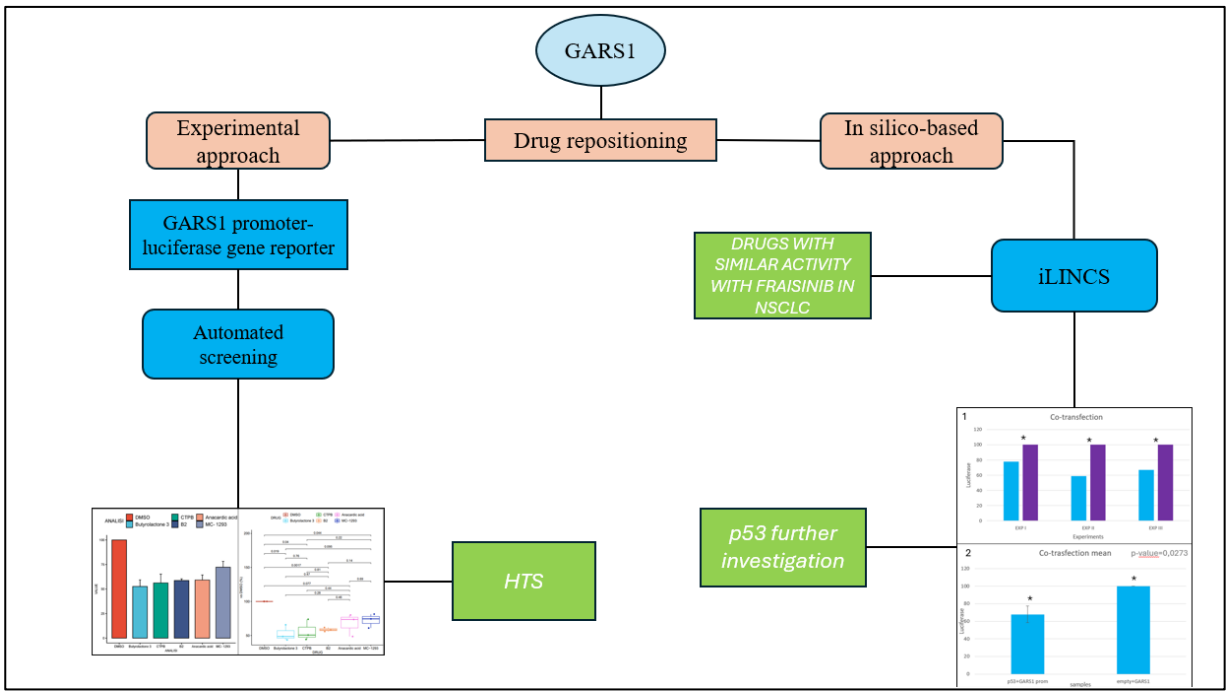

Figure 30. This image summarizes the main results obtained from the work, also highlighting the new perspectives generated by the analysis (green) which could lead to further investigation in NSCLC and beyond.

Finally, the interactome of *GARS1*, except for its role in protein synthesis, is today at study yet, and particularly in tumor context. In the last few years always more work shown is dysregulation in different cancer type and in different stage of them, indicating *GARS1* as a novel target for treatment and for further investigation for its role in tumor proliferation and in tumor microenvironment.

6 REFERENCES

1. Alexandrova J, Paulus C, Rudinger-Thirion J, Jossinet F, Frugier M. Elaborate uORF/IRES features control expression and localization of human glycyl-tRNA synthetase. RNA Biol. 2015;12(12):1301-13. doi: 10.1080/15476286.2015.1086866. PMID: 26327585; PMCID: PMC4829322.
2. Ben Toumia, Iméne, Tiziana Bachetti, Leila Chekir-Ghedira, Aldo Profumo, Marco Po nassi, Alessandro Di Domizio, Alberto Izzotti, et al. “Fraisinib: A Calixpyrrole Derivat ive Reducing A549 Cell-Derived NSCLC Tumor in Vivo Acts as a Ligand of the Glyci ne-TRNA Synthase, a New Molecular Target in Oncology.” Frontiers in Pharmacology 14 (January 3, 2024). https://doi.org/10.3389/fphar.2023.1258108.
3. Carvalho AL, Annoni R, Silva PR, Borelli P, Fock RA, Trevisan MT, Mauad T. Acute, subacute toxicity and mutagenic effects of anacardic acids from cashew (Anacardium occidentale Linn.) in mice. J Ethnopharmacol. 2011 Jun 1;135(3):730-6. doi: 10.1016/j.jep.2011.04.002. Epub 2011 Apr 12. PMID: 21511024.
4. Delavan, Brian, Ruth Roberts, Ruili Huang, Wenjun Bao, Weida Tong, and Zhichao Li u. “Computational Drug Repositioning for Rare Diseases in the Era of Precision Medic ine.” Drug Discovery Today 23, no. 2 (February 2018): 382–94. https://doi.org/10.101 6/j.drudis.2017.10.009.
5. Furukawa, Ryutaro, Shin-ichi Yokobori, Riku Sato, Taimu Kumagawa, Mizuho Nakag awa, Kazutaka Katoh, and Akihiko Yamagishi. “Amino Acid Specificity of Ancestral Aminoacyl-TRNA Synthetase Prior to the Last Universal Common Ancestor Common ote Commonote.” Journal of Molecular Evolution 90, no. 1 (January 27, 2022): 73–94. https://doi.org/10.1007/s00239-021-10043-z.
6. Joerger, Andreas C., and Alan R. Fersht. “The P53 Pathway: Origins, Inactivation in C ancer, and Emerging Therapeutic Approaches.” Annual Review of Biochemistry 85, no . 1 (June 2, 2016): 375–404. https://doi.org/10.1146/annurev-biochem-060815-014710.
7. Juárez-López, Daniel, and Alejandro Schcolnik-Cabrera. “Drug Repurposing: Consider ations to Surpass While Re-Directing Old Compounds for New Treatments.” Archives of Medical Research 52, no. 3 (April 2021): 243–51. https://doi.org/10.1016/j.arcmed.2 020.10.021.
8. Khosh Kish E, Gamallat Y, Choudhry M, Ghosh S, Seyedi S, Bismar TA. Glycyl-tRNA Synthetase (GARS) Expression Is Associated with Prostate Cancer Progression and Its Inhibition Decreases Migration, and Invasion In Vitro. Int J Mol Sci. 2023 Feb 21;24(5):4260. doi: 10.3390/ijms24054260. PMID: 36901698; PMCID: PMC10001614.

9. Levine, Arnold J., and Moshe Oren. “The First 30 Years of P53: Growing Ever More C omplex.” Nature Reviews Cancer 9, no. 10 (October 2009): 749–58. https://doi.org/10. 1038/nrc2723.
10. Mudge SJ, Williams JH, Eyre HJ, Sutherland GR, Cowan PJ, Power DA. Complex organisation of the 5'-end of the human glycine tRNA synthetase gene. Gene. 1998 Mar 16;209(1-2):45-50. doi: 10.1016/s0378-1119(98)00007-9. PMID: 9524218.
11. Park MC, Goughnour PC, Jun S, Cho S, Song E, Kim SB, Kim HY, Hyun JK, Kim P, Jung HS, Kim S. Two distinct receptor-binding domains of human glycyl-tRNA synthetase 1 displayed on extracellular vesicles activate M1 polarization and phagocytic bridging of macrophages to cancer cells. Cancer Lett. 2022 Jul 28;539:215698. doi: 10.1016/j.canlet.2022.215698. Epub 2022 May 9. PMID: 35523311.
12. Parvathaneni, Vineela, Nishant S. Kulkarni, Aaron Muth, and Vivek Gupta. “Drug Rep urposing: A Promising Tool to Accelerate the Drug Discovery Process.” Drug Discove ry Today 24, no. 10 (October 2019): 2076–85. https://doi.org/10.1016/j.drudis.2019.06. 014.
13. Pilarczyk, Marcin, Mehdi Fazel-Najafabadi, Michal Kouril, Behrouz Shamsaei, Juozas Vasiliauskas, Wen Niu, Naim Mahi, et al. “Connecting Omics Signatures and Revealin g Biological Mechanisms with ILINCS.” Nature Communications 13, no. 1 (August 9, 2022). https://doi.org/10.1038/s41467-022-32205-3.
14. Reck, M., S. Popat, N. Reinmuth, D. De Ruysscher, K.M. Kerr, and S. Peters. “Metasta tic Non-Small-Cell Lung Cancer (NSCLC): ESMO Clinical Practice Guidelines for Di agnosis, Treatment and Follow-Up.” Annals of Oncology 25 (September 2014): iii27–39. https://doi.org/10.1093/annonc/mdu199.
15. Seve, Pascal, and Charles Dumontet. “Chemoresistance in Non-Small Cell Lung Cance r.” Current Medicinal Chemistry-Anti-Cancer Agents 5, no. 1 (January 1, 2005): 73–88 . https://doi.org/10.2174/1568011053352604.
16. Shimizu H, Nakayama KI. A 23 gene-based molecular prognostic score precisely predicts overall survival of breast cancer patients. EBioMedicine. 2019 Aug;46:150-159. doi: 10.1016/j.ebiom.2019.07.046. Epub 2019 Jul 26. PMID: 31358476; PMCID: PMC6711850.
17. Sung, Yulseung, Ina Yoon, Jung Min Han, and Sunghoon Kim. “Functional and Pathol ogic Association of Aminoacyl-TRNA Synthetases with Cancer.” Experimental & Molecular Medicine 54, no. 5 (May 2, 2022): 553–66. https://doi.org/10.1038/s12276-022-00765-5.

18. Toumia, I. B., Ponassi, M., Barboro, P., Iervasi, E., Vargas, G. C., Banelli, B., Fiordoro , S., Ghedira, L. C., Kohnke, F. H., Izzotti, A., & Rosano, C. (2022). Two calix[4]pyrro les as potential therapeutics for castration-resistant prostate cancer. Investigational Ne w Drugs, 40(6), 1185–1193.
19. W. He, H. Zhang, Y.E. Chong, M. Guo, A.G. Marshall, & X. Yang, Dispersed disease-causing neomorphic mutations on a single protein promote the same localized conformational opening, Proc. Natl. Acad. Sci. U.S.A. 108 (30) 12307-12312, https://doi.org/10.1073/pnas.1104293108 (2011).
20. Wang J, Yang B, Wang D, Han R, Bi Z, Lin L. GARS is implicated in poor survival and immune infiltration of hepatocellular carcinoma. Cell Signal. 2022 Jun;94:110302. doi: 10.1016/j.cellsig.2022.110302. Epub 2022 Mar 7. PMID: 35271987.
21. Wu, Patrick, QiPing Feng, Vern Eric Kerchberger, Scott D. Nelson, Qingxia Chen, Bingshan Li, Todd L. Edwards, et al. “Integrating Gene Expression and Clinical Data to Identify Drug Repurposing Candidates for Hyperlipidemia and Hypertension.” Nature Communications 13, no. 1 (January 10, 2022). https://doi.org/10.1038/s41467-021-27751-1

7 Acknowledgements

I would like to sincerely thank Dr. Tiziana Bachetti for giving me the opportunity to take part in this thesis project and for her continuous support and guidance throughout every phase of the work, which made this experience truly enriching.

I am also deeply grateful to Dr. Chiara Cervetto and Dr. Elena Grasselli for their valuable advice and assistance during the final stages of the project, which allowed me to further enhance the content of this work.

Finally, I would like to thank all the members of the laboratory of Proteomics and Mass Spectrometry Laboratory of the Advanced Biotechnology Center (CBA) at IRCCS Ospedale Policlinico San Martino (Camillo, Marco, Aldo, Erika, Gabriela, Katerina, and Deianira) for warmly welcoming me from the very beginning and for helping me bring this project to completion.