University of Nottingham
UK | CHINA | MALAYSIA

Division of Advanced Drug Delivery and Tissue Engineering

School of Pharmacy

Understanding the roles and impact of Solute Carrier Transporters OCT1, OCTN1, and OCTN2 in lung epithelia

Paulyna G. Magaña BSc, MSc

Thesis submitted to the University of Nottingham for the completion of Doctor of Philosophy degree

August 2023

ABSTRACT

The organic cation transporters (OCT) 1, 2, and 3, along with the novel organic cation transporters (OCTN) 1 and 2, are members of the Solute Carrier 22 (SLC22) family, these transporters play significant roles in mediating the intracellular movement of endogenous compounds, including neurotransmitters, L-carnitine, and ergothioneine. Furthermore, numerous drugs used in the management of respiratory disorders exhibit cationic properties at physiological pH, making them potential substrates of OCT/Ns. Additionally, the involvement of OCT/Ns in lung diseases, suggests their possible relevance as novel drug targets Moreover, these proteins significantly influence the efficacy of specific anti-cancer drugs, as the transport of these compounds is mediated by OCT/Ns.

This thesis aimed to investigate the expression of OCT1/SLC22A1, OCTN1/SLC22A4, and OCTN2/SLC22A5 in lung epithelia, with a focus on understanding their roles in various lung pathophysiology. Expression profiles of these transporters were analysed in asthma, chronic obstructive pulmonary disease (COPD), lung adenocarcinoma, and lung squamous carcinoma, where OCTN1 and OCTN2 were downregulated in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) using datasets from TCGA and GEO. Analysis revealed a downregulation of OCTNs in lung malignancies, whereas OCTN2 was only found differentially expressed in COPD samples.

Functional impact of OCT/OCTNs was explored using CRISPR-Cas9 to target SLC22 genes in HEK293 cells. Migration, proliferation, and adhesion assays on edited cells suggested a role for SLC22 genes in cell development.

Regulatory mechanisms of OCTN1 expression were investigated, suggesting a potential role for RUNX1 transcription factor binding within the promoter region.

Overall, this work sheds light on the expression patterns and functional roles of SLC22A1, SLC22A4, and SLC22A5 in lung epithelia under various pathological conditions. However, further research is necessary to fully elucidate the mechanisms of gene regulation in lung cells and their relationship to lung cancer development.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my funders, the CONAHCYT and the University of Nottingham, for providing the means to carry out this project. I am sincerely grateful to my supervisors, Cynthia Bosquillon and Sebastiaan Winkler, for their unwavering support throughout my Ph.D. journey. Their guidance and mentorship have been invaluable, particularly during the challenging times presented by the COVID pandemic. Their expertise and encouragement have been instrumental in shaping this research.

I extend my heartfelt appreciation to my family for their immense support throughout this endeavour. To my brother, thank you for providing the motivation and push when I needed it the most, and for teaching me the importance of trusting the process. To my sister, thank you for always being by my side, offering your unwavering support, and helping me navigate this journey.

To my mum, you have been an incredible anchor of support. Your encouragement, perseverance, and belief in me have been an inspiration. I am deeply grateful for having you as the best role model of a strong and intelligent woman.

To my dearest friends, thank you for standing by my side and providing the much-needed moments of respite during this intense journey. Your understanding and encouragement have kept me going even during the most challenging times. I extend my gratitude to everyone who has lent their support throughout this process. Your help and guidance have been instrumental in

bringing this thesis to fruition. Without your contributions, this work would not have been possible.

I also wish to express my thanks to the GRRB team, Ryan, Fahad, Angie, and Luc, for their assistance and support. I am indebted to Chris for his invaluable help with cloning and to Hilary for her assistance with transfection. Special thanks to Sasha for providing valuable insights into CRISPR/Cas9 and for helping me develop my gene editing skills.

I would also like to express my gratitude to the faculty and staff at University of Nottingham, whose support and resources have been instrumental in shaping my academic and research experience.

While these acknowledgments capture some of the key contributors, I am indebted to many others whose influence and support have left an indelible mark on this journey. Everyone mentioned here has played an integral role in shaping this research and my growth as a scholar. I am humbled and grateful for their contributions, and I carry the knowledge and experiences gained during this journey into future endeavours with immense appreciation.

TABLE OF CONTENTS

ABBREVIATIONS

ABC	ATP-Binding Cassette
ADME	Drug Absorption, Distribution, Metabolism and Excretion
ASL	Airway Surface Liquid
ATCC	American Type Culture Collection
BCRP	Breast cancer Resistance Protein
Calu-3 cells	Human lung cancer cell line
COPD	Chronic Obstructive Pulmonary Disease
CRISPR-Cas9	Clustered Regularly Interspaced Short Palindromic Repeat- associated protein (Cas) system
DSB	Double-strand Break
DEG	Differentially Expressed Genes
E. coli	Escherichia coli
Fwd	Forward
GEO	Gene Expression Omnibus
HEK 293 cells	Human embryonic kidney 293 cells
ILP	Isolated lung perfusion
LB	Luria-Bertani
LLF	Lining Fluid in the Lung
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
min	minutes
OAT	Organic Anion Transporter
OATP	Organic Anion Transporting Polypeptide
OCT	Organic Cation Transporter

OCTN	Organic Cation and Carnitine Transporter
pX459	pSpCas9(BB)-2A-Puro plasmid
RNAi	RNA interference
s	seconds
sgRNA	single guide RNA
siRNA	Small interference RNA
SLC	Solute Carrier
TCGA	The Cancer Genome Atlas (TCGA)

LIST OF TABLES

LIST OF FIGURES

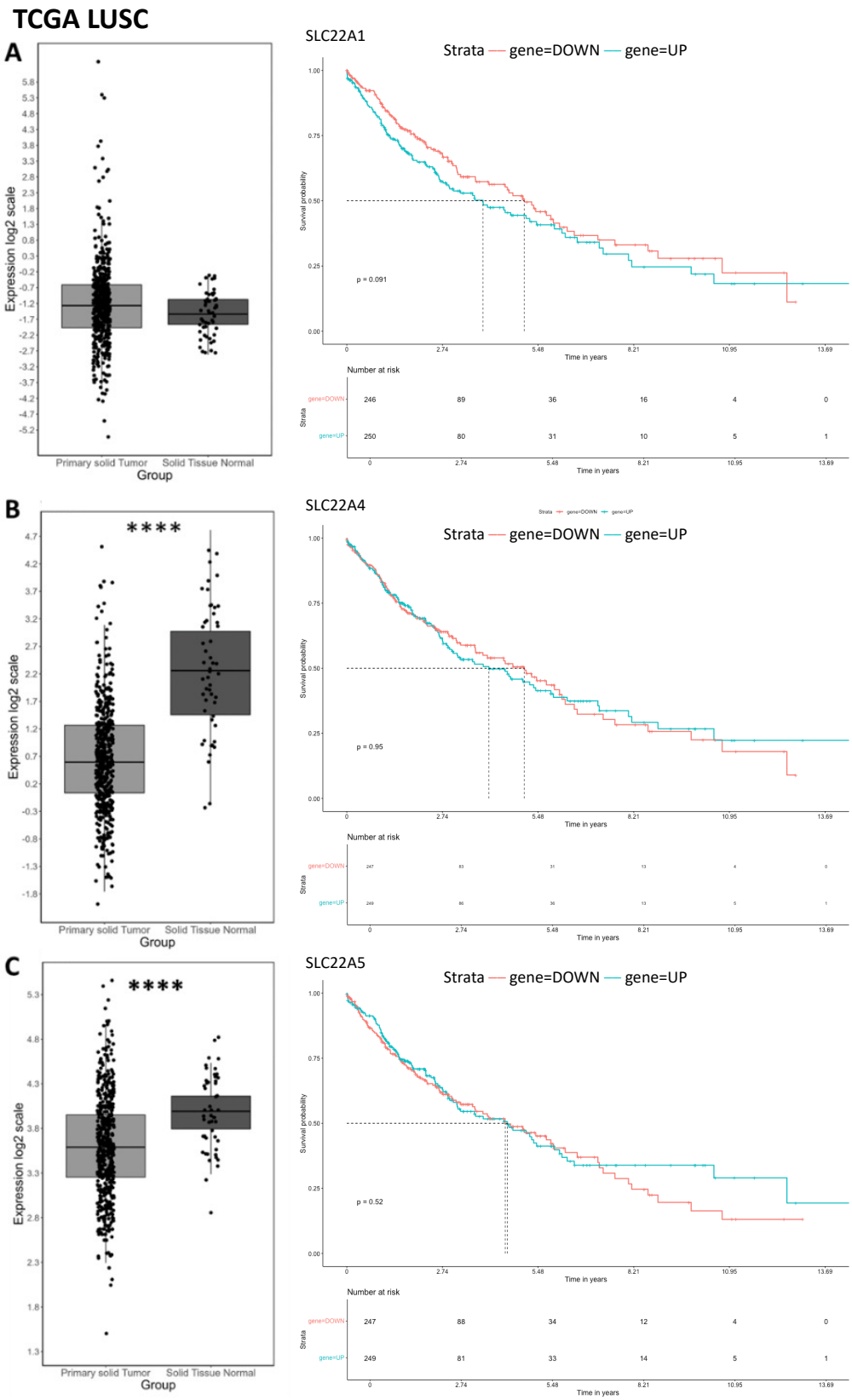


FIGURE 3-8. EXPRESSION AND SURVIVAL ANALYSES OF THE THREE SLC22A GENES IN LUNG SQUAMOUS CELL CARCINOMA (LUSC). (A-C). EXPRESSION OF SLC22A1, SLC22A4 AND SLC22A5 IN LUSC TISSUES AND ADJACENT NON-TUMOUR TISSUE. (D-F) KAPLAN-MEIER

CHAPTER 1
GENERAL INTRODUCTION

CHAPTER 1. GENERAL INTRODUCTION

Drug transporters are transmembrane proteins that facilitate the flux of molecules through the cell membrane. These proteins play a critical role in drug absorption, distribution, metabolism and excretion (ADME) (Nicholls et al., 2016). Drug transporters are currently a topic of intense study due to their important role in the successful development of drug candidates. However, research has largely been focused on transporters in the brain and the main organs involved in drug distribution (intestine, liver and kidney),(Bosquillon, 2010) despite reports that drug transporters play an important role in the drug absorption in the respiratory epithelium (Bosquillon, 2010; Gumbleton et al., 2011). Moreover, several drugs used in the treatment of respiratory disorders are cations at physiological pH and potential substrates of OCT/Ns.

The inhaled route is of interest in the pharmaceutical field not only for the treatment of respiratory diseases but also as an alternative for drug administration to the bloodstream. It has become an attractive non-invasive route to administer drugs as it overcomes first-pass metabolism while improving bioavailability with high local drug concentrations.(Patil & Sarasija, 2012) Furthermore, the lungs receive ample blood supply, and the pulmonary epithelium has high permeability and a large surface area (approximately 70-140 m2)(Groneberg et al., 2003) with low enzymatic activity (Hou et al., 2015; Ibrahim & Garcia-Contreras, 2013). While advances have been made in the field of inhalation product development, especially in the formulation of dry powders and inhalation systems,(Pilcer & Amighi, 2010) further improvements have been

hampered by a lack of understanding of the mechanisms underlying mechanisms of protein transporters and drug absorption in the lungs.

1.1. Lung Structure and Function

The main function of the lung is to perform gaseous exchange with the environment by delivering oxygen and removing carbon dioxide from the blood; which occurs in the alveolar region (parenchyma) (Chaudhry & Bordoni, 2019). The respiratory tract can be divided into two major parts, airway anatomy and lung anatomy, both of which are organised into a tree-like system which extends to a terminal zone where the actual exchange of gases occur. As the airways progress, they assume different structural features as well as a decrease in diameter, as shown in Figure 1-1.

Figure 1-1. Structure of the human respiratory tract, based on the Weibel model (Weibel, 1963). As the conducts progress, they gradually become smaller in diameter before reaching the alveoli. Created with BioRender.com

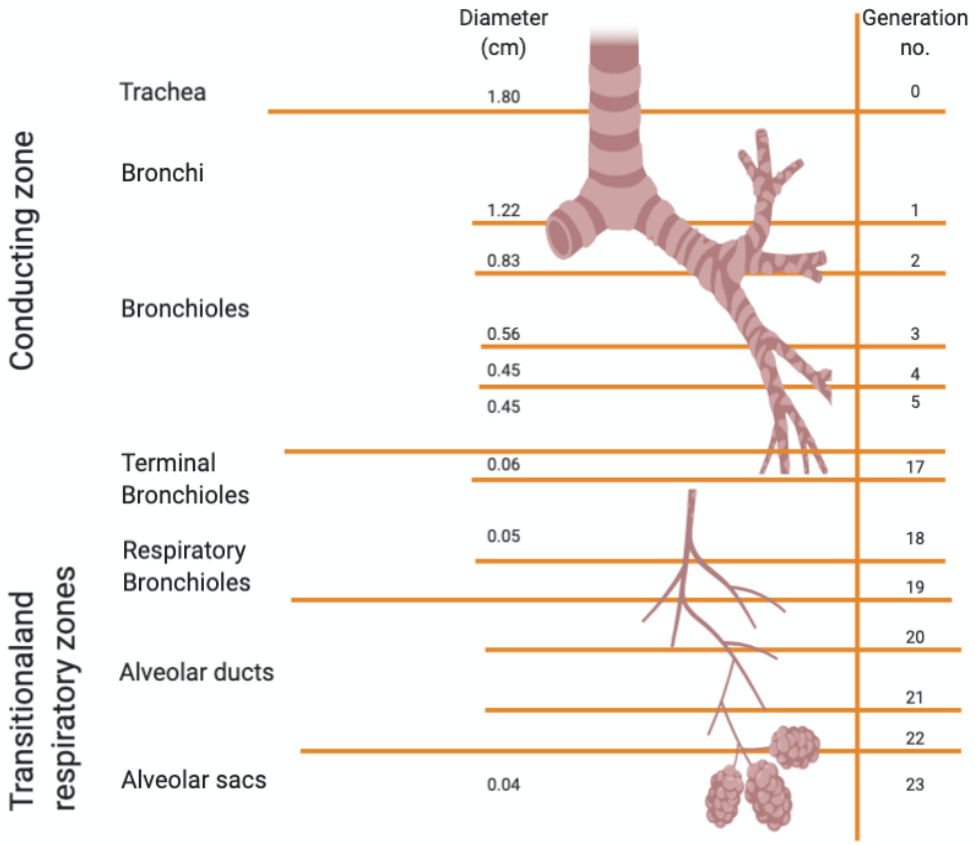


Figure 1-1. Structure of the human respiratory tract, based on the Weibel model (Weibel, 1963). As the conducts progress, they gradually become smaller in diameter before reaching the alveoli. Created with BioRender.com

The airway anatomy includes the glottis regions, trachea, mainstem bronchi and multiple bronchial generations. These structures filter, humidify, and warm the inhaled air while conducting it to the alveolar surface. They bifurcate roughly 16-18 times before reaching the respiratory zone (Figure 1-1). The lung anatomy also includes the parenchyma, where the respiratory exchange takes place, and is further subdivided into the alveoli, alveolar ducts and respiratory bronchioles (generations 17-23 in the airway model Figure 1-1). The alveoli are the functional unit of gas exchange and they are in intimate contact with blood vessels (Knudsen & Ochs, 2018). Over 300 million alveoli exist in the human lung, with a surface area of approximately 70 m^2, comprising about 90% of the

lung volume (Chaudhry & Bordoni, 2019; Plopper, 1996). In comparison to the total mucosal surface of the digestive tract interior of around 32 m^2 (Helander & Fändriks, 2014).

1.1.1. Epithelial lining fluid

The lining fluid in the lung (LLF) has a heterogeneous thickness that decreases gradually along the lung epithelial anatomy in the alveoli (Patton, 1996). The LLF lubricates and protects the underlying pulmonary epithelium. Its composition varies from the larger to the smaller airways, mainly consisting of water (96%), proteins, salts, phospholipids and mucins. In the large airways the fluid resembles a mucus-like solution, while in the alveolar region consists of LLF and surfactant (phospholipids and proteins) (Bicer et al., 2012; Moliva et al., 2014; Olsson et al., 2011).

Therefore, lining fluid and binding proteins in the different parts of the pulmonary epithelium are relevant parameters for the availability and absorption rates of inhaled drugs (D. A. Smith et al., 2010). The fluid can induce aggregation of the particles or enhance solubility of small lipophilic drugs (Patton, 1996). Hence, delivering therapeutic drugs to the lung requires a deeper understanding of the biological and physiological system. The selected model for absorption studies should also replicate the mucus-like layer that affects the lung deposition of inhaled compound and pharmacokinetic profiles.

1.1.2. Lung Epithelium

As a result of air passage through the respiratory system, the lung is a target for airborne particulates, pathogens and pollutants. The pulmonary epithelium is

lined with distinct cellular populations that fulfil different functions host defence, hydration or gas exchange (Plopper, 1996). The vast majority of insoluble particles (>6 μm) become trapped in a mucus lining produced by secretory cells and the movement of cilia of the airways (Rackley & Stripp, 2012).

The lung consists of over 40 different types of cells, including lymphocytes, mast cells and leukocytes (Franks et al., 2008; McDowell et al., 1978). The cellular composition and diversity of the respiratory epithelium varies according to the airway region (Figure 1-2).

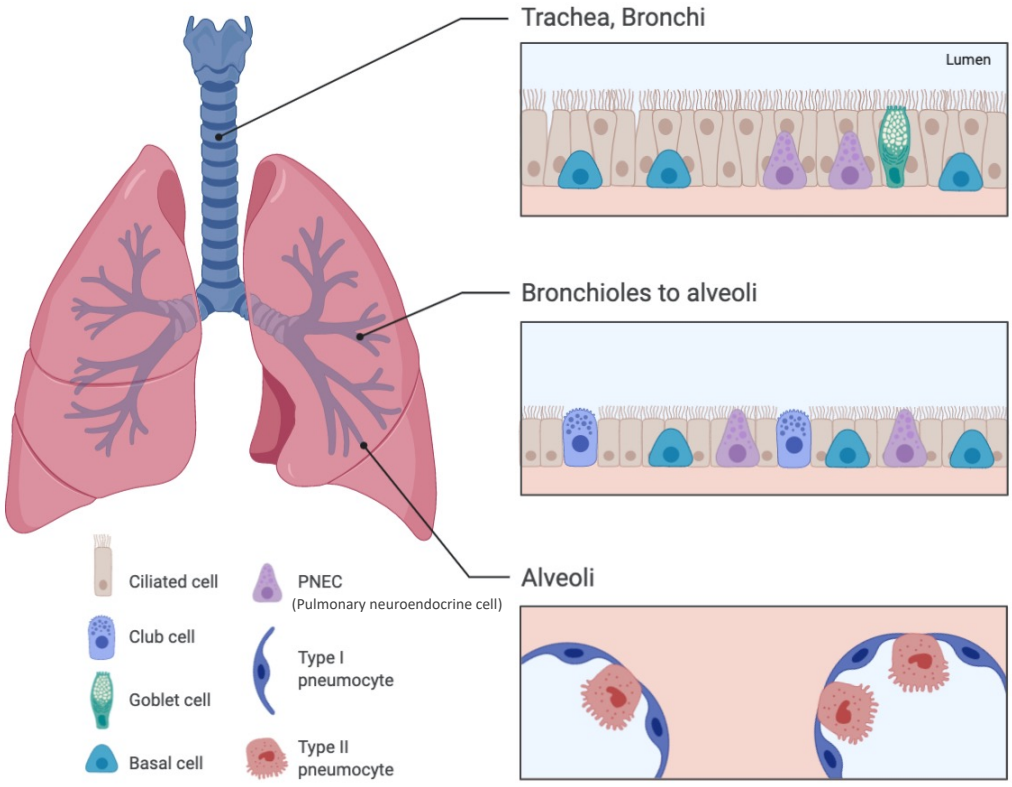


Figure 1-2. Schematic representation of a healthy respiratory epithelium. The respiratory epithelium is mainly composed of ciliated cells, goblet cells, brush cells and basal cells. Goblet cells become fewer further down the respiratory tract and are absent in the bronchioles where club cells are then present. Pulmonary neuroendocrine cells (PNECs) harbour neuronal and endocrine characteristics. Pneumocytes line the alveolar compartment of the lungs (McDowell et al., 1978; Pinkerton et al., 2015; Plopper, 1996; Reynolds et al., 2015). Adapted from (Camelo et al., 2014), created with BioRender.com

The most distal section of the airway epithelium or bronchioles has one of the highest degrees of diversity between species of any region of the lung (Franks et al., 2008; Plopper, 1996; Plopper & Hyde, 2015). The bronchiolar epithelium is a simple cuboidal epithelium, mainly composed of ciliated cells and Club cells, also known as non-ciliated non-mucous secretory cells. Its main role is to provide physical protection to the lung by secreting surfactant (Rokicki et al., 2016). This allows bronchioles to expand during inspiration and prevents their collapse during expiration (*Polydisperse Microparticle Transport and Deposition to the Terminal Bronchioles in a Heterogeneous Vasculature Tree | Scientific Reports*, n.d.).

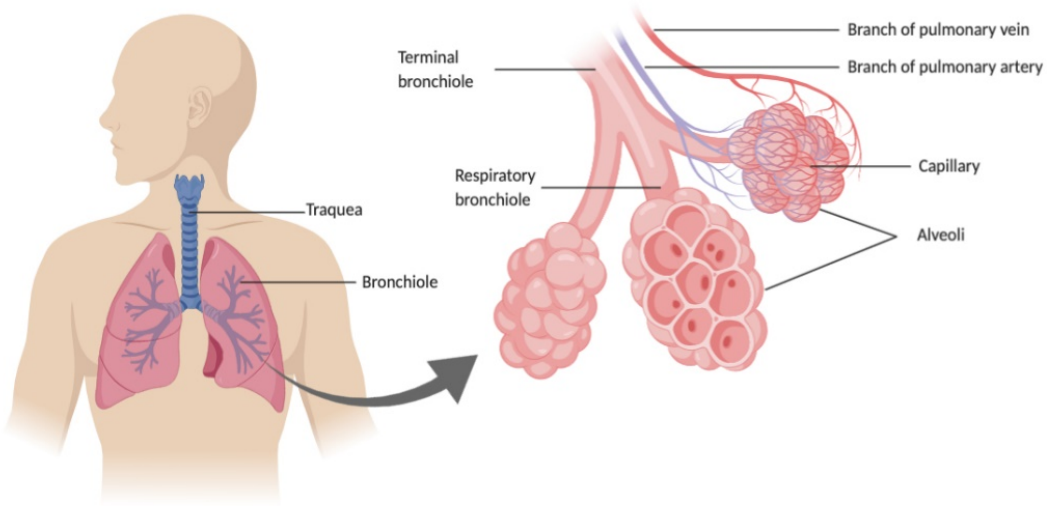


Figure 1-3. Anatomy of the human lung. The airways bifurcate in a tree-like system before reaching the respiratory zone. Capillaries cover 70% of the outside of alveoli, providing a large surface area for gases to diffuse across. Created with BioRender.com

In the alveolar area, in addition to macrophages, two other epithelial cell types can be found. Squamous or *type I alveolar* epithelial cells cover approximately 95% of the surface area, have an average thickness of 0.26 µm (Plopper, 1996; Sporty et al., 2008), and constitute the air-blood barrier. The remainder of the interalveolar surface is covered by the granular alveolar epithelial cells or *type II alveolar* cells, which synthesize and secrete pulmonary surfactants that line alveolar air spaces, reducing surface tension (Pinkerton et al., 2015; Plopper, 1996; Sporty et al., 2008). These sac-like structures are surrounded by a mesh like network of thin blood vessels (capillaries), as seen in Figure 1-3.

1.2. Membrane Transport proteins

Transporters can generally be classified into two major superfamilies. The solute carrier family (SLC) is involved in the influx and efflux of substrates, while the ATP-binding cassette (ABC) family actively secretes compounds out

of the cell (Figure 1-4). The ABC transporters are primary active transporters that utilize energy from ATP hydrolysis to transport substrates across the cell membrane. The best known and studied of these is P-glycoprotein (P-gp), encoded by the multidrug resistance protein (MDR1/ABCB1), the multidrug resistance associated protein 1 (MRP1/ABCC1) and breast cancer resistance protein (BCRP) (K. G. Chen & Sikic, 2012; Z. Chen et al., 2016; Grandjean-Forestier et al., 2009; Hong, 2017; International Transporter Consortium et al., 2010). SLC transporters are the largest family of transmembrane transporters with over 400 genes identified in the human genome. They can either facilitate diffusion across the membrane; or act as a secondary active transporter in that they rely on ion gradients generated by ATP-dependent pumps to transport substrates against the concentration gradient.

Most drug transporters belong to the SLC transporter family. Many of its members play key roles in human disease as well as in ADME properties in drug discovery (L. Lin et al., 2015; Y. Zhang et al., 2019), making them important targets in drug therapy (Qosa et al., 2016). Like other barriers in the body, the lung expresses a variety of ABC and SLC transporters, though expression levels are highest in the liver and kidney (Bleasby et al., 2006). Research thus far has mainly focused on transporters in the intestine, liver, kidney, brain and their relevance to drug disposition in those organs (Bosquillon, 2010).

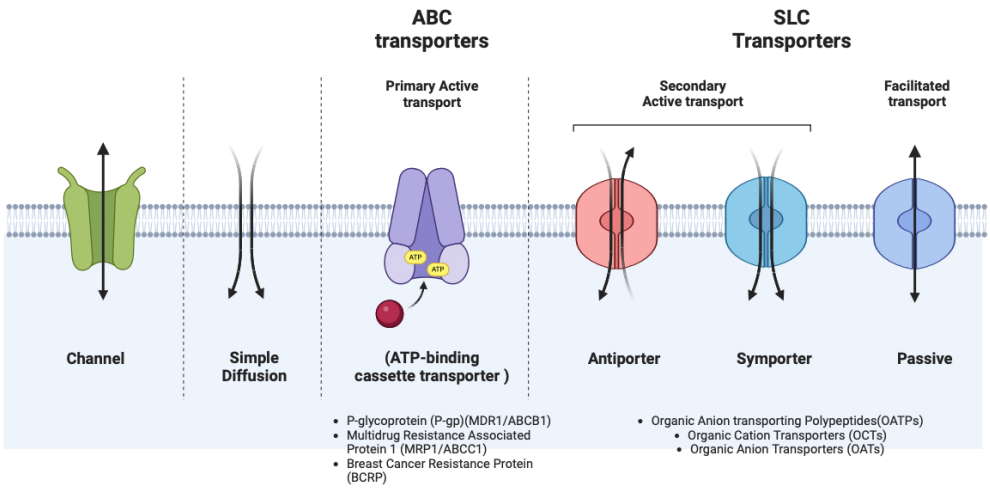


Figure 1-4. General classification of Solute Carriers (SLC), ATP-dependent carriers (ABC-transporter), in the plasma membrane. ABC transporters use the energy of ATP binding and hydrolysis to facilitate transport cross membranes, whereas SLC transporters use the ion electrochemical gradients. Created with BioRender.com

The pulmonary epithelium has been demonstrated to be a more permeable barrier to drugs, with a higher molecular polar surface area than the intestinal mucosa and blood-brain barrier (Bosquillon, 2010). Yet, inhaled drugs remain poorly investigated with regard to their absorption, distribution and elimination processes in the lung.

1.2.1. Solute Carrier Link/ Organic Cation Transporters

There are more than 400 SLC transport proteins with over 60 families that are classified based on their biological function and substrate. They interact with a broad array of substrates including amino acids, peptides, inorganic and organic ions, ad range in selectivity from transporters that interact with a narrow group of substrates to those that accept a diverse range of chemically diverse substrates. Further, transporters within a particular family can differ in specificity (L. Lin et al., 2015). For example, the SLC38 family consists of 11 transmembrane

proteins with different substrate specificity. SNAT1, SNAT2, SNAT4 prefer small neutral amino acids, whereas SNAT3 and SNAT5 favour glutamine, asparagine and histidine (Bröer, 2014) .

This MSA was employed to construct a phylogenetic tree of the SLC22 family using the ape package (version 5.7-1) with ClustalW alignment (Figure 1-6). The phylogenetic tree depicts the evolutionary relationships among members of the SLC22A transporter family based on amino acid sequence homology. The tree was constructed using using a neighbor-joining algorithm. The major clades are designated as OCT (Organic Cation Transporters) and OAT (Organic Anion Transporters). The tree reveals several subclades within each major clade, potentially reflecting functional specializations of the transporters. SLC22A18 appears as a distinct branch, diverging from the main cluster of SLC22A transporters. SLC22A31, SLC22A23, and a group related to OAT transporters form a separate subclade within the major OCT clade. SLC22A17 clusters with SLC22A13 and SLC22A14 within the OAT-like subclade. Several transporters (SLC22A24, SLC22A10, SLC22A25, SLC22A9) are interspersed along the major branches of the tree, suggesting a less defined evolutionary relationship with other family members. The majority of the SLC22A transporters (SLC22A12, SLC22A11, SLC22A8, SLC22A6, SLC22A15, SLC22A16, SLC22A5, SLC22A4, SLC22A3, SLC22A2, SLC22A1) cluster within the major OCT clade.

.

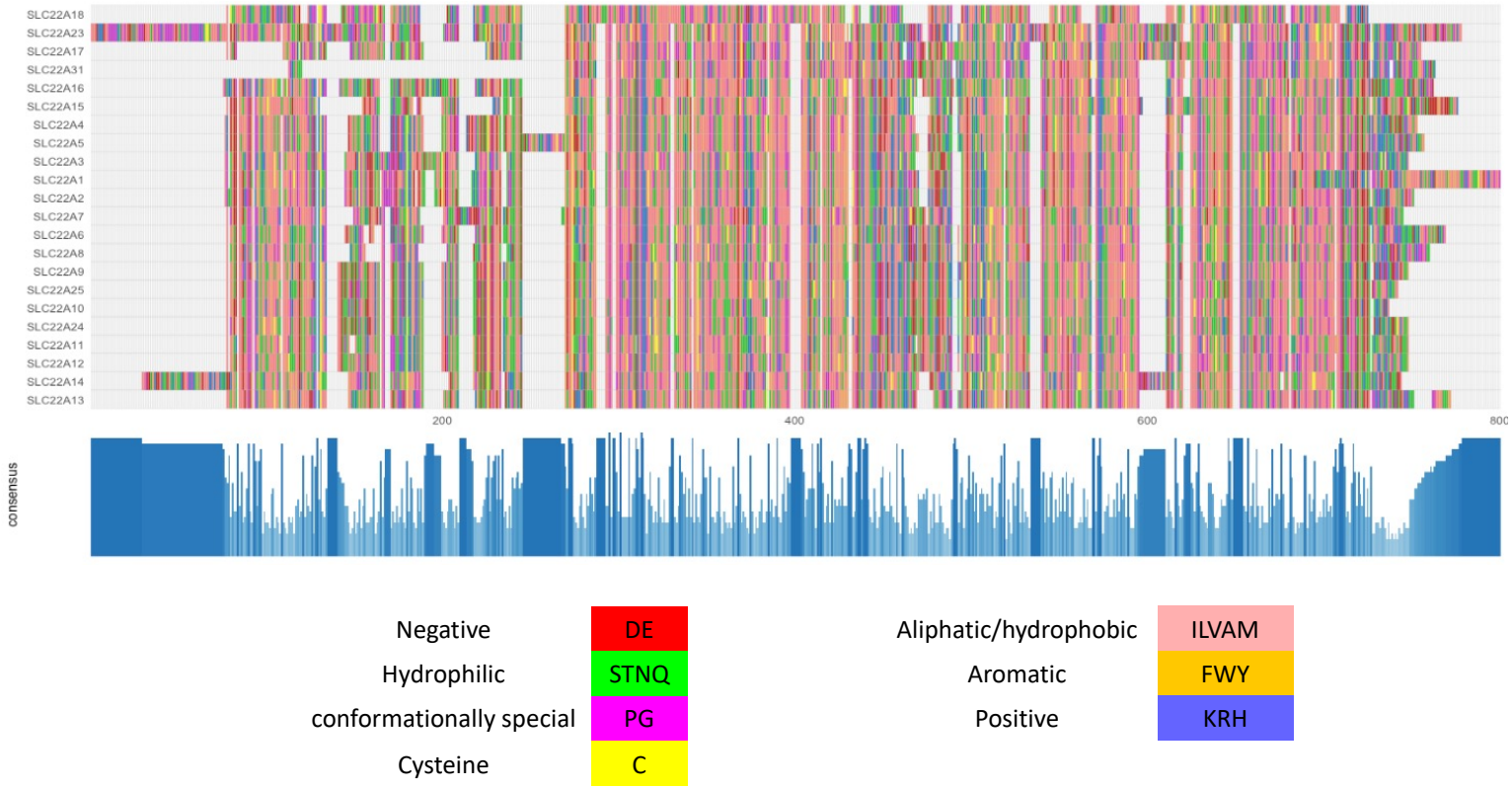


Figure 1-5. Multiple sequence alignment of SLC22 amino acid sequence aligned using ClustalW, constructed in R/Bioconductor using msa v1.32.0 and ggmsa package v1.6.0, coloured according to their physicochemical properties. Blue bars show sequence conservation.

The SLC superfamily appears understudied, and structural characterization of human proteins is poorly understood (César-Razquin et al., 2015). The SLC22 family is one of the most studied SLC transporters in the lung. The SLC22 family, contains the organic cation transporters (OCTs), carnitine/organic cation transporters (OCTNs) and organic anion transporters (OATs) (DeGorter et al., 2012; Hong, 2017; You & Morris, 2014). These function as uniporters that facilitate diffusion in either direction (OCTs), as anion exchangers (OAT1, OAT3 and URAT1), and as sodium/carnitine transporters (OCTN). A multiple sequence alignment (MSA) of SLC transporters was generated using R software (version 2023.12.1) (Figure 1-5). SLC transporters are known to possess 12 transmembrane domains, and due to their role as membrane-embedded solute carriers, their sequences exhibit a high degree of conservation, particularly within regions defined by the physicochemical properties of their constituent amino acids.

In 2010, The International Transporter Consortium (ITC) highlighted the clinical importance of OCTs transporters in drug disposition, and recommended in vitro assays to study transporter-drug interactions (International Transporter Consortium et al., 2010)

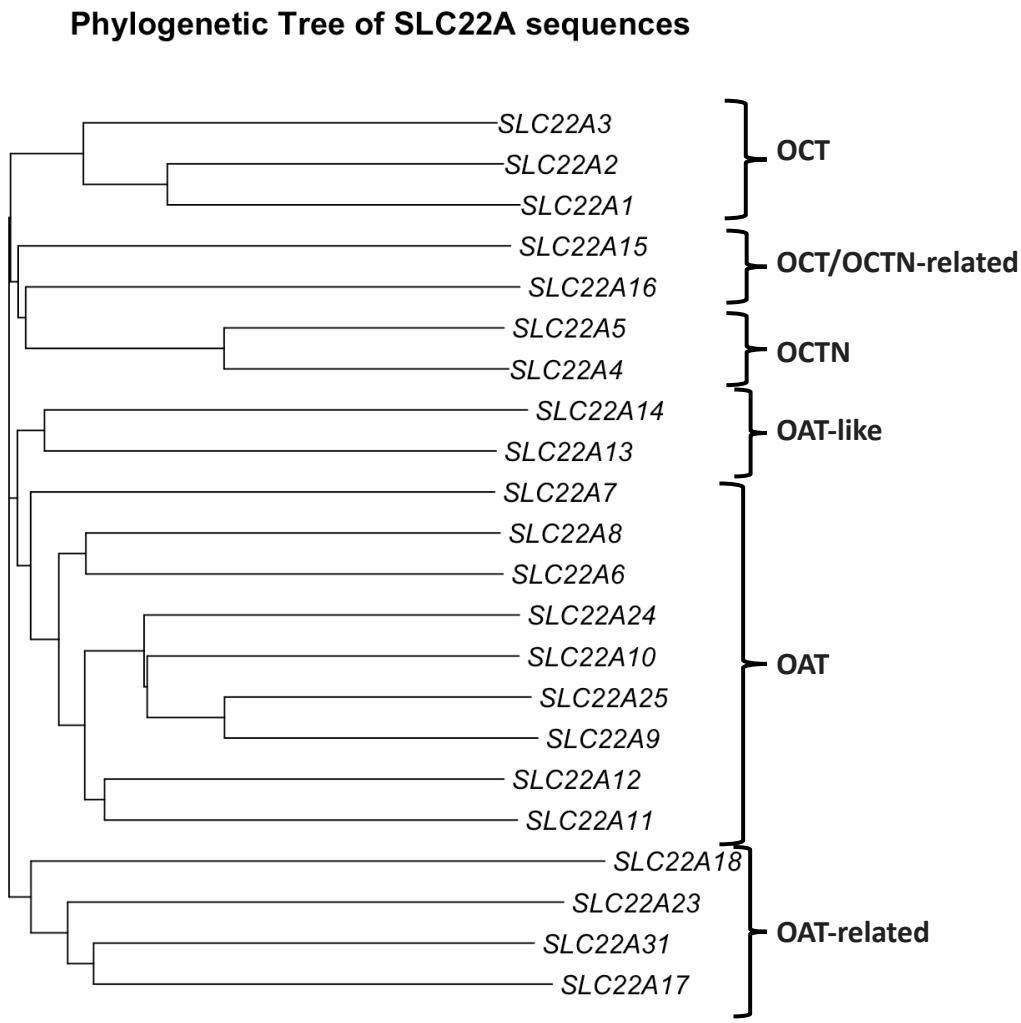


Figure 1-6. Subfamilies of SLC22 transporters. Phylogenetic tree was build using canonical sequences from UniProt, multiple sequence analysis was performed using ClustalW (Edgar, 2004). The SLC22 family consist of two major clades: OAT (Organic Anion Transporter) and OCT (Organic Cation Transporter) which can be further subclustered in clades, as OAT, OAT-like, OAT-related, OCT, OCTN (Novel Organic Cation Transporter), and OCT/OCTN-related, according to (Nigam, 2018).

OCTs are facilitative diffusion systems; they translocate a range of organic cations in a bidirectional Na^+-independent electrogenic manner. Their involvement in the absorption of cationic drugs in kidney, liver and intestine is well-established (Lozano et al., 2013; Motohashi & Inui, 2013). Figure 1-7 illustrates the classification of SLC transporters. Since several inhaled drugs

such as ß$_2$- agonists and anticholinergic bronchodilators are either hydrophilic or cationic, they are substrates for OCT/OCTN (Salomon et al., 2015).

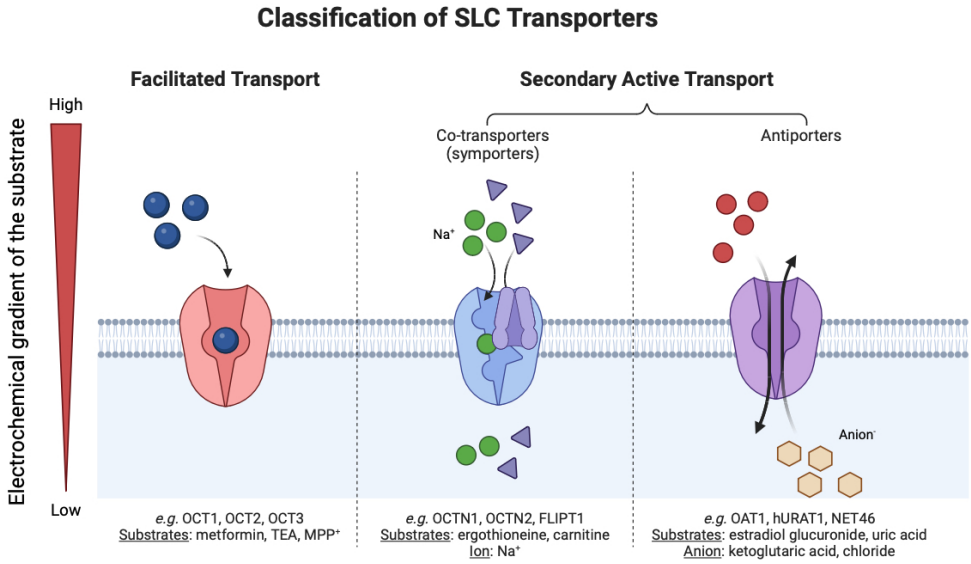


Figure 1-7. Classification of solute carrier transport (SLC) mechanisms. Arrows show the general direction of flux. Sodium and anions are ions that provide a driving force for transport by moving down their concentration gradients. Created with BioRender.com

Transporters of this superfamily present tissue-specific profiles which may imply that pharmacokinetics of pulmonary absorption differs from oral administration. Several studies have demonstrated OCTs and OCTNs to be moderately expressed in lung tissue samples and cell culture models (Berg et al., 2014; Bleasby et al., 2006; Ingoglia et al., 2015a; Lips et al., 2005; Mukherjee et al., 2012; T. Wang et al., 2007). This tissue-specific profile could be important in local pharmacokinetics and disposition of inhaled drugs. Table 1-1 summarises the tissue distribution and general information of SLC22A1-SLC22A5.

Table 1-1. OCT1-3, OCTN1 and OCTN2 general information on tissue distribution, main substrates and splice variants characterised.

Human gene symbol	Protein Name (additional)	Tissue distribution (Bleasby et al., 2006)	Substrate	Human gene locus	Splice variants	GenBank ID
SLC22A1	OCT1	Liver, small intestine, kidney, lung, skeletal muscle, brain, adipose tissue, immune cells.	Organic cations	6q26	6 variants	NM_003057.3
SlC22A2	OCT2	Kidney, small intestine, lung, placenta, thymus, brain (neurons, blood-brain barrier), inner ear	Organic cations	6q26	5 variants	NM_003058.4
SLC22A3	OCT3	Heart, skeletal muscle, brain (neurons, glial cells, plexus choroideus), small intestine, liver, lung, kidney, urinary bladder, mammary gland, skin blood vessels	Organic cations	6q26-q27	-	NM_021977.4
SLC22A4	OCTN1	Kidney, intestine, spleen, heart, skeletal muscle, brain, mammary gland, thymus, prostate, airways, testis, eye, fetal liver, sperm, immune cells.	Ergothioneine, zwitterions, organic cations	5q31.1	2 variants	NM_003059.3
SLC22A5	OCTN2	skeletal muscle, kidney, prostate, lung, pancreas, heart, small intestine, adrenal gland, thyroid gland, liver, etc.	Zwitterions (L-carnitine), organic cations	5q31	8 variants	NM_003060.4

Interaction with Endogenous/Exogenous Compounds

SLC transporters have a widespread presence throughout the body, with particularly high expression in the epithelial cells of vital organs like the intestine, liver, and kidneys. These transporters are essential in drug absorption, metabolism, distribution, and elimination, and their clinical significance cannot be overstated. It is now widely acknowledged that SLC drug transporters play a crucial role in determining drug disposition and response, which has implications not only for clinical drug therapy but also for drug development. Table 1-2 summarises the identified endogenous and exogenous substrates of OCT and OCTN transporters..

Genetic variations in SLC22A1, encoding OCT1, can influence the pharmacokinetics of metformin. Individuals carrying specific reduced-function alleles exhibit significantly different drug behavior compared to those with only reference alleles (Shu et al., 2008).

Ergothioneine (ET), a highly hydrophilic zwitterion at physiological pH, has been demonstrated as a substrate for OCTN1 transport (Gründemann et al., 2005). Interestingly, studies using OCTN1 knockout models, devoid of tissue ET, haven't revealed any major outward signs of deficiency. However, lower circulating ET levels have been associated with an increased risk of various disorders, including Parkinson's disease (PD) (Hatano et al., 2016), mild cognitive impairment (MCI) (Cheah & Halliwell, 2021), and Crohn's disease (CD) (Lai et al., 2019). Conversely, higher blood ET levels correlate with a lower risk of cardiometabolic disorders and associated mortality (E. Smith et al., 2020).

Loss-of-function (LOF) variants within the SLC22A5 gene, encoding the organic cation transporter OCTN2, are the causative factor for Carnitine Transporter Deficiency (CTD). CTD, while a rare inborn error of metabolism, presents with potentially lethal consequences if left undiagnosed (Koleske et al., 2022).

Table 1-2. Substrates of the OCT and OCTN transporters.

Transporter	Endogenous	Drugs	References
OCT	cyclo(His-Pro), salsolinol, agmatine, MPP+ (1-methyl-4-phenylpyridinium), thiamine	Metformin, lamuvidine, acyclovir, ganciclovir, pentamidine, furamidine, berberine, oxaliplatin, picoplatin, cis-diammine(pyridine)chloroplatin(II), irinotecan, paclitaxel	(Hendrickx et al., 2013) (Koepsell et al., 2007) (Koepsell & Endou, 2004)
OCT2	Acetylcholine, dopamine, epinephrine, norepinephrine, serotonin, histamine, putrescine, choline, cyclo(His-Pro), salsolinol, agmatine,	Memantine, amantadine, oxaliplatin, picoplatin, cisplatin, ifosfamide, cimetidine, famotidine, ranitidine, zalcitabine, lamuvidine, amiloride, metformin, and berberine	(Hendrickx et al., 2013) (Koepsell et al., 2007)
OCT3	Epinephrine, norepinephrine, histamine, agmatine, cyclo(His-Pro), salsolinol	Lidocaine, quinidine, metformin, etilefrine, oxaliplatin, lamuvidine	(Koepsell et al., 2007)
OCTN1	Acetylcholine, ergothioneine, glycine-betaine, L-carnitine	Quinidine, pyrilamine, verapamil, ipratropium, tiotropium, mitoxantrone, doxorubicine, stachydrine, betonicine, gabapentin	(Pochini et al., 2011) (Urban et al., 2008)
OCTN2	L-Carnitine, choline	Mildronate, cephaloridine, emetine, pyrilamine, verapamil, spironolactone, oxaliplatin, ipratropium, tiotropium, colistin	(Visentin et al., 2017) (Koepsell et al., 2007)

Ongoing research investigates the impact of organic cation transporters/novel organic cation transporters on the pulmonary distribution of inhaled drugs. Even though organic cation transporters have been implicated in the absorption of inhaled drugs, (Gumbleton et al., 2011; Nakanishi et al., 2013) there is a lack of clarity as to whether SLC22A transporters play a role in pulmonary absorption. Ehrhardt et al., (2005) contributed to the current understanding of salbutamol absorption, used to relieve symptoms of asthma and chronic obstructive pulmonary disease (COPD), by demonstrating, for the first time, its active net absorption across human airway epithelial barriers in vitro. Proposing the involvement of OCT/Ns in the active transport across monolayers. Al-Jayyoussi et al., (2015) showed that OCT/OCTNs were not involved in the transport of ipratropium and L-carnitine in pulmonary epithelia. Instead, absorption was demonstrated to be driven by passive diffusion. Therefore, this study highlights the potential of the inhaled route for drugs that undergo low absorption by oral administration (Al-Jayyoussi et al., 2015; Tronde et al., 2003). It is worth noting that these studies employed the use of an ex-vivo lung model which may lead to significantly distorted results as the perfusion process fails to reflect the distribution of the drug from airway to lung submucosal tissue. Fenoterol, a widely used anti-asthmatic has been reported as a substrate of SLC22A1(Tzvetkov et al., 2018).

The β2-adrenergic receptor agonists, formoterol and salbutamol (albuterol), and the muscarinic antagonist ipratropium, commonly employed in the treatment of asthma and chronic obstructive pulmonary disease (COPD), have been shown to inhibit the organic cation transporters SLC22A1 (OCT1), SLC22A3 (OCT3), SLC22A4 (OCT4), and SLC22A5 (OCT6) in bronchial epithelial Calu-3 cells

(Mukherjee et al., 2012). Anticholinergic drugs, i.e. ipratropium bromide and tiotropium, commonly prescribed for treating asthma and chronic obstructive pulmonary disease, are primarily taken up by the organic cation transporter OCTN2 (Table 1-2) in bronchial epithelial cells, BEAS-2B cells. OCTN1 also plays a role, albeit to a lesser extent (Nakamura et al., 2010). According to previous results, in airway epithelial cells, the primary members involved in the interaction with β2-agonists are OCT1 and OCTN2, while OCT3 primarily plays a role in the uptake and clearance of β2-agonists in and out of airway smooth muscle cells.

Structure-Function of Organic Cation Transporters

Members of the SLC22A transporter family share a similar structure that consists of 12 predicted α-helical trans-membrane domains (TMDs) with both N- and C-termini localised intracellularly. A large glycosylated extracellular loop located between TMDs 1/2, and a large intracellular loop between TMDs 6 and 7 containing phosphorylation sites (Koepsell & Endou, 2004). The extracellular loop mediates homo-oligomerization which may be important for membrane trafficking, with non-influence in transport activity (Keller et al., 2011). The large intracellular loop is involved in posttranscriptional regulation (Koepsell et al., 2007). Various attempts have been made to elucidate the structure of the binding site of SLC22 transporters. A recent study has determined the structure of the human OCT3 transporter at a resolution of 3.2 Å. This study provides an insight into the ligand binding pocket of OCT3 through the examination of its structure when bound to two inhibitors, corticosterone and decynium-22 (Khanppnavar et al., 2022).

It is supposed that OCTs follow an alternating-access transport mode. An organic cation binds to the outward-open conformation to the transporter. This induces a conformational change where the cavity closes on the extracellular space and opens on the intracellular side. The substrate is released to the cytoplasm and the transporter switches back to the outward open conformation. This translocation probably requires conformational changes of the TMDs during transition from the outward and inward facing orientation (Gorbunov et al., 2008). Pedersen et al., (2013) constructed a model of OCT1 (SLC22A1) and OAT3 (SLC22A8) from a homologous fungal transporter, which suggested that highly conserved residues on the TMD7, Tyr328 and Trp320 might be involved in the substrate-binding mechanism in the SLC22 family. Additional data suggests that a hinge domain in TMD11 is crucial for the structural changes in substrate movement. OCT1 has also been modelled with an outward-facing conformation (Volk et al., 2009). Comparison of both conformations revealed that the general structure of the innermost cavities is conserved, but specific amino acids are involved in substrate affinity.

Predicted structures are available at the AlphaFold database (https://alphafold.ebi.ac.uk/) (accessed on 12 February 2023) and shown in Figure 1-8.Although, predicted 3D structures have been modelled with an inward-facing conformation and caution should be exercised when conducting *in-silico* binding analysis. This is due to the inherent complexities of predicting protein structures, which may affect the accuracy of the analysis. It is recommended that the results of the *in-silico* binding analysis be critically evaluated and verified through additional experimental methods.

Organic Cation Transporter 1 (OCT1/SLC22A1)

The SLC22A member 1 gene (Gene ID: 6580) is mapped in chromosome 6q25 and is comprised of 11 exons and 10 introns. OCT1 was the first organic cation transporter to be cloned from a rat renal cDNA library (Gründemann et al., 1994). Table 1 1 summarises general information for OCT1. The human gene has 7 transcript variants with 4 variants coding for a functional transporter (SLC22A1 Solute Carrier Family 22 Member 1 [Homo sapiens (Human)] - Gene - NCBI, n.d.). OCT1 transporter contains 554 amino acids with a predicted membrane structure that correspond to the SLC22A family.

In order to determine the domains involved in substrate recognition and transport of the rat oat3/oct, Feng et al., (2001) constructed a chimera consisting of TMDs 1–5 from the oat3 and TMDs 6–12 from the oct1. Rat oat3/oct1 chimera displayed a selectivity for cations, rather than anions, suggesting that the site of substrate selectivity of OCTs is within the C-terminal half of the protein transporters. Moreover, Gorbunov et al., (2008a) performed a functional characterization of OCT1 by replacing several amino acids within the transmembrane α-helix (TMH) 11. Their results suggested that aspartic acid 475, located in the TMH 11 in OCT might play a pivotal role in the transport function of rOCT1.

As mentioned earlier, OCT1 transporter is primarily expressed in the liver, brain, skeletal muscle, and immune system cells.(Bleasby et al., 2006; Nishimura & Naito, 2005) However, it has also been found in the apical membrane of ciliated epithelial cells in bronchi (Olsson et al., 2011; *Tissue Expression of SLC22A1 - Summary - The Human Protein Atlas*, n.d.). OCT1 has been shown to be able to

transport a large number of drugs (Lozano et al., 2013; Salomon et al., 2015) including neurotransmitters (Lips et al., 2005), anthracyclines (Andreev et al., 2016) and ranitidine (Meyer et al., 2017). Indeed, recent studies have shown that OCT plays an important role in the absorption of inhaled drugs. In 2015, Salomon et al,. (Salomon et al., 2015) analysed the influence of OCT/OCTN on the absorption of the β2-agonists salbutamol, formoterol and salmeterol in the presence of [^{14}C]-TEA in alveolar, bronchial, and bronchiolar epithelial cells. Moreover, they transfected cells to express OCT1-3 individually. Absorption in hOCT1-transfected cells indicated salbutamol to be a substrate for OCT1.

It might be argued, however, that such a conclusion would require confirmation by assessing the effect of silencing transporter expression *per se*. As in a previous study indicated OCT3 to be the main transporter for 1-Methyl-4-phenylpyridinium (MPP(+)), however significant OCT1 activity was observed at the apical side of bronchial epithelial (Ingoglia et al., 2015a). As a result, gene silencing confirmed OCT3 to be a high-affinity transporter for MPP+, while OCT1 was shown to transport the substrate with a lower affinity.

Novel Organic Cation Transporter 1 (OCTN1/SLC22A4)

The solute carrier family 22 member 4 (Gene ID: 6583) codes for an integral membrane protein, OCTN1, which has been found to transport inhaled drugs such as anticholinergics (Nakamura et al., 2010). This protein is encoded in 11 exons, located in chromosome 5q31.1 Table 1-1 summarises the general information for OCTN1. SLC22A4 has 3 splice variants with only 1 protein-coding variant. OCTN1 was discovered in 1997 by Tamai et al (Tamai et al., 1997), who characterised it as a pH-dependant transporter. It was later

demonstrated to be also a multi-specific and bidirectional transporter (Yabuuchi et al., 1999). SLC22A4 codes for a 551-amino acid protein with 12 putative transmembrane domains and one intracellular nucleotide binding site between transmembrane 4 and 5 involved in the regulation of transport activity (Pochini et al., 2011).

OCTN1 has been found in several adult tissues, mainly expressed in the kidney, and moderately expressed in the liver (Bleasby et al., 2006; Koepsell, 2013; Nishimura & Naito, 2005; Tamai, 2013). A comprehensive study reported OCTN1 to be the most abundant protein (2.08 ± 1.19 fmol/μg) among all drug transporters in human lung tissues (Sakamoto et al., 2013) and is expressed primarily on the apical side of the bronchiolar, bronchial and alveolar epithelium, as well as in alveolar macrophages. Further, Horvath et al (Horvath et al., 2007) reported high mRNA levels of OCTN1 in human airway epithelia. In 2009, Nakamura et al(Nakamura et al., 2010) showed that OCTN1 contributed to the uptake of the cationic anticholinergic drugs ipratropium and tiotropium in human airway epithelial cells by silencing the expression of OCTN1 and OCTN2.

Novel Organic Cation Transporter 2 (OCTN2/SLC22A5)

The solute carrier family 22 member 5 gene codes for OCTN2 (Gene ID: 6584). It is located on chromosome 5, downstream of the OCTN1, and consists of 10 exons and 9 introns. (*SLC22A5 Solute Carrier Family 22 Member 5 [Homo Sapiens (Human)] - Gene - NCBI*, n.d.). In 1998, Wu et al(X. Wu et al., 1998) were the first to isolate and clone this transporter from a human placental trophoblast cell line cDNA library. Table 1-1summarises the general

information for OCTN2. The sequence has 7 splice variants, including 4 protein-coding variants. The OCTN2 gene codes for a 557 amino-acid protein, displaying the common topological structure of the SLC22A family, with a predicted molecular mass of 63 kDa(X. Wu et al., 1998). The extracellular loop consists of 107 amino acids with 3 potential sites for N-glycosylation. The nucleotide binding site is located in the intracellular loop between fourth and fifth transmembrane domains, which is involved in the regulation of transport activity(Pochini et al., 2011). Based on their respective amino-acid sequences, OCTN2 is more closely related to OCTN1 than to OCTs.

OCTN2 is expressed at the highest levels in the kidney and in skeletal muscle (Bleasby et al., 2006; Koepsell, 2013; Nishimura & Naito, 2005; Tamai, 2013). Expression of both OCTN transporters have been demonstrated in the apical portion of human lung airway epithelial cells (Horvath et al., 2007). In a recent analysis, the expression profile of OCTN2 among different *in vitro* lung epithelial cell lines was consistently low (25-50% quartile (Endter et al., 2009). Nakamura et al(Nakamura et al., 2010) investigated the contribution of OCTNs to the uptake of ipratropium by knocking-down the expression of both transporters in human bronchial epithelial cells. The results suggested involvement of OCTN1 and OCTN2 in the absorption of ipratropium in human airway epithelia. Furthermore, a significant reduction of uptake was observed when OCTN2 was silenced, suggesting it might have a dominant role in the pulmonary absorption of ipratropium.

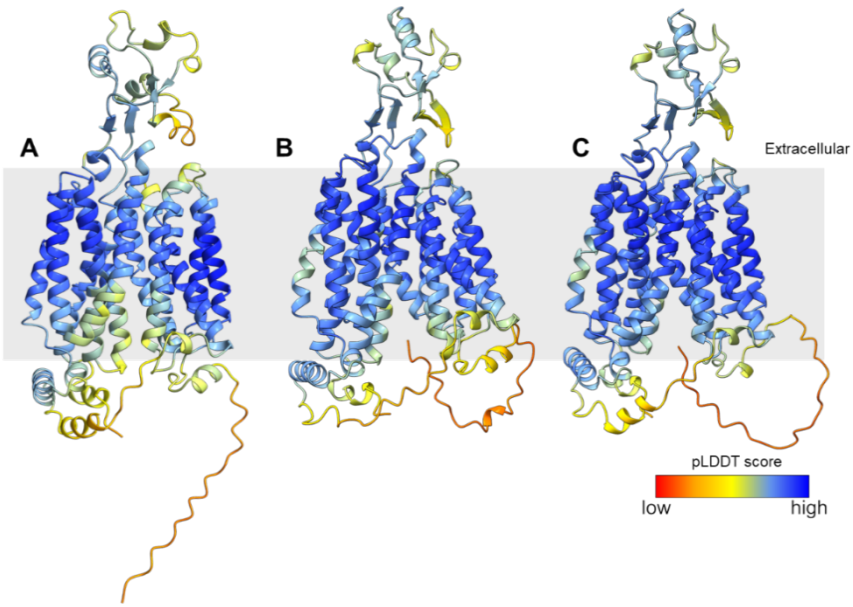


Figure 1-8. Ribbon representation of the A) SLC22A1 B) SLC22A4, C) SLC22A5 structural model retrieved from Alphafold Database (Varadi et al., 2022b). Includes 12 α-helical transmembrane domains (TMDs), a large extracellular loop between TMDs 1 and 2, and a large cytoplasmic loop between TMDs 6 and 7. AlphaFold produces a per- residue confidence score (pLDDT). Structures are coloured according to their pLDDT score.

Data from a site-directed mutagenesis performed in OCT1 indicated that Asp475, located in the middle of in TM11 is important for cation selectivity (Gorbunov et al., 2008). Replacement of this aspartate by a neutral or basic amino acid resulted in a significant reduction in the transport rate. However, affinity to some transported cations (TEA, NMN and choline) was significantly increased when Asp475 was replaced by glutamate, whereas affinity to other cations (MPP) remained unchanged. This data suggests that Asp475 is close to the cation binding site and stabilises the conformation of the cation-binding pocket.

Expression and Localisation of OCT

In a comprehensive study, Bleasby et al., (2006) examined tissue expression profiles of different transporter genes in humans. OCT/OCTN transporters were detected in a wide variety of tissues, as seen in Figure 1-9. OCT1 has been shown to be primarily expressed in the liver, brain, skeletal muscle, and immune system cells. OCT2 has been found to be strongly expressed in kidney and skeletal muscle (Bleasby et al., 2006; Nishimura & Naito, 2005). Although, OCT2 is not generally found in lung tissue or immortalised cells. Further, studies have shown a lack of expression of OCT3 in both lung tissue and primary cells (Sakamoto et al., 2013). Others have reported low expression in cell culture models, though profiles seem to be dependent on culture interface and passage (Endter et al., 2009; Ingoglia et al., 2015a; Mukherjee et al., 2012; Salomon et al., 2012). OCT3 is expressed predominantly in the prostate, skeletal muscle, salivary glands (Bleasby et al., 2006) and some areas of the brain (C. Hu et al., 2019; Lozano et al., 2018). OCTN1 is primarily expressed in the kidney, and to a modest degree in the liver. Meanwhile, OCTN2 is predominantly expressed in the kidney and skeletal muscle (Bleasby et al., 2006; Koepsell, 2013; Nishimura & Naito, 2005; Tamai, 2013).

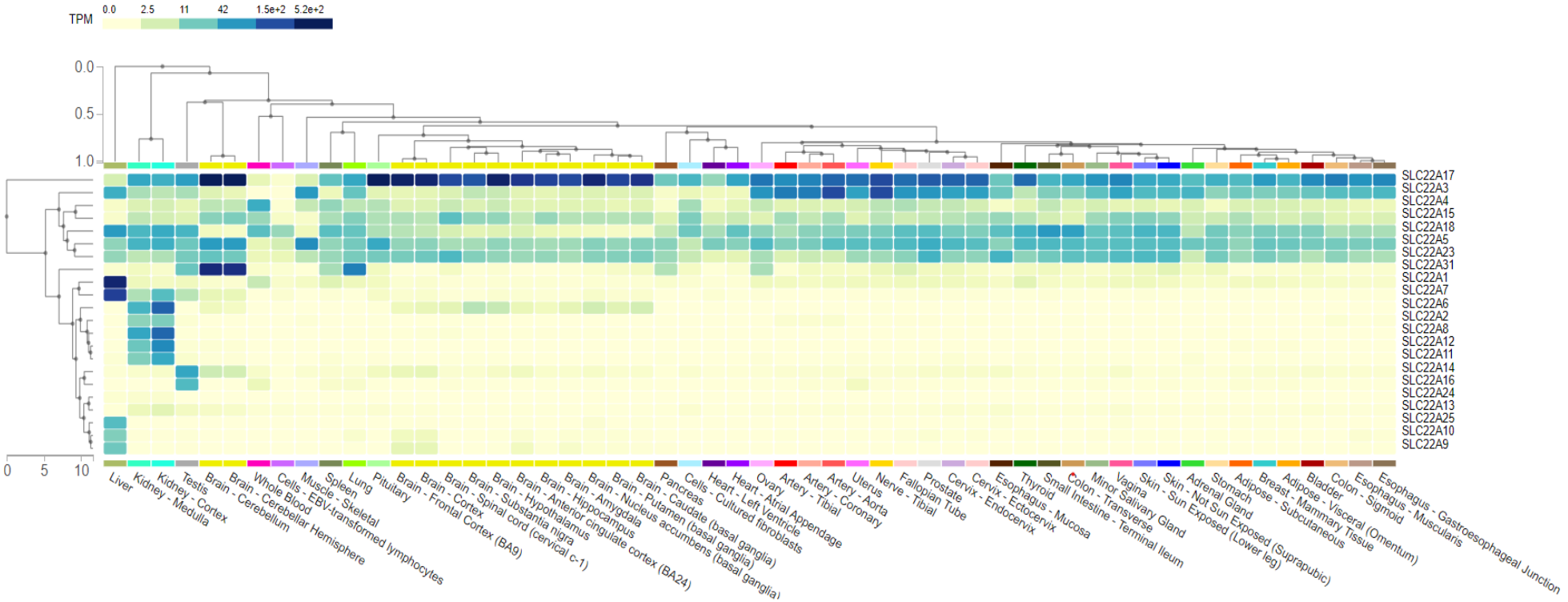


Figure 1-9. Transcriptomic expression of 22 members of the solute carrier 22 family across 54 tissues. Transporters are clustered together according to their abundant expression. SLC22A17 is on top as is ubiquitously expressed across almost all tisues (dark blue). The data used in this figure were obtained from the Genotype-Tissue Expression (GTEx) Multi Gene Query page (https://www.gtexportal.org/home/multiGeneQueryPage) (GTEx Portal, n.d.).

Studies has been mainly focused on liver, kidney and blood-brain barrier (Koepsell et al., 2007). There remain several aspects of the role of these drug transporters in the pulmonary epithelium about which relatively little is known. Expression patterns of the transporters OCT1, OCT3, OCTN1, and OCTN2 in the lung epithelium have been widely studied and there is a general agreement that these transporters are widely distributed throughout the lung. However, the expression of OCT2 is less well-established, and many lung-derived cell lines have been found to lack the transporter, with the exception of NCl-H441 (Courcot et al., 2012; Endter et al., 2009; Ingoglia et al., 2015a; Mukherjee et al., 2012; Sakamoto et al., 2015b). Table 1-3 highlights the expression profiles of SLC22A family in different lung cell models.

Using RT-PCR, Lips et al., (2005) demonstrated OCT1-3 expression in the epithelial cells of the trachea and bronchi in humans. While immunohistochemistry (IHC) of human OCT1 transporter was present in the intracellular space and the luminal membrane of ciliated epithelial cells in human bronchi. The study also found that OCT2 was mainly present in the apical membrane of ciliated cells and less intensely at the plasma membrane of basal cells. Expression of OCT3 was weakly observed in the apical membrane of ciliated cells, while the entire plasma membrane of basal cells and the basolateral membrane of intermediate cells showed more intense immunoreactivity. Bleasby et al., (2006) showed that expression profiles of OCT1 and OCT3 in the human epithelia fell within 25-50% quartile, while OCT2 showed a lower degree of expression (0-25% quartile). In contrast, a higher degree of expression of OCTN1 and OCTN2 has been demonstrated in the human lung (50-75% quartile). In line with these results, Horvath et al., (2007) found OCT2

transporter to be nearly undetectable, while reporting a relatively high degree of expression of OCT1 and OCTN2 in human airway epithelial cells.

A previous study (Berg et al., 2014) found that OCTN1 showed the strongest expression among all the OCT/N transporters in human bronchi, however, it was expressed at a lower level in peripheral lung tissue. Regional expression and subcellular localization of OCT/N transporters in the airways is still somewhat unclear. Due to their biomedical relevance and expression profiles, the present study therefore focusses on the absorption mechanism of OCT1, OCTN1 and OCTN2 transporters of the polyspecific SLC22A family. Table 1-3 summarises the mRNA and protein expression profiles of OCT1-3, OCTN1 and OCTN2 in lung tissue and cell culture models.

Expression levels in NHBE cells were found to be low for OCT1 (0.48 ± 0.294 fmol/µg), high for OCTN1 (2.01 ± 0.35 fmol/µg) and undetectable for OCTN2. Expression of OCTs/OCTNs in A549 has been previously analysed by Western Blot; it was found that all five organic cation/carnitine transporters were present in A549 cells (Salomon et al., 2012). However, a recent study conducted by Ingoglia et al., (2015a) found mRNA expression of OCT1 to be barely detectable, and OCT2 to be undetected in alveolar cells. Endter et al., (2009) found 16HBE14o- cells to display low-moderate levels of OCT1, OCTN1 and OCTN2 transporters. These expression profiles were later confirmed by Western Blot analysis, which showed that 16HBE14o- cells expressed OCT1 and both carnitine transporters, OCTN1 and OCTN2 (Salomon et al., 2012). Data from recent studies suggest that the four isoforms of the human SLC22A organic cation/carnitine transporter family, OCT1, OCT3, OCTN1 and OCTN2 are

present in the bronchial epithelial cell line Calu-3 (Endter et al., 2009; Ingoglia et al., 2015a; Mukherjee et al., 2012; Salomon et al., 2012). However, there is contradictory data regarding the expression of OCT3 in Calu-3 and primary cells, and its expression profile has been shown to fluctuate over passages.

Expression in non-healthy tissue

Polymorphisms of organic cation transporters (OCTs) can alter physiological functions due to their role in the uptake, excretion, and tissue distribution of compounds such as neurotransmitters, essential nutrients, and toxic substances. This can impact the emergence and progression of diseases, as well as the elimination of toxins, highlighting the importance of studying the functional effects of OCT polymorphisms.

Table 1-3. Comparison of mRNA and Protein Expression Profiles of SLC22A transporters in Lung Tissue and Pulmonary Cell Lines.

	Lung Cell Models						
	Calu-3		A549		NHBE		
SLC transporter	mRNA Expression by Hutter et al., (2014)	Protein Expression by Sakamoto et al., (2015b)	mRNA Expression by Ingoglia et al., (2015a)	Protein Expression by Sakamoto et al., (2015b)	mRNA Expression by Hutter et al., (2014)	Protein Expression by Sakamoto et al., (2013)	Protein Expression in Lung Tissue by Sakamoto et al., (2013)
OCT1	++	+	+	+	+	+	+
OCT2	-	-	-	-	-	+	+
OCT3	+	+	+++	+	++	-	-
OCTN1	-	+	N.A.	+	-	+++	++
OCTN2	++	+	N.A.	-	++	-	-***

The quantitative data of protein expression by mass spectrometry (fmol/µg protein) in the present study: −, not detectable; +, low expression (1 fmol/µg protein > 0); ++, moderate expression (2 fmol/µg proteint>1); +++, high expression (fmol/µg proteint>2). Data from Hutter et al., (2014) represent gene expression by qPCR Signal intensity represented as: − (<0.001) 'negligible expression'; + (0.001–0.02) 'low expression', ++ (0.02–0.5) 'moderate expression, +++ (>0.5) 'high expression as compared to the two house-keeping genes RPLP0 and MVP. Calu-3, passage 25–30. A549, passage 36-41 below the LQ (0.0336 fmol/µg protein). N.A., not available.

Lower expression of SLC22A1 was linked to poor patient survival in liver cancer. This downregulation was found to be significantly connected to more advanced stages of hepatocellular carcinoma (HCC), characterized by a greater number of T3 tumours with larger diameters, lower differentiation, and higher alpha-fetoprotein levels (Heise et al., 2012). In addition, western blot analysis revealed a distinct protein expression pattern in tumour samples with a more widespread staining seen in immunofluorescence, suggesting that OCT1 may not be functional in advanced HCC.

OCTN2 is transporter with a high affinity for carnitine. Deficiencies in the OCTN2 carnitine transporter cause autosomal recessive primary carnitine deficiency, leading to decreased carnitine accumulation within cells, increased carnitine loss in urine, and low serum carnitine levels (Longo et al., 2016). It has been reported that variations in the genes coding for the organic cation transporters, SLC22A4 (OCTN1) and SLC22A5 (OCTN2), located at the inflammatory bowel disease 5 locus, can raise the likelihood of developing Crohn's disease (CD) (Leung et al., 2006). According to another study (Otter et al., 2022), the organic cation transporters SLC22A1, SLC22A3, SLC22A4, SLC22A5, and SLC22A16 were expressed in both normal breast tissue and breast cancer tissue. However, the study revealed that there was a significant increase in the expression of SLC22A1 in breast cancer tissue as compared to normal breast tissue.

In a recent study, significant association was observed between overall survival in patients with clear cell renal cell carcinoma (KIRC) and kidney renal papillary cell carcinoma (KIRP) with SLC22A4 and SLC22A5. Furthermore, SLC22A5

was found to be linked with tumour size and progression in KIRC (Whisenant & Nigam, 2022). Previous research has also established the association of SLC22A1 with overall survival in KIRC (D. G. Hu et al., 2020).

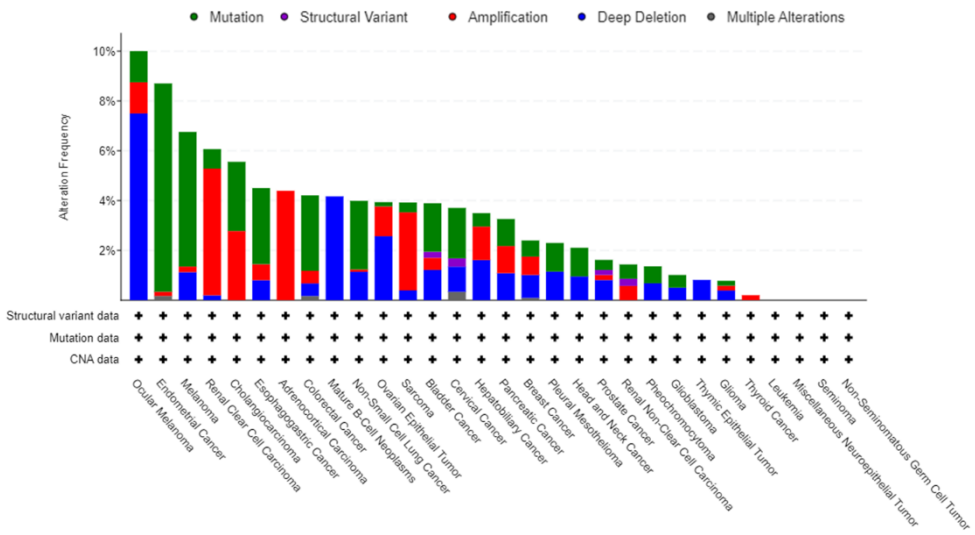


Figure 1-10. Alteration frequency of SLC22A1, SLC22A4 and SLC22A5 in different cancer types. Data was obtained from cBioPortal (https://www.cbioportal.org) (accessed on 13 February 2023) and summarises samples from 32 studies according to cancer type and genomic alteration types.

A study suggested that there was a correlation between the SLC22A5 polymorphism and the emergence of extra-articular manifestations in individuals with rheumatoid arthritis (RA). (Pawlik et al., 2019). Cells derived from patients with Squamous Cell Carcinoma (SCC) were found to have a high expression level of OCTN2, according to a study conducted by Sun et al., (2021).

According to data obtained from the cBioPortal (Figure 1-10), in Renal Clear Cell Carcinoma, alterations in SLC22A1, SLC22A4 and SLC22A5 genes were observed in 6.07% of 511 cases. The alterations included mutation in 0.78%, amplification in 5.09%, and deep deletion in 0.2%. While in adrenocortical

carcinoma, SLC22A genes were altered in 4.4% of 91 cases with all alterations being amplifications. Surprisingly, ocular Melanoma, 10% (8 out of 80) cases showed gene alterations. Alterations include mutation in 1.25%, amplification in 1.25%, and deep deletion in 7.5%. Lastly, in non-small cell lung cancer SLC22A1, SLC22A4, and SLC22A5 genes showed alterations in 3.99% of 1053 cases, with mutations observed in 2.75%, amplifications in 0.09%, and deep deletions in 1.14% of cases.

A small number of studies have been conducted to examine the connection between expression of OCT/Ns and lung diseases. A genome-wide association studies (GWAS) have identified SLC22A5 variants being linked to asthma (Moffatt et al., 2010; Shrine et al., 2019). Expression of OCTN1/2 was found to be decreased in the lungs of rats in vivo and in a human alveolar cell line in vitro after the induction of chronic obstructive pulmonary disease (COPD) through exposure to cigarette smoke extract (CSE) and lipopolysaccharide (LPS) (Qi et al., 2020). Although Berg et al., (2014) reported no differences in mRNA expression levels of OCT1, OCT3, OCTN1, and OCTN2 between healthy subjects and ex-smokers with severe COPD, they were able to detect a distinct difference in mRNA expression of genes encoding membrane transporters between central airways and peripheral tissue. SLC22A4 (OCTN1) exhibited a higher expression in the central airways, indicating its potential importance in transporting bronchodilators like muscarinic antagonists and β2-agonists.

To simulate asthmatic-like conditions in the epithelium in vitro, 19 days old Calu-3 cells were subjected to air-interfaced culture (AIC) conditions and exposed to pro-inflammatory lipopolysaccharide (LPS) or house dust mite

extract (HDM) (Mukherjee et al., 2017). The LPs challenge significantly increased expression of SLC22A1, SLC22A3, SLC22A4 and SLC22A5 on mRNA and protein levels. Exposure to HDM showed similar results in the first 8 h of exposure where levels returned to basal within 24 h.

Regulation of Organic Cation Transporters

The regulation of transport activity in response to endogenous and exogenous signals may occur at various levels such as transcription, mRNA stability, translation, and posttranslational modification. In general, transcriptional regulation and posttranslational modification are believed to be responsible for long-term and short-term regulation, respectively (Ciarimboli & Schlatter, 2005), Figure 1-11. We are interested in the transcriptional regulation of drug transporters, as changes in transport activity are dynamically regulated by increases or decreases in levels of mRNA expression. Table 1-4 summarises transcription factors and regulatory mechanisms that have been associated with OCT/OCTNs.

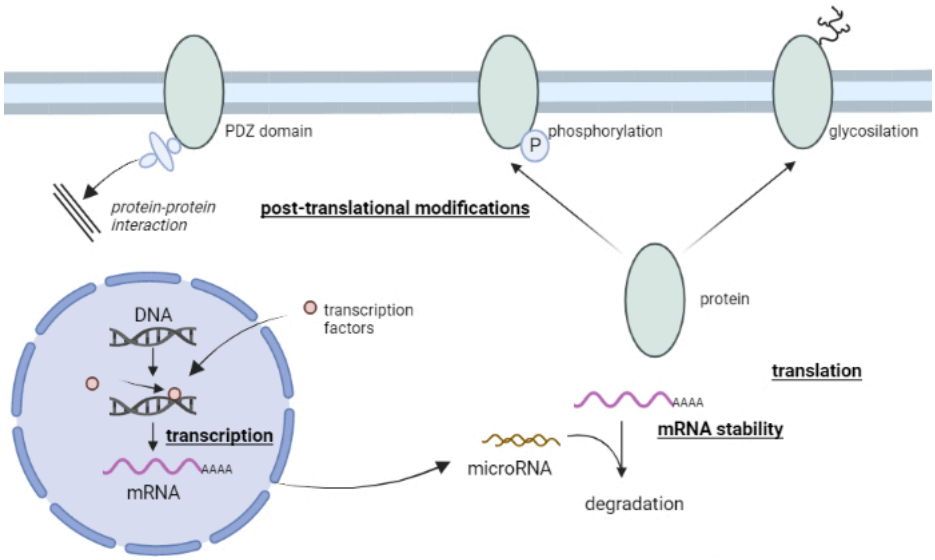


Figure 1-11. Various factors influencing the expression of drug transporters. The activity of drug transporters may be regulated at various levels including transcription, mRNA stability, translation, and posttranslational modification. Posttranslational modification may involve glycosylation, phosphorylation, and protein–protein interaction. Moreover, transcriptional regulation is of particular interest, because many extra- and intracellular signals eventually alter the activity of transcription factors. In addition to the regulation of various signals, the tissue-specific expression of drug transporters is also under transcriptional control.

The regulation of hOCT1 expression occurs at various levels, including transcription, intracellular trafficking, and the modification of functional properties. Expression of OCT1 in normal human hepatocytes is primarily regulated by the hepatocyte nuclear factor-α (HNF4) nuclear receptor through two direct repeat (DR)-2 format sites (Saborowski et al., 2006).

Involvement of the upstream binding stimulating factors (USF) 1 and USF2 in the SLC22A1 promoter region was characterized in liver cell lines Huh7 and HepG2 (Kajiwara et al., 2008). In a previous study, liver cells were used in reporter gene assays that revealed a specific binding of HNF1 to the evolutionary conserved region (ECR) in intron 1 of SLC22A1. In addition, a strong

relationship was identified between high levels of HNF1 expression and high levels of OCT1 expression in human liver samples (O'Brien et al., 2013).

Table 1-4. Protein Factors and Mechanisms Controlling SLC Drug Transporter Gene Transcription.

Transporter	Regulatory Mechanism/ Protein Factor	References
OCT1	DNA methylation, PXR, RXR, RAR, PPAR*α,* HNF4α	(Schaeffeler et al., 2011) (Martovetsky et al., 2013) (Rulcova et al., 2013)
OCT2	Androgen receptor (AR), P*,* Histone modification, DNA methylation, TFII-I	(Aoki et al., 2008; Freitas-Lima et al., 2020; Martovetsky et al., 2013; Q. Zhu et al., 2019)
OCT3	DNA methylation, MZF1	(L. Chen et al., 2013; Gu et al., 2022; Nies et al., 2009)
OCTN1	DNA methylation, RUNX1, NF-κB	(Maeda et al., 2007) (Yamada et al., 2004)
OCTN2	Estrogen receptor (ER), PPAR*γ,* RXR, PPAR*α*,	(C. Wang et al., 2012) (Qu et al., 2013)

In another study by Rulcova et al., (2013), it was observed that dexamethasone significantly increased the expression of SLC22A1 mRNA and protein in normal primary human hepatocytes, but not in the hepatocyte-derived tumour cell lines HepG2 and MZ-Hep1. The induction of HNF4α by dexamethasone

was similarly observed in primary human hepatocytes, but not in the hepatocyte tumour-derived cell lines. Thereby, suggesting a relationship between dexamethasone, HNF4α, and the expression of SLC22A1. The role of transcription factors in the regulation of OCT1 gene was further explored. Results showed a significant contribution of the HNF4α, CCAAT/enhancer binding proteins β (C/EBPβ), and peroxisome proliferator-activated receptor-γ coactivator 1α (PGC1α) factors in the regulation of the OCT1 gene. Additionally, the expression of OCT1 mRNA in human livers was found to positively correlate with the expression of C/EBPβ and HNF4α mRNAs. C/EBPβ was also found to stimulate the OCT1 gene reporter construct in HepG2 cells. However, it was noted that neither C/EBPβ nor PGC1α were upregulated in human hepatocytes by dexamethasone.

DNA methylation of SLC22A1 has been found to be correlated with a decrease in the expression levels of SLC22A1 in cases of hepatocellular carcinoma (HCC) (Schaeffeler et al., 2011). Whereas, upregulation of SLC22A1 in chronic myeloid leukemia (CML) cells was observed after incubation with the tyrosine kinase inhibitors, imatinib (Gromicho et al., 2011; Sreenivasan Tantuan & Viljoen, 2018).

An evaluation of the post-transcriptional regulation of the SLC22A1 gene revealed the existence of a negative modulation by the ischemia/reperfusion-inducible protein (IRIP) in conjunction with its interaction with the regulatory protein RS1, as described in the study by LI et al., (2013), This interaction has been shown to regulate the activity of membrane transporters through both endocytotic and exocytotic pathways (Korn et al., 2001; Kroiss et al., 2006).

It has been reported that the binding of the Runt-related transcription factor 1 (RUNX1) to intron 1 of the SLC22A4 gene is involved in its transcriptional regulation, as documented in the study by Tokuhiro et al., (2003). Additional studies demonstrated regulation by inflammatory cytokines such as interleukin-1β, and nuclear factor-κB (NF-κB), in fibroblast-like synoviocyte cell line MH7A, derived from rheumatoid arthritis (RA) patients (Maeda et al., 2007).

Transcriptional regulation of SLC22A5 has been linked to the presence of a heat-shock transcription factor (HSF)-binding element (HSE), located 207 bp upstream of the start codon (Peltekova et al., 2004). Promoter methylation has been established as a factor responsible for the down-regulation of transcription. The effects of the demethylating agent decitabine (DCA) on the expression of OCTN2 were examined in four different cell lines: HepG2 (hepatoma), LS174T (colon cancer), QBC-939 (bile duct cancer), and U251 (glioma). The results showed that treatment with DCA led to an increase in OCTN2 mRNA and protein levels in HepG2 and LS174T cells, but no significant changes were observed in QBC-939 and U251 cells. An analysis of the CpG islands revealed a methylated CpG site in promoter Region-1 (which spans from -325 to -92 base pairs), which was significantly hypermethylated in HepG2 and LS174T cells compared to QBC-939 and U251 cells (Qu et al., 2013).

In a previous study (C. Wang et al., 2012), it was demonstrated that the transcription of hOCTN2 in breast cancer cells is modulated by estrogen. This regulation involves an intronic estrogen-responsive element (ERE) and an enhancer region that contains a binding site for the nuclear receptor related 1 (NR4A2/Nurr1). In colon cells, evidence has been provided that the peroxisome

proliferator-activated Receptors (PPAR)γ stimulates transcription of OCTN2 via functional PPRE located in the first intron (D'Argenio et al., 2010).

1.3. Models to assess pulmonary absorption and disposition of inhaled therapeutics

Over recent years, the lungs have emerged as an interesting route of administration for both topical and systemically acting drugs. A comprehensive characterisation of the inhaled experimental drug is required to determine the efficacy, dosing, deposition, absorption and safety of the molecule. Several models have been used for the preclinical evaluation of inhaled drug products and have recently been improved to better mimic the human respiratory epithelia. Choosing the right model to study drug transport *in vitro* is quite challenging. Whereas animal models provide a complete assessment of pulmonary drug delivery, cell cultures remain the ideal model for the study of drug transport and absorption (Fernandes & Vanbever, 2009).

1.3.1. *In vivo* models -whole animal

The use of animal models is a longstanding practice for medical and biological research. *In vivo* experiments provide information on drug disposition, metabolism, absorption and pharmacokinetic profile as well as formulation tolerability (Fernandes & Vanbever, 2009). Animal species are chosen based on the pharmaceutical study being performed. Small rodents are cost-efficient and commonly used for broader, initial assessments of pulmonary delivery, whereas larger animals are used for inhalation pharmacokinetics (Cryan et al., 2007; Sakagami, 2006).

Although studies on lung clearance can be carried out *in vivo*, detailed mechanisms of drug transport across the lung epithelium cannot be delineated. Further, not all results obtained from pre-clinical animal studies can be directly extrapolated to humans since transporter expression, distribution profiles and substrate specificity vary across species (Bleasby et al., 2006). In addition, anatomical and physiological differences among species require consideration.

1.3.2. *Ex vivo* models - lung tissue models

Animal models have recently been partially substituted by isolated perfused lungs (IPL) for their closeness to *in vivo* processes. This minimises the influence of metabolic activity of other organs that might overwhelm lung activity, thereby maintaining the anatomy and functionality of the whole lung. IPL is therefore a powerful tool to investigate the effectiveness of particle size, drug uptake, metabolism, deposition and distribution profile (Niemeier, 1984).

The technique involves the isolation of the organ from the animal, after which it is housed in an artificial system. The main advantage of this preparation is the ability to control and measure lung ventilation and perfusion parameters during an experiment (Mehendale et al., 1981). Further, the integrity of the epithelia and transcellular transport are preserved. Limitations of the model include the high levels of expertise required in the setting-up of the experiment, short-tissue viability, high costs and complexity of design(Ibrahim & Garcia-Contreras, 2013; Mehendale et al., 1981; Sakagami, 2006).

1.3.3. *In vitro* models

While *in vivo* and *ex vivo* approaches continue to be the main models used to assess dosing administration and lung-regional distribution, cell line models offer significant details on the mechanism of drug transport. Much effort has been spent on the development of cell culture systems towards the modelling of trachea-bronchial epithelia for the evaluation of the deposition and absorption profiles of inhaled formulations (Hermanns et al., 2004). Table 1-5 summarises the main characteristics of both primary and immortalised cell lines previously used to study pulmonary absorption.

The culture of lung cells is a simpler, relatively more straightforward and less expensive alternative in drug transport and permeability research due the high reproducibility of results. It is therefore the best approach for the study of drug absorption and metabolism. Cell cultures also provide an opportunity to reduce animal testing, hence shortening the development time for new drug products (Steimer et al., 2005). Important features of the pulmonary epithelium including drug transport systems, efflux pumps, metabolic pathways and morphological characteristics have to be present in cell models (Bur & Lehr, 2008). Therefore, the use of cell culture in drug absorption studies precludes hypothetical extrapolation by avoiding interspecies differences in expression and substrate recognition.

Table 1-5. Comparison of physiological features of human bronchial epithelial cells in primary cell culture and immortal cell lines.

	NHBE (Haghi et al., 2014; Rayner et al., 2019)	**Calu-3 (Forbes et al., 2003)**	**16HBE14o- (Forbes et al., 2003)**	**A549 (J. Wu et al., 2017)**
Tissue	Healthy	Adenocarcinoma	Healthy	Adenocarcinoma
Phenotype	Multilayer Ciliated*, goblet cells	Monolayer Non-ciliated, pseudo-stratified	Multilayer Non-ciliated, cuboidal	Alveolar Type II
Mucus production	Yes	Yes	No	Yes
Air interface culture	Yes	Yes	No	Yes
Tight epithelial barrier	Yes	Yes	Yes	No

**Until passage sixth.*

However, specific adaptations are required to emulate the conditions of pulmonary tissue. Lung epithelial cell models can be cultured either as an LCC (liquid-covered culture) or ALI (air-liquid interface). In LCC, cells are cultured submerged in media; however, in order to mimic pulmonary epithelia, cells may be cultured in ALI conditions. In an ALI culture, cells are grown on a porous membrane in which the basal surface of the cells is in contact with liquid media, while the apical side is exposed to air. As shown in Figure 1-12, ALI mimics the oxygen exchange environment in the *in vivo* airway. Table 1-5 depicts the main features of the most commonly used human epithelial cell lines.

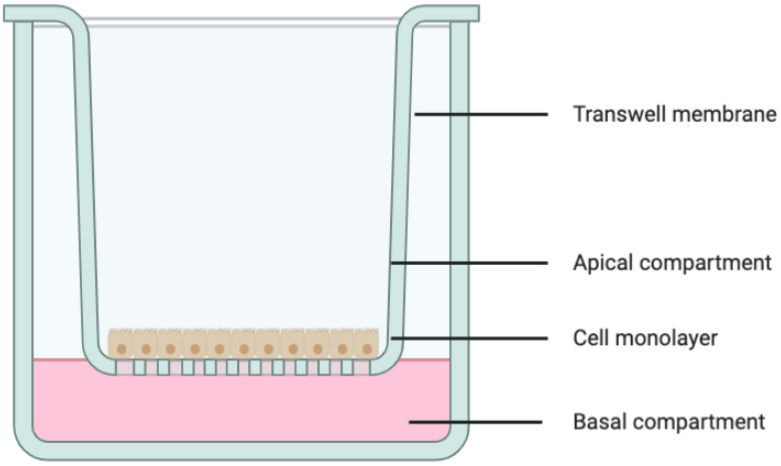


Figure 1-12. Schematic representation of a cell monolayer cultured under ALI conditions. Media is located on the basolateral chamber and the apical surface is exposed to air.

The function of the lung epithelia as a barrier depends on so-called tight junctions (TJ). Tight junctions are dynamic protein structures that hold cells together, forming a regulated barrier. This tightness prevents the lateral diffusion of molecules and ions through the space between the apical and lateral/basal surfaces. Therefore, TJ formation is a key factor when performing drug transport studies in ALI conditions (Wittekindt, 2017). Transepithelial electrical resistance (TEER) is a rapid, non-invasive and widely used to assess the integrity of tight junctions of a cellular monolayer grown on a porous support. Hence, TEER values indicate physical structural integrity and permeability of epithelial cultures (Hickman, 2016; Narai et al., 1997). Measurement is relatively straightforward. Electrodes are placed on both sides of a cell layer and electrical signals are applied to measure voltage and current in order to calculate the electrical resistance of the barrier.

Primary Cells

Primary cells are directly isolated from human lung tissue and grown into media culture. They maintain resemblance with native bronchial epithelium by their ability to differentiate into ciliated cells, goblet cells and Club cells. Culture of primary cells in ALI results in a polarised cell layer with retained characteristics that resemble the airway epithelium (Hiemstra et al., 2018).

Normal Human Bronchial Epithelial Cells (NHBE) are isolated from the epithelial lining of airways above the bifurcation of the lungs. A comprehensive study analysed expression levels of transporter proteins in cultured epithelial cells using LC-MS/MS. NHBE represent a good model to study cell-cell and cell-matrix interactions (Sakamoto et al., 2013). However, they can only be passaged up to 6-8 times (Rayner et al., 2019), and a gradual disintegration of the epithelium has been demonstrated after passage six. In addition, they form a multilayer phenotype when cultured in ALI conditions (Haghi et al., 2014; Min et al., 2016). The finite expansion highlights the limitations of using NHBE cells as models to study drug transport in pulmonary epithelia. Further, primary cells are costly, display high variability, have a short lifespan and are not easily available. They may also be more difficult to transfect, and obtaining well-differentiated cultures can be challenging (Stewart et al., 2012).

Lung cell lines

Continuous or immortalised cells are widely used to study human respiratory processes and diseases (Y. Zhu et al., 2010). They have an extended life span, are increasingly available and inexpensive and demonstrate great homogeneity.

Cell lines are easier to culture and surpass the primary cells in terms of reproducibility.

A549 cells

A549 cells are adenocarcinoma human alveolar basal epithelial cells with Type II characteristic morphology. They were obtained from a 58-year-old Caucasian male in 1972. This model is commonly used to assess respiratory cytotoxicity and to study lung cancer and the development of drug therapies against it (Foster et al., 1998). Further, this cell line does not form tight junctions and do not show microvilli when exposed to ALI conditions (Öhlinger et al., 2019).

16HBE14o- cells

This human bronchial epithelial cell line is widely used to model barrier function of the airway epithelium and to study respiratory ion transport. 16HBE14o- cell lines are immortalised cells derived from a male heart-lung patient, and retain the characteristic features of normal differentiated bronchial epithelial cells including a cuboidal morphology, cytokeratin expression and the ability to form tight junctions (Forbes et al., 2003). However, a decrease in TEER values has been shown in 16HBE14o- cells when cultured in ALI conditions. Further, even when this model resembles the bronchial epithelium, a failure to produce mucus when cultured in air-liquid has been demonstrated.(Ehrhardt et al., 2002)

Calu-3 cells

Calu-3 cells are derived from human bronchial submucosal glands and are widely available. It is a well-differentiated and characterized cell model that is easy to culture, relatively inexpensive, and display the main features of lung

epithelial cells (Florea et al., 2003; Macdonald et al., 2013; Y. Zhu et al., 2010). Cells resemble native airway epithelia in that they secrete polarized mucin, and express apical chloride and bicarbonate ion transporters, secretory IgA receptors and Na+-glucose transporters (Mathias et al., 2002).

When cultured at ALI, Calu-3 form a pseudostratified columnar epithelium that mimics the native bronchial epithelium by exhibiting longer microvilli and excreting mucus (Kreft et al., 2015). Further, absorption has been shown to correlate strongly with ex-vivo models of the lungs, suggesting that this cell culture might provide insight into the absorption rate of pulmonary drugs (Bosquillon et al., 2017; Mathias et al., 2002).

1.4. Strategies for transporter analysis

To study the activity of protein transporters, several inhibitors have been used in combination with substrates (Oostendorp et al., 2009; Panduga et al., 2017; Weiss et al., 2007). However, data must be interpreted with care as OCT/OCTNs have demonstrated overlapping substrate and inhibitor affinities. Genetic suppression of a drug transporter might clarify the specific interactions between drugs and transporter. Several technologies have been used to suppress the function and to illustrate the role of transporters in physiology and ADME of drugs and endogenous compounds. (Simoff et al., 2016)

1.4.1. Stable transfected cells

Rapid progress in understanding the participation and substrate specificity of different OCT has been made by stably transfecting cells with a single OCT transporter. (Salomon et al., 2015; Sato et al., 2008)

In a previous study, Müller et al,. (Müller et al., 2005) compared the uptake of OCT1-3 transporters in small intestinal cells (Caco-2) with transfected kidney cells (HEK293) and chinese hamster ovary (CHO). Data suggested that transfected cells were not resembling the native expression levels and cellular localisation of the organic transporters. In addition, Salomon et al,.(Salomon et al., 2015) compared interaction of $ß_2$-antagonists with organic cation transporters (OCT) in lung epithelial cells versus HEK-293 cells transfected with OCT1−3. However, they failed to consider the expression levels of the organic transporters between transfected cells and lung epithelial cells, in relation to drug uptake. Indeed, Ciarimboli and Schlatter (Ciarimboli & Schlatter, 2005) have shown that transfected cells exhibit different regulation and affinity compared to intact cells. Therefore, studies should be performed by using lung epithelial cells that express the specific transporter.

1.4.2. RNA interference (RNAi) knockdown

RNAi is a widely used gene knockdown approach to study gene function in mammalian cells. It was discovered in the nematode *C. elegans,* gene silencing is activated by induced small double-stranded RNA (dsRNA)(Fire et al., 1998). Its success is attributed to an endogenous pathway that regulates gene expression via small RNAs. The mechanism is triggered by introducing synthethic small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs), or by suppressing translation of specific mRNAs, induced by microRNA (miRNA)(O'Keefe, 2020; Sledz & Williams, 2005).

By far the most important advantages are that cells can be transfected with relatively high efficiency, low-priced and they can be modified to reduce the

likelihood of off-target effects(Campeau & Gobeil, 2011). Nevertheless, their greatest difference is that RNAi only reduces the gene expression at mRNA level instead of a permanent gene disruption. This method, however, is useful for studying the effect of essential genes (Han, 2018).

1.5. CRISPR-Cas9 technology

Recently, an accessible, easily adapted and programmable tool revolutionised genome editing. The form of clustered regularly interspaced short palindromic repeat (CRISPR)/ CRISPR-associated protein (Cas) system has been adapted from a naturally occurring genome editing system in many archaea and bacteria. The technology enables a fast and accurate alteration of genomic information in mammalian model systems and human cells (R. M. Gupta & Musunuru, 2014; Sander & Joung, 2014). The system relies on two parts: a single-guide RNA (sgRNA) and Cas9 enzyme. The two components form a complex to cleave target DNA sites, as shown in Figure 1-13. The sgRNA consists of 20-nt complementary to target DNA sequence and a protospacer adjacent motif (PAM, sequence 5'-NGG-3'), bind to a trans-activating crispr RNA (tracrRNA), a constant component that forms a stem-loop for Cas9 binding (black in Figure 1-13). The Cas 9 is then guided by the 20-nt sequence which is directly adjacent to the protospacer motif. Thus, the system can be easily retargeted to new DNA sequences by simply modifying the 20-nt guide sequence (Cui et al., 2018).

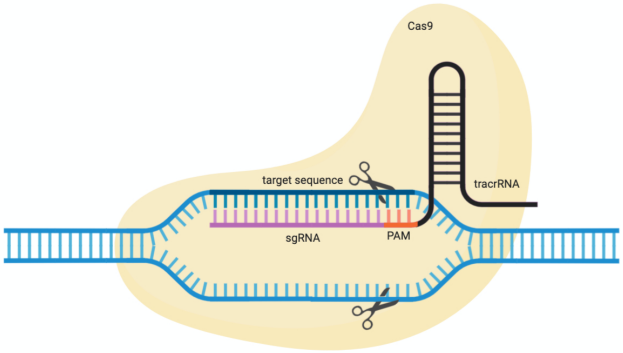


Figure 1-13. Schematic representation of the RNA-guided Cas9 nuclease. The Cas9 protein is targeted to a genomic DNA by a sgRNA that consists of 20-nt guide sequence (orange) and a scaffold (black). The guide sequence pairs with the DNA target (dark blue) that as a requisite has to precede a PAM (5'-NGG, red) sequence motif. Cas 9 then mediates a double-stranded break (DSB) upstream of the PAM.

Upon cleavage, Cas9 induces a double-stranded breaks (DSBs) at the specific genomic loci. DNA then undergoes one of two major pathways for DNA damage repair, the error-prone NHEJ or the high-fidelity HDR pathway. For NHEJ, DSBs are re-ligated, leaving scars in form of insertions or deletions (indels), leading to frameshift mutations and premature stop codons. HDR is an alternative major DNA repair pathway, an introduced exogenous DNA template will fill the gap. It occurs at lower and substantially more variable frequencies than NHEJ. HDR is generally active only in dividing cells, it provides an effective and simple method for making small edits in the genome(Ran et al., 2013).

Over the last years, CRISPR-Cas9 system has been applied to uncover cellular mechanisms and to identify or validate new drug targets (Fellmann et al., 2017; Y. Yang et al., 2016). In a recent study, a stable and complete knockout of the canine Mdr1 was performed using the CRISPR-Cas9 technology in MDCK II cells line (Simoff et al., 2016). It showed up as an excellent and strong tool for

drug transport studies. More recently, major progress has been made in bioinformatics(L. J. Zhu, 2015), allowing an easy finding of targeting sites by scanning the whole genome for the PAM sequence. One major drawback of this technology is to design gRNA with high efficacy and great specificity, for which many computational tools have been developed in order to predict potential off-targets.

1.6. Aim and objectives

The aim of this research is to understand the expression and physiological roles of OCT/OCNs transporters in human airway epithelium in order to elucidate its participation in cellular proliferation, cancer development and characterisation of the mechanisms underlying the transcription of SLC22A1, SLC22A4 and SLC22A5 in the lung epithelia.

Thus, the specific objectives are:

- To explore the expression signature of SLC22A family genes in lung pathophysiology by analysing mRNA microarray datasets from the Gene Expression Omnibus (GEO) database and The Cancer Genome Atlas (TCGA).

- To design and validate CRISPR/Cas9 constructs for generating edited alleles of the SLC22A1, SLC22A4 and SLC22A5 transporters in a well-known cell culture model (HEK293) and lung cell models A549 and Calu-3, and characterise the cellular phenotypes of the knock-out cell lines.

- To characterise the minimal promoter mechanisms of SLC22A1, SLC22A4 and SLC22A5 expression in lung epithelial cells A549, Calu-3 and HEK293 cells. This will be accomplished through luciferase assays integrated with methylation analysis of publicly available data.

CHAPTER 2.
MATERIALS AND METHODS

CHAPTER 2. MATERIALS AND METHODS

2.1. Materials

2.1.1. Buffers

10X Tris/Borate/EDTA (TBE) Buffer.	Tris, 108 g Boric acid, 55 g 0.5 M EDTA pH 8.0 20 ml
Antibiotics (1000x)	50 mg/ml Ampicillin
2X HEPES Buffered Saline (HBS)	280mM NaCl, 16.36 g 10mM KCl, 0.74 g 1.5mM Na_2HPO_4 anhydrous, 0.212 g 12mM dextrose, 2.16 g 50mM HEPES pH7.05, 11.92 g Filter sterilize
Super Optimal Broth (SOB) medium	Tryptone, 20 g Yeast extract, 5g NaCl, 0.5 g KCl, 0.816 g $MgCl_2$ anhydrous, 0.952 g $MgSO_4$ anhydrous, 1.204 g
CCMB80 buffer	1M KOAc, 10 ml $CaCl_2 \cdot 2H_2O$, 11.8 g $MnCl_2 \cdot 4H_2O$, 4.0 $MgCl_2 \cdot 6H_2O$, 2.0 g 10% Glycerol, 100 ml
10X DNA Gel Loading	10% (w/v) SDS, 500 µL 0.5 M EDTA, 200 µL Bromophenol blue, 0.025 g Xylene cyanol, 0.025 g
Luria-Bertani (LB) medium	Tryptone, 10 g NaCl, 10 g Yeast extract, 5 g pH adjusted to 7.0, autoclaved
Luria-Bertani (LB) agar	LB medium Bacto-agar, 7.5 g
SDS	

2.1.2. Plasmids

Plasmids were purchased from Addgene, plasmid maps can be found in the Appendix Section. pGL3-Basic plasmid was purchased from Promega®.

Table 2-1. List of plasmids used for the present study, product code and characteristics.

Plasmid	Details
pX459 (SpCas9-2A-Puro V2.0) Addgene plasmid ID: 62988	Vector type: Mammalian Expression, CRISPR Total vector size (bp): 9200 Selectable markers: Puromycin Bacterial Resistance(s)Ampicillin, 100 µg/mL
pX461- pSpCas9n(BB)-2A-GFP Addgene plasmid ID: 48140	Vector type: Mammalian Expression, CRISPR Cas9 D10A nickase mutant Total vector size (bp): 9300 Bacterial Resistance(s)Ampicillin, 100 µg/mL
pGL3-Basic Promega, cat #E1751	Vector Type: Luciferase Total vector size (bp): 4818 Bacterial Resistance(s)Ampicillin, 100 µg/mL
pJET1.2/blunt ThermoFisher, cat #K1231	CloneJET PCR Cloning Kit Total vector size (bp): 2974 Bacterial Resistance(s)Ampicillin, 100 µg/mL

2.1.3. Cell lines

Various cell lines were employed to achieve the aim of the current study. The human epithelial cell line derived from a lung carcinoma tissue, A549, the non-small-cell lung cancer cell, Calu-3 and the human embryonic kidney cells, HEK293 cells were obtained from the ATCC. Calu-3 cell line (Merck SCC438) was kindly gifted by Dr Emilia Moradi from Optic and Photonic group/Faculty of Engineering/University of Nottingham.

2.2. Methods

2.2.1. Differential Expression and Methylation Analysis

All data was processed and analysed using R version 4.2.2 and various Bioconductor packages, including limma (v 3.54.1), edgeR (v 3.40.2), and DESeq (v 1.38.3) for normalization, differential expression analysis and multiple testing correction.

The LIMMA package in R is a package for linear modelling of gene expression data. It is particularly useful for analysing microarray and RNA-seq data. The package provides a variety of functions for pre-processing and normalizing the data, as well as for fitting linear models and performing statistical tests to identify differentially expressed genes. The linear model used in LIMMA is based on the general linear model (GLM), which is a flexible framework for modelling the relationship between a response variable and one or more predictor variables. The package also includes a number of methods for multiple testing correction and for identifying and visualizing differentially expressed genes.

Microarray datasets

The GSE67472 dataset contained 62 samples with mild-to-moderate asthma and 43 healthy airway epithelial brushings analysed using the Affymetrix Human Genome U133 Plus 2.0 Array (GPL16311) , the GSE76925 dataset contained 111 COPD surgically-resected lung tissue and 40 control smokers tissue analysed using the Illumina HumanHT-12 V4.0 expression beadchip array (Barrett et al., 2013).

Pre-processing and normalisation

Data preprocessing is an important step in differential expression analysis with microarray data. Microarray data normalization aims to eliminate systematic variations arising from technical factors unrelated to biological differences. This ensures consistent and comparable gene expression measurements across samples.

For Illumina HumanHT data (GSE76925), neqc() function from the limma package (Shi et al., 2010) was employed. This function performs both background correction and quantile normalization in a single step. Background correction utilizes negative control probes to remove non-specific signal, while quantile normalization adjusts the intensity distribution across samples, ensuring comparability.

Previous studies have shown that neqc() offers superior precision and reduced bias in fold change estimates compared to alternative methods (Shi et al., 2010).

The affy package was used to normalize the Affymetrix data (GSE67472). This package implements Robust Multi-array Average (RMA) normalization, a well-established method specifically designed for Affymetrix data (Bolstad et al., 2003). RMA employs mismatch probes for background correction and applies a multi-step approach, including quantile normalization, to address intensity-dependent biases and technical variations. This method adjusts the distribution of the data so that it is the same across all the slides.

Analysis

Data was fitted to a linear model using the function LmFit from the limma package (Smyth, 2005). The function requires the transformed and normalized data as input and the design matrix as an argument. These fitted linear models can then be used for further statistical analysis to identify differentially expressed genes (Law et al., 2018). Results were corrected for multiple testing using the eBayes function.

A threshold of 0.05 for the Benjamini-Hochberg adjusted p-value was used to determine the significantly different genes (Benjamini et al., 2001). All expression plots were created with the R package, ggplot2 v.3.4.1.

Table 2-2. Clinical characteristics of patients in the included datasets used in the analysis obtained from TCGA and the GEO database (Barrett et al., 2013).

Characteristic	TCGA-LUSC	TCGA- LUAD	GSE67472	GSE76925
Lung disease	Lung squamous cell carcinoma	Lung Adenocarcinoma	Asthma	Chronic obstructive pulmonary disease (COPD)
Platform	Illumina Hiseq2000 RNA sequencing platform	Illumina Hiseq2000 RNA sequencing platform	Affymetrix Human Genome U133 Plus 2.0 Array	Illumina HumanHT-12 V4.0 expression beadchip
Tissue	Bronchus and lung	Bronchus and lung	Airway epithelial brushings	Lung tissue
Samples	**553 (100.0%)**	**598 (100.0%)**	**105 (100.0%)**	**151 (100.0%)**
Case	502 (90.8%)	539 (90.1%)	62 (59.0%)	111 (73.5%)
Normal	51 (9.2%)	59 (9.9%)	43 (41.0%)	40 (26.5%)
Gender	**553 (100.0%)**	**598 (100.0%)**	**105 (100.0%)**	**151 (100.0%)**
Female	145 (26.2%)	325 (54.3%)	51 (48.6%)	84 (55.6%)
Male	408 (73.8%)	273 (45.7%)	54 (51.4%)	67 (44.4%)
Stage	**498 (90.1%)**	**531(88.8%)**		
I	245 (44.3%)	295 (49.3%)		
II	162 (29.3%)	126 (21.1%)	NA	NA
III	84 (15.2%)	84 (14.0%)		
IV	7 (1.3%)	26 (4.3%)		
Survival status	**553 (100.0%)**	**598 (100.0%)**		
Survival	304 (55.0%)	379 (63.4%)	NA	NA
Death	249 (45.0%)	219 (36.6%)		

Abbreviations: TCGA, The Cancer Genome Atlas; NA, not available

RNAseq

Specific tumour type and matched normal tissue gene expression and clinical data were obtained from the TCGA via the GDC portal (https://portal.gdc.cancer.gov/) using TCGAbiolinks v.2.26.0 (Colaprico et al., 2016), raw data was further processed in R. Transcriptomic profiles and relevant clinical information of patients with LUAD of a total of 539 LUAD tissues and 59 adjacent non-tumour tissues were retrieved. LUSC dataset contained transcriptomic profiling and clinical information from 502 solid tumours and 51 adjacent normal tissue. Table 2-2 lists the datasets used for analysis and clinical information of the samples.

Pre-processing

RNA-seq data analysis often involves pre-processing and filtering steps to ensure reliable downstream analysis. The raw RNA-seq read counts were normalized using the counts per million (cpm) method. This method adjusts for differences in library size between samples, allowing for meaningful comparisons across the dataset. After normalization, a filtering step was implemented to remove genes with low expression levels. Two criteria were employed in conjunction, to filter lowly expressed genes. A CPM threshold of 1.0 was used to identify genes with potentially meaningful expression levels. This threshold ensured that genes were expressed at a minimum level across a substantial portion of the samples. To further enhance the quality of the filtered dataset, a requirement for a gene to be expressed with CPM > 1.0 in at least 30 samples was imposed. This criterion aimed to capture genes with more consistent and reliable expression patterns.

Kaplan-Meier estimator

To assess the potential association between SLC22A gene expression and patient survival in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) patients, a Kaplan-Meier (KM) estimator was employed. The KM estimator is a non-parametric statistic used to estimate the probability of event-free survival (survival) over time, often employed in survival analysis studies (Goel et al., 2010).

Clinical data for LUAD and LUSC patients, including vital status, were retrieved from The Cancer Genome Atlas (TCGA) database. Patients were categorised into two groups (low and high) based on the median expression value of the SLC22A gene. The survfit function from the survival package (version 3.4-0) in R software (version 4.2.2) was used to estimate the KM survival curves for each patient group. The survminer package (version 0.4.9) was then utilized to generate the corresponding visual representations of the survival curves.

The log-rank test was employed to compare the survival curves between the two patient groups and assess the statistical significance of any observed differences in survival probabilities.

DNA methylation

To measure DNA methylation, one of the most used techniques in the BeadChip Arrays, which covers around 450,00 CpG sites in different gene regions, including TSS1500, TSS200, 5'UTR, 1stExon, body and 3'UTR. The term "TSS1500" refers to a genomic region spanning 200–1500 bases upstream of the transcriptional start site (TSS) of a gene. Similarly, "TSS200" denotes the region

0–200 bases upstream of the TSS. The "5′UTR" designates the 5′ untranslated region, which is defined as the segment between the TSS and the ATG start site. "1stExon" is an abbreviation for the first exon of the gene, while the "Body" denotes the region between the ATG start site and the stop codon. Furthermore, "3’ UTR" represents the 3’ untranslated region, which lies between the stop codon and the poly-A tail of the gene.

In the context of CpG sites, methylation levels are quantified using the beta value (b), calculated as the ratio of methylated signal intensity (M) to the sum of methylated (M), unmethylated (U), and an offset value (a). The variables M and U represent the respective signal intensities for methylated and unmethylated states, with the constraint that both M and U are positive values (Weinhold et al., 2016).

All data from the Cancer Genome Atlas Program (TCGA) was retrieved from Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov/). The raw data of gene expression profiles for asthma and chronic obstructive pulmonary disease (COPD) were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) via getGEO, raw data was further processed and analysed in R. Table 2-2 lists the datasets used for analysis and characteristics of the samples. Gene expression analysis from these datasets was performed using RNA sequencing (RNAseq) and microarray data, accordingly.

Statistical Analysis

All statistical analyses were performed using the R programming language (version 4.2.2). A two-sided p-value < 0.05 was regarded as statistically significant.

2.2.2. Methylation Analysis

All data analyses in Chapter 4 were performed using R (R version 4.2.2). DNA methylation and clinical data for LUAD and LUSC were downloaded from the Cancer Genome (TCGA) portal.

The methylation level is expressed as β value which represents the methylation intensity and total array intensity, between 0 (lower level of methylation) and 1 (higher level of methylation). Poor performing probes, cross reactive probes, Y chromosomes probes and SNP probes were excluded in the data processing.

Correlation analysis

Correlation analysis was performed between DNA methylation data and expression data, using Pearson's correlation with an FDR-corrected p-value threshold of 0.05.

2.2.3. General Molecular Biology Techniques

All DNA oligos used in this study were purchased from Sigma-Aldrich.

Isolation of genomic DNA

At least 1 x 10^6 cells were pelleted, DNA was isolated using the GenElute™ Mammalian Genomic DNA Miniprep kit (Merck, cat #G1N350), according to the manufacturer's instructions. Genomic DNA was resuspended in water before quantification by Nanodrop™.

Polymerase Chain Reaction (PCR)

PCRs were run for several different experiments, including genotyping, colony screening, T7 endonuclease assay, cloning and generation of inserts. Phusion® High-Fidelity DNA Polymerase (NEB, cat #M0530) was used for all reactions, conditions varied depending on the oligo sequence and length of fragment. Unless otherwise stated, PCR reactions were prepared in a final volume of 50 µl with 0.1 µM primers and 250 ng of genomic DNA or 10 ng of plasmid DNA, depending on the template required. Thermocycling conditions for a routine PCR consisted of 98°C for 5 min followed by 35 cycles of 10 sec denaturation step, 30 sec of annealing temperature, determined according to the T_m of oligos used and a 72 °C extension step, time was calculated depending on the product size (30 sec per kb). Following a final 5 min extension step at 72°C, samples were then cooled down to 4°C.

PCR products were verified on a 1% - 2% (w/v) agarose gel, depending on product size. Primer excesses were removed from reactions using Monarch® PCR & DNA Clean-up Kit (NEB, cat #T1030S), amplicons were eluted in 10 µl of nuclease-free water.

DNA Restriction digest and Gel extraction

As required for some experiments, DNA was digested by endonuclease restriction. Restriction sites were either inserted by PCR or used from the multiple cloning site (MCS) region in plasmid vectors, according to the experiment.

Digestions were performed for 2 hours using the appropriate enzyme, recommended buffer and incubation temperature according to the manufacturer. Digestions were run on a 2% agarose gel. Bands were carefully excised under UV light and further extracted using Monarch® DNA Gel Extraction Kit (NEB, cat #T1020S),and eluted in 10 µl of nuclease-free water and quantified by Nanodrop™.

DNA Ligation

Digested purified DNA fragments or annealed and phosphorylated oligos were used as inserts in ligation reactions. A mass ratio of 1:5 of plasmid to insert was used for all ligations with T4 DNA Ligase (NEB, cat #M0202), to a final volume of 10 µl. Reactions were incubated overnight on a thermocycling ligation consisting of 10°C for 30 sec followed by 30°C for 30 sec (adapted from Lund et al., 1996).

Phosphorylation and annealing of sgRNAs

Top and bottom stands of oligos for each sgRNA design were resuspended to a final concentration of 100 µM and mixed according to conditions in Table 2-3 using T4 Polynucleotide Kinase (T4 PNK) (NEB, cat # M0201S).

Table 2-3. Preparation of mixtures for sgRNA annealing and phosphorylation

Component	Volume (µl)
sgRNA top (100 µM)	1
sgRNA bottom (100µM)	1
T4 ligation buffer, 10X	1
T4 PNK	1
H_2O	6

Oligos were phosphorylated and annealed in a thermocycler by using the following parameters: 37°C for 30 min, 95°C for 5 min: ramped down to 2°C at 5°C/min. Double-stranded oligos were diluted to 1:200 ratio with RNA-free water.

Transformation protocols

Preparation of competent cells

SOB medium was inoculated with an overnight culture and grown at 25°C to an OD_{600} of 0.3. Once reached, bacterial culture was kept on ice for 10 min before centrifugation at 3000 rpm for 10 min in a chilled centrifuge. Cells were resuspended in ice-cold CCMB80 buffer and incubated on ice for 20 min. After incubation, cells were spun down and gently resuspended in ice-cold CCMB80 buffer. Suspension was then aliquoted and stored in -80°C until further use.

Transformation of competent E.coli cells

Cells were allowed to thaw on ice before use. 1-5 µl containing 1 pg-100 ng of plasmid DNA or ligation was added to an aliquot of competent cells. Mixture was incubated on ice for 30 min. Cells were then placed into a 42°C water bath for 30 sec and placed on ice for 5 min. Luria Broth (LB) media was added to a final volume of 1 ml and cells were placed in a shaking incubator at 37°C for 45

min to recover. Finally, cells were spread onto a selection plate and incubated overnight at 37°C.

Colony screening

Up to 20 colonies were analysed for the presence of the insert using specific primers according to the plasmid DNA. Colony PCR was performed with Taq DNA Polymerase (NEB, cat #M0273S) thermocycling conditions were as following: 95°C, 5 min; 30 cycles of 95°C for 30 sec for denaturation; annealing temperature was calculated according to the T_M of oligos, 30 sec at 68°C (1min/kb) and a final extension at 68°C for 5 min. PCR products were further analysed on agarose gel. Colonies were then propagated overnight in a shaking incubator at 37°C in LB medium with its respective antibiotic.

Isolation of Plasmid DNA

Plasmid DNA was isolated from propagated *E.coli*. Isolation was performed using GenElute® Plasmid Miniprep Kit (Merck, cat #PLN350), or GenElute™ HP plasmid maxiprep kit (Merck, cat #NA031), depending on the culture volume. Plasmid isolation is based on the alkaline lysis method and was performed according to manufacturer's instructions. Plasmid DNA was finally eluted in nuclease-free water and quantified by Nanodrop™.

Sanger Sequencing

PCR products and plasmids were sent to Source BioScience for Sanger sequencing. Oligos were selected according to the vectors, as listed on Table

2-4. Samples sent for genotyping where sequenced with their respective forward oligo.

Table 2-4. Oligos used for sequencing of vectors.

Oligo	Sequence (5' – 3')	Vector
U6-Fwd primer	ACTATCATATGCTTACCGTAAC	pX459, pX461
PGL3_seq	CTAGCAAAATAGGCTGTCCC	PGL3
pJET1.2_REV	AAGAACATCGATTTTCCATGGCAG	pJET1.2

2.2.4. General Tissue Culture

Maintenance of cell lines

Calu-3, HEK293 and A549 cells were grown in Dulbecco's Modified Eagle's Medium/Nutrient Mixture F-12 Ham (Merck, cat #D8437-500ML) supplemented with 10% Fetal Bovine Serum (FBS) (Merck, Non-USA origin, cat #F0804), 100 UI/ml penicillin and 100 µg/ml streptomycin (ThermoFisher, cat #11528876), 20 mM L-glutamine solution (Merck, cat # G7513-100ML), and 1% non-essential amino acids (Merck, cat #M7145-100ML). All cells were maintained in a humidified 5% CO_2 environment at 37°C and fresh medium was provided every 1-2 days. All experiments with Calu-3, HEK293 and A549 cells were performed between passages 26-31, 14-25 and 27-31, respectively.

Upon reaching 90% confluence, cells were trypsinised using 1X Trypsin (ThermoFisher, cat # 11538876). Calu-3, HEK293 and A549 cell lines were sub-cultured maintaining a split ratio of 1:3, 1:6, and 1:6, respectively.

Dose response curve for antibiotic selection of mammalian cells

Cells transfected with pX459 vectors were screened against puromycin resistance at 48 h post-transfection. The optimal antibiotic concentration for selecting stable cell colonies was determined by mammalian cell sensitivity.

Assessment of antibiotic concentration for cell selection

Cells were subjected to increasing amounts of antibiotic to determine the minimum antibiotic concentration needed to select only transfected cells with resistance.

Briefly, cells were seeded in a 96-well plate at a density of $1x10^4$ cells/well. Media containing antibiotic was replaced every 2 days for up to a week. Cell viability was assessed by MTT assay. Viable cells with active metabolism convert MTT into formazan. The MTT assay is a measure of the metabolic activity of the cells analysed; the more metabolic activity in the sample, the higher the signal.

10 µl of MTT reagent (5mg/ml) (Merck, cat #M5655) was added to cells after incubation with puromycin and reaction was incubated for 3 h at 37°C, 5% CO2. Media was removed and crystals were solubilised with Dimethyl Sulfoxide (DMSO) (Merch, cat # C6164). Plates were placed on an orbital shaker for 10 min in the dark. Absorbance was measured at 570 nm.

The optimal puromycin concentration is the lowest dose that will kill 90-99% of non-selected cells within 7 days. HEK293 and A549 viability decreased in

response to a dose of 0.5 µl/ml puromycin, while Calu-3 viability decreased with 1.5 µl/ml puromycin, as presented in Figure 2-1.

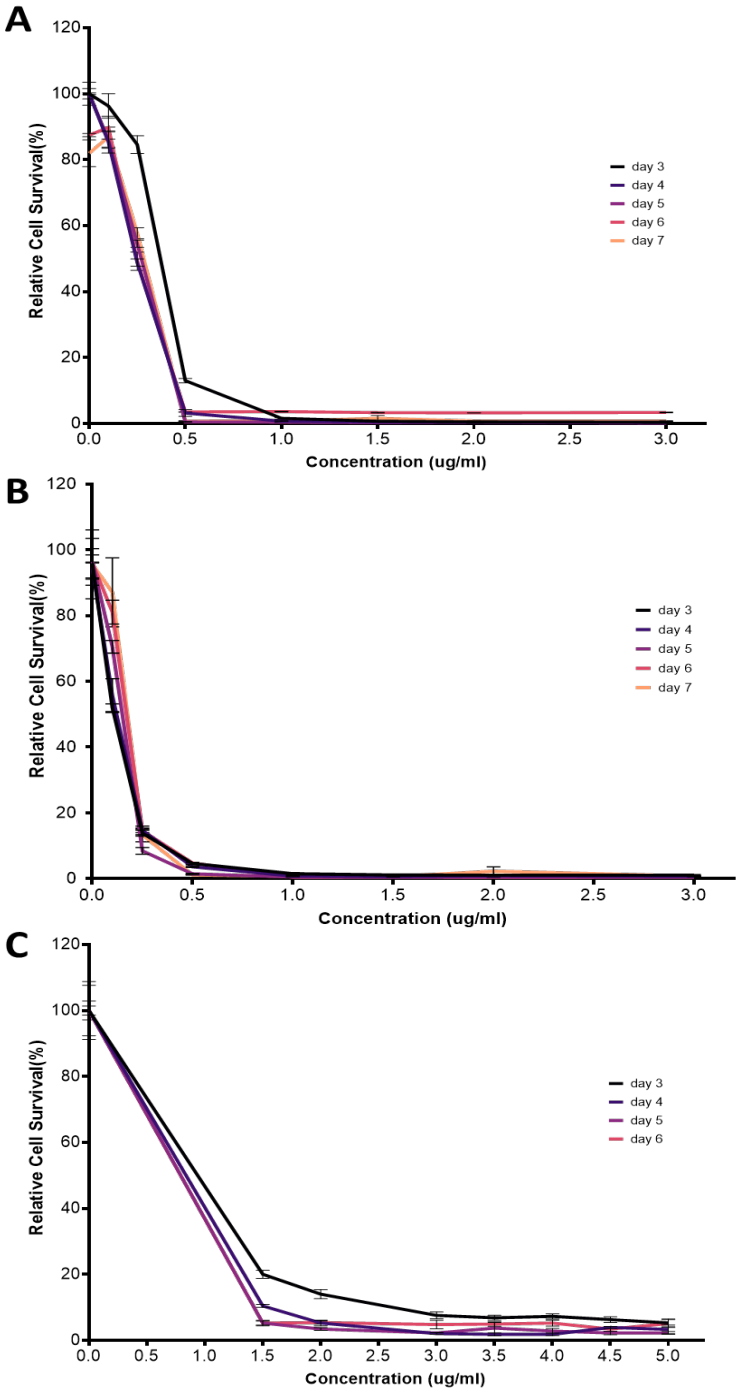


Figure 2-1. Determination of dose-response for puromycin selection of A) A549, B) HEK293 and C) Calu-3 . Curves were constructed from day 3-7 using a range of puromycin concentrations. Data is presented as the mean ± SEM. Each represent technical triplicates for 2 biological replicates.

Antibiotic Selection

Briefly, culture medium was replaced with fresh medium containing the respective concentration of puromycin, according to the cell line; to remove non-transfected cells, medium was replaced every 2 days. After 1 week of selection, medium was replaced with fresh medium without antibiotic. Cells were expanded on a 60 mm dish and allowed to grow again for further genome editing analysis.

Cryopreservation and Thawing cells

Cultured cells were frozen and stored for preservation. Briefly, cells were suspended with 1X trypsin and centrifuged at 2000rpm for 5 min, pellet was resuspended in freezing medium consisting of 90% FBS and 10% Dimethyl sulfoxide (DMSO) (Merck, cat # D1435). Cells were then transferred into 2 mL cryogenic vials and placed into a Mr. Frosty Freezing Container (ThermoFisher, cat #5100-0001) which was then stored at -80°C overnight. The freezing container is filled with isopropanol to allow the samples to achieve a slow cooling rate of -1°C / minute which is the optimal rate for cell preservation. Following overnight storage, cryogenic vials were transferred into liquid nitrogen for long-term storage.

To defrost cells, cryogenic vials were taken out of the liquid nitrogen tank and thawed in a 37°C water bath. Cells were added to a 15 mL conical tube containing warm medium and then centrifuged at 2000rpm for 5 min. Cells were then resuspended with fresh culture medium and placed into a T25 flask to optimize recovery.

2.2.5. Generation of CRISPR components

Design of targeting components and selection of vector for CRISPR-Cas9

All-in one plasmids have been previously used for the simple and rapid construction and simple delivery into the cells.

To generate the expression construct, the sgRNAs were cloned into the pSpCas9(BB) vector. The plasmid contains the Cas9 protein, a sgRNA scaffold and a BbsI cloning site for insertion of the guide sequence (Ran et al., 2013). A selective resistance to Ampicillin that allows bacterial screening and Puromycin resistance for selection of transfected mammalian cells. Plasmid map can be found in the Apendix.

The CRISPRseek package was used to design the CRISPR targets and analyse the sgRNAs for genomic editing of SLC22A1, SLC22A4 and SLC22A5 (L. J. Zhu et al., 2014). The package identifies candidates for CRISPR-Cas9 within a given input. Prediction of the relative off-target cleavage rates is based on the Cutting Frequency Determination (CFD) scoring, calculated by using the percent activity values provided in a matrix of penalties based on mismatches of each possible type at each position within the guide RNA sequence. (Doench et al., 2016).

cDNA was obtained from NCBI, in order to design guides that would target the exonic regions. Criteria used to choose the sgRNAs per gene were, the position of the sgRNA should be in an exonic region, with complete match and efficacy should be as high as possible. The last criteria was relaxed if sequence showed

a perfect-match for the intended target. Target sequences obtained for the SLC22A1, SLC22A4 and SLC22A5 transporters can be found in Table 2-5.

Table 2-5. Targeted sequences obtained from CRISPRseek for the pX459 plasmids. Bold letters indicate the 3'PAM sequence (NGG). Genomic location is refered for the hg38 genome assembly.

gRNA name	Target sequence	Genomic location	Strand
SLC22A1_Exon5	TTGTGATAACAGCCACCGAG**GGG**	chr6:160136234	-
SLC22A1_Exon4	GGTGGCGATCATGTACCAGA**TGG**	chr6:160133991	+
SLC22A4_Exon3	GAACGTTCTCTTCGCAACCA**TGG**	chr5:132313626	+
SLC22A4_Exon2	CACGCCTACGAAGAACAGGG**AGG**	chr5:132312204	-
SLC22A5_Exon4	AATGTGCTGTTCGTGACCAT**GGG**	chr5:132384160	+
SLC22A5_Exon3	TCCAGTCGTCCTCACACACC**AGG**	chr5:132378383	-

Two pairs of sgRNA were designed to target SLC22A1, SLC22A4, SLC22A5 individually. Each pair of sgRNA targets a different exon per gene, as listed in Table 2-5.

Paired sgRNAs were used to excise an exon in the human Homeobox protein, encoded by the EMX1 gene. The EMX1 gene has been widely used as a validated CRISPR editing control, which serves as an experimental control for the wt Cas9 (Duan et al., 2014; Ran et al., 2013). EMX1 cells were used as a control for genome editing and functional analysis.

In order to clone the sgRNA in the plasmid, the sequence 5'-CACCG-3' was added to each top of the sgRNA to generate an overhang compatible with BbsI the digested sites, where the guanine (G) is added for efficient U6 transcription of the sgRNA (D. Wang et al., 2019). In addition, all reverse complement sequences 5'-AAAC-3' were added to the 5' end and a C in the 3' end as shown in Figure 2-2

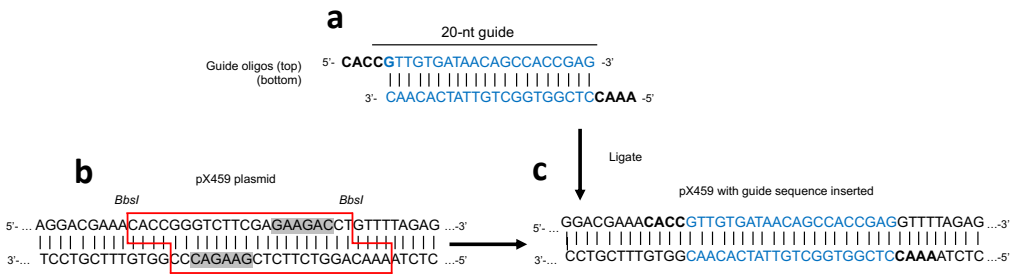


Figure 2-2. Schematic representation of the insertion of guide sequence oligos into the plasmid a) The oligos (blue) contains the overhangs (bold) for ligation into the pair of BbsI sites in the vector, both matching the ones in the plasmid (the top oligo is the 20-bp sequence preceding the 5'-NGG in genomic DNA). b) Digestion of pX459 with BbsI allows the replacement of the restriction sites (red outline) with a direction insertion of annealed oligos. Recognition sites are marked with grey rectangles c) pX459 plasmid with guide sequence (blue).

Design of targeting components and selection of vector for CRISPR-Cas9 nickase

In comparison to wild type Cas9, which needs only one sgRNA to cut both strands of the target DNA, Cas9 nickases use two adjacent sgRNAs to generate a double-strand break. Guides must target opposite strands of the genomic DNA and can be oriented with their PAM sites facing toward each other (PAM-in), or apart from each other (PAM-out).

Plasmid pSpCas9n(BB)-2A-GFP (PX461) was used for Cas9 nickase strategy. This plasmid contains the Cas9n (D10A nickase mutant) from *S. pyrogenes,* a sgRNA scaffold with the BbsI restriction site as cloning site and 2A-EGFP as reporter gene (Ran et al., 2013). Plasmid map can be found in the Appendix.

Guides were designed with PAM-out conformation, as previous studies have concluded that this orientation yields a higher rate of indel formations when used with D10A nickase (Amo et al., 2022; Schubert et al., 2021). Table 2-6 lists the location of the target sites chosen for the CRISPR nickase strategy.

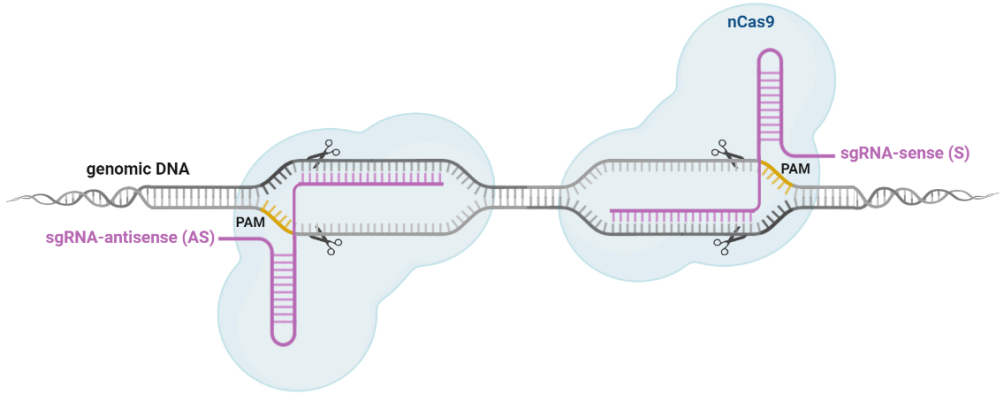


Figure 2-3. Paired sgRNAs with PAM sites facing outwards (PAM-out). The PAM-out configuration supports robust genome editing. Created with BioRender.com

As two sets of sgRNAs are required for the nickase strategy, a pair of sense(S) and antisense (AS) sgRNAs were designed for the double-nicking strategy (Figure 2-3). Similar to the design of sgRNAs for pX459, the sequence 5'-CACCG-3' was added to each top guide to generate compatibility the BbsI restriction site while, 5'-AAAC-3' was added to the 5' end and a C to the 3' end to the bottom guides as noted in Table 2-7.

Table 2-6. Targeted sequences obtained CRISPRseek for the pX461 nickase plasmids. Bold letters indicate the 3'PAM sequence (NGG). Genomic location is refered for the hg38 genome assembly.

gRNA name	Target sequence	Genomic location	Strand
SLC22A1_nick_Exon1_S	TGTAGACCCCCTGGCTAGCC**TGG**	chr6:160122241	+
SLC22A1_nick_Exon1_AS	GCGCCTGCACTGGCCAAGGA**AGG**	chr6:160122183	-
SLC22A5_nick_Exon1_S	GACCGCCTTCCTGGGCGAGT**GGG**	chr5:132369993	+
SLC22A5_nick_Exon1_AS	AGGCCCACAGAGCGCGGCCT**GGG**	chr5:132369939	-

Cloning of sgRNAs constructs into the CRISPR-Cas9 vectors

Firstly pX459 and pX461 plasmids were linearized with BbsI (ThermoFisher, cat #ER1011) at 37°C, digestion was ran on a 1% agarose gel and band was extracted according to previous protocol. Ligation was set up with the phosphorylated double-strand oligos as insert and the linearized pX459 or px461 as vector, accordingly. Figure 2-2 shows the representation of the scarless cloning of the guide sequence oligos into the plasmid, oligos designed for the guide sequences contain matching sequences to the overhangs left after digestion.

Table 2-7. Oligos used to build the sgRNA. Bold letters highlight the nucleotides added for complementarity with the overhangs created after digestion with BbsI.

sgRNA name	Sequence (5' – 3')
SLC22A1_Exon5	Top - **CACCG**TTGTGATAACAGCCACCGAG
	Bottom - **AAAC**CTCGGTGGCTGTTATCACAA**C**
SLC22A1_Exon4	Top – **CACCG**GGTGGCGATCATGTACCAGA
	Bottom - **AAAC**TCTGGTACATGATCGCCACC**C**
SLC22A4_Exon3	Top - **CACCG**GAACGTTCTCTTCGCAACCA
	Bottom - **AAAC**TGGTTGCGAAGAGAACGTTC**C**
SLC22A4_Exon2	Top - **CACCG**CACGCCTACGAAGAACAGGG
	Bottom - **AAAC**CCCTGTTCTTCGTAGGCGTG**C**
SLC22A5_Exon4	Top - **CACCG**AATGTGCTGTTCGTGACCAT
	Bottom - **AAAC**ATGGTCACGAACAGCACATT**C**
SLC22A5_Exon3	Top - **CACCG**TCCAGTCGTCCTCACACACC
	Bottom - **AAAC**GGTGTGTGAGGACGACTGGA**C**
SLC22A1_nick_Exon1_S	Top - **CACCG**TGTAGACCCCCTGGCTAGCC
	Bottom - **AAAC**GGCTAGCCAGGGGGTCTACA**C**
SLC22A1_nick_Exon1_AS	Top - **CACCG**GCGCCTGCACTGGCCAAGGA
	Bottom - **AAAC**TCCTTGGCCAGTGCAGGCGC**C**
SLC22A5_nick_Exon1_S	Top - **CACCG**GACCGCCTTCCTGGGCGAGT
	Bottom - **AAAC**ACTCGCCCAGGAAGGCGGTC**C**
SLC22A5 nick_Exon1_AS	Top - **CACCG**AGGCCCACAGAGCGCGGCCT
	Bottom – **AAAC**AGGCCGCGCTCTGTGGGCCT**C**
EMX1	Top – **CACCG**GTCACCTCCAATGACTAGGG
	Bottom - **AAAC**CCCTAGTCATTGGAGGTGAC**C**

Plasmids adopted the name according to the respective sgRNA inserted into the multiple cloning site.

Generation of Cas9-sgRNA ribonucleoproteins (RNPs)

Direct delivery of CRISPR/Cas9 system as a ribonucleoprotein (RNP) complex has emerged as a powerful method for genome editing. This can accomplish a higher efficiency and specific genome editing when compared to traditional Cas9 approaches (DeWitt et al., 2017; S. Zhang et al., 2021).

To generate RNP complexes, sgRNAs were generated by *in vitro* transcription, where PCR products were used as a template. As seen in Figure 2-4, the forward primer introduces the T7 promoter into the template for transcription to occur.

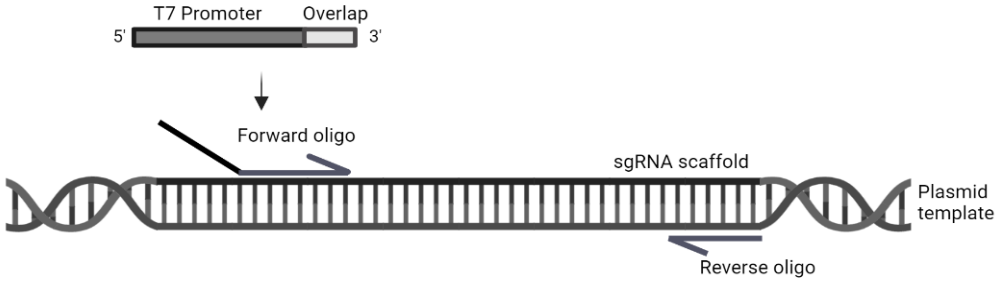


Figure 2-4. Overview of oligo design for in vitro sgRNA generation. The forward-specific oligo contains the T7 promoter sequence, and ~20 nucleotides for overlap with the sgRNA from the plasmid template.

Three forward oligos were designed for each gene, using the pX459 plasmids as templates. Table 2-8 lists the target-specific oligos used for the generation of *in vitro* sgRNA, gRNA_RV oligo was used as reverse sequence for the assembling all sgRNAs. Forward oligos contain the T7 promoter sequence and an overlap sequence complementary to the sgRNA scaffold from the pX459 plasmids.

Table 2-8. Primers used for in vitro transcription (IVT) of sgRNAs from their respective template. gRNA_RV oligo was used as reverse oligo for all PCR reactions. T7 promoter sequence is underlined.

Oligo	Sequence (5' – 3')	Plasmid template
gRNA_RV	AAAAGCACCGACTCGGTGCCAC	
T7_22A5_FWD	TCCTTAATACGACTCACTATAGGAATGTGCTGTTCGTGA	pX459_SLC22A5_Exon4
T7_22A1_FWD	TCCTTAATACGACTCACTATAGGTTGTGATAACAGCC	pX459_SLC22A1_Exon5
T7_22A4_FWD	TCCTTAATACGACTCACTATAGGGAACGTTCTCTTCG	pX459_SLC22A4_Exon3

PCR was carried according to protocol described in Section 2.2.3. Amplicons were subsequently cleaned-up and used for the *in-vitro* RNA transcription.

In-vitro RNA transcription

In vitro transcription is a simple procedure that allows for DNA template to be transcribed by a RNA polymerase in the presence of ribonucleoside triphosphates (rNTPs).

PCR fragments amplified from pX459 plasmids were used as template for in vitro RNA transcription using the HiScribe™ T7 Quick High Yield RNA Synthesis Kit (NEB, cat # E2050S), according to the manufacturer's protocol. The reaction was set up in a total volume of 30 µl containing the NTP buffer mix, T7 RNA Polymerase mix and amplified fragments. Reactions were incubated overnight at 37°C.

Double-stranded DNA (dsDNA) was removed with DNase I and RNA was purified with Monarch® RNA Cleanup Kit (NEB, cat # T2040L) by following the spin column manufacturer's instructions. Spin columns remove unincorporated nucleotides, proteins and salts. RNA was eluted in nuclease-free water and quantified by Nanodrop™.

Assessment of CRISPR: detection of mismatches by the T7-E1 assay

To determine gene targeting efficiency of the CRISPR guides, a T7 endonuclease I (T7E1) assay was performed. T7E1 enzyme recognizes and cleaves mismatched dsDNA. The T7E1 assay can be used as a mutation

detection assay based on the ability of T7E1 to determine between homo- and heteroduplex DNA.

When Cas9 generates a double-stranded break (DSBs) in the targeted region, the area is repaired by non-homologous end joining (NHEJ) if a template is not present. This process is error-prone and often leads to gene alterations such as insertions or deletions (indels) that can be recognized and cleaved by T7E1.

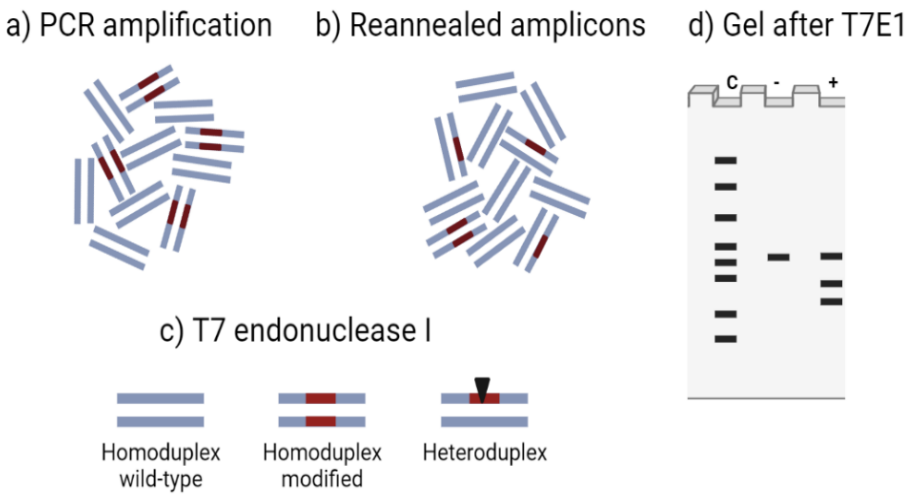


Figure 2-5. T7 Assay. Amplified DNA from edited and wild-type cells are denatured and re-annealed, forming homoduplex wild-type, homoduplex modified and heteroduplex fragments. T7 endonuclease I recognises mismatched DNA and cleaved heteroduplexes complex, resulting in short fragments. Mixtures are then analysed on agarose gel

After recovery from antibiotic selection, cells were harvested by trypsinization and DNA was isolated according to Section 2.2.2. Non-transfected cells were used as a negative control for each pair of primers.

Table 2-9. Primers used for amplification and genotyping of the CRISPR-Cas9 efficiency for each targeted transporter, according to the plasmid used for gene editing.

Genotyping	Oligo (5' – 3')	Product size (bp)
pX459_SLC22A1	FWD, CATCCTCTTGAGGGATTACAGC	239
_Exon5	REV, TTAAATCAGCAGGAGGCAACTT	
pX459_SLC22A4	FWD, GGGAGGATGTGACAGAGAAAAA	206
_Exon3	REV, CCCACGATGACAAATAACACAG	
pX459_SLC22A5	FWD, TTTTCCAGCTGGTTATCTGTCA	250
_Exon4	REV, AAAGGTAGGTGATGGGATGATG	
pX459_SLC22A1	FWD, ATAACGTCCACACCTCCTGTTT	250
_ Exon4	REV, GAAGGGCCTCACCAGTAGTAGA	
pX459_SLC22A4	FWD, CGTGCTAATATTCCCTCAGAGC	262
_ Exon2	REV, CCAAGATAGGCAGTGATTGACA	
pX459_SLC22A5	FWD, TGTTCTGACTTCATTTTCCAGG	288
_ Exon3	REV, TGCGCTGAAAGAATACTACCAA	
pX461_SLC22A1	FWD, TCTGGAGCAGGTTGGGGAGTCTG	516
_nick_Exon1	REV, ACACCCACGAACTGCACAATAAACA	
pX461_SLC22A5	FWD, TTCCCTGGTCGTGCGCCCTATGTAA	495
_nick_Exon1	REV, TCCAGCCCAAGCGCCGAGAAGTT	
pX459_EMX1_E	FWD, ACTACAGTGGTGCCTGGAAA	538
xon3	REV, AGTTTCTCATCTGTGCCCCT	

To analyse CRISPR efficiency, the targeted regions were amplified by PCR, according to their respective gene target as listed on Table 2-9. Prior to digestion, the products were analysed on gel to verify size and appropriate amplification and a fraction was kept for comparison with T7E1. Amplicons were denatured at 95°C for 10 min and re-annealed at −2 °C per second temperature ramp to 85 °C, followed by a −1 °C per second ramp to 25 °C.

Briefly, 5 µl of T7E1 (NEB, cat #M0302S) was added to the heteroduplexed PCR products to a final volume of 10 µl, and reactions were incubated at 37°C for 30 min. If present, T7 endonuclease I will cleave DNA heteroduplex; efficiency will depend on the number of mismatched nucleotides. PCR and

T7E1 samples were analysed on a 2% (wt/vol) agarose gel for the presence or absence of cleaved products.

In order to estimate gene modification levels a densitometry analysis was performed with ImageJ. Apparent percentage of NHEJ of the cell poll was estimated using Equation 2-1 (Chiang et al., 2016; Guschin et al., 2010):

Equation 2-1. Quantification of mutations that result from DNA double-strand break repair via non-homologous end joining.

$$\%\ NHEJ\ events = 100\ \times \left[1 -\ (1 - fraction\ cleaved)^{1/2}\right]$$

where the fraction cleaved is defined as

$$\frac{density\ of\ digested\ products}{density\ of\ digested\ products\ +\ undigested\ parental\ band}$$

2.2.6. Dual-Luciferase Assay

Genetic reporter systems are widely used to study eukaryotic gene expression. These assays involve placing a genetic regulatory element upstream of a reporter gene. The Dual-Luciferase assay relies on two different reporter genes, *Renilla* luciferase (Rluc) and *Firefly* luciferase (Fluc), to evaluate the regulation of gene expression. Rluc is fused to a constitutive promoter, whereas the firefly gene is fused to a test promoter. The assay is performed by subsequentially measuring the luminescence of the *Firefly* and *Renilla* luciferase on the same sample; results are expressed as the ratio of *Fluc* to *Rluc*.

Design and assembling of the luciferase reporter plasmids

A luciferase-containing plasmid is most often used for monitoring the effect of regulatory elements. The pGL3-Basic vector (Promega), contains the firefly luciferase reporter with no promoter and a MCS that allows the insertion of the putative regulatory sequences under investigation.

Upstream sequences and putative regulatory sequences were identified using the UCSC genome browser (Lee et al., 2021). In order to analyse the cell-type-specific functional regions of the SLC22A1, SLC22A4 and SLC22A5 proteins, luciferase assays were carried out in HEK293, A549 and Calu-3 cells. For the purpose of analysing the ability of the first intron to activate expression, the first intron of SLC22A1 gene was cloned into the pGL3 plasmid and were identified with the suffix INT. Table 2 9, lists the position of the amplicons in relation to the transcription start site (TSS) of the protein of interest.

Table 2-10. Location of Luciferase constructs for SLC22A1, SLC22A4 and SLC22A5 in relation to the TSS per protein.

Protein	Construct	Relative position from TSS
SLC22A1		
	pGL3-A1_PF1	-2539 to +18
	pGL3- A1_PF2	- 163 to +18
	pGL3- A1_PF3	- 92 + 18
	pGL3- A1_PF4	- 59 to +18
	pGL3- A1_INT1	+ 772 to + 1791
	pGL3- A1_INT2	+1995 to + 3515
	pGL3- A1_INT3	+2912 to + 3515
SLC22A4		
	pGL3-A4_ENH	- 601 to - 421
	pGL3-A4_PD1	- 334 to + 12
	pGL3-A4_PD2	-61 to + 12
SLC22A5		
	pGL3-A5_P3	-505 to - 289
	pGL3-A5_INT	+ 8524 to + 9562

Figure 2 9, Figure 2 10 and Figure 2 11 show a schematic representation of the localisation of the sequences under investigation with the intron-exon structure of the genes.

To assemble the plasmids, sequence upstream the gene of interest was amplified by PCR from genomic DNA of using the oligos from Table 2-11. To maximise ligation into the final vector, each of the amplified products were individually ligated into the pJET1.2/blunt cloning vector CloneJET PCR Cloning Kit (ThermoScientific™, cat # K1232), according to the manufacturer's protocol. The pJET1.2/blunt is a linearized cloning vector that supports the cloning of blunt-end PCR products. The 5'-end of the vector contains phosphoryl groups so that phosphorylation of the PCR primers is not required. The cloning vector contains a lethal gene that is disrupted by ligating an insert into the MCS. Plasmid map can be found in the Apendix.

In the case of pGL3_A1_INT1, the amplicon has a restriction site for NcoI localised in the sequence, therefore, restriction sites for NheI and XhoI, were added on the forward and reverse oligos, respectively for insertion into the pGL3 vector.

Firstly, PCR products were run on a 2% agarose gel and extracted prior to ligation with the pJET1.2/blunt vector. Clones were tested for orientation of the amplicons before expansion, using the forward oligo and the oligo for sequencing. Products were ran on a 2% agarose gel for confirmation and further expansion in *E.coli*

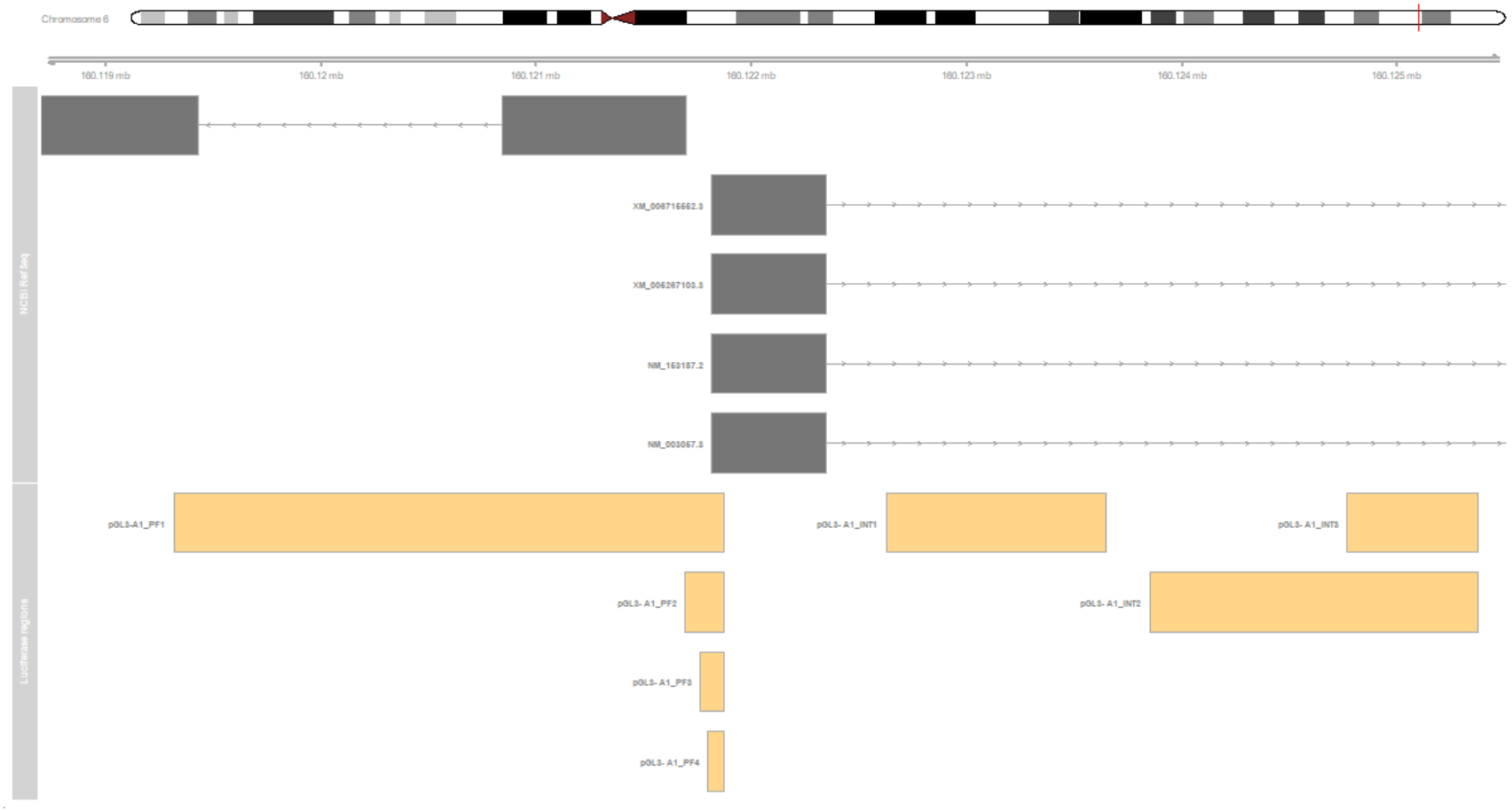


Figure 2-6. Illustration of the promoter regions that were inserted into each of the pGL3 plasmids for the SLC22A1 protein analysis. Exons are shown as boxes, grey lines between each exon indicate the intron.

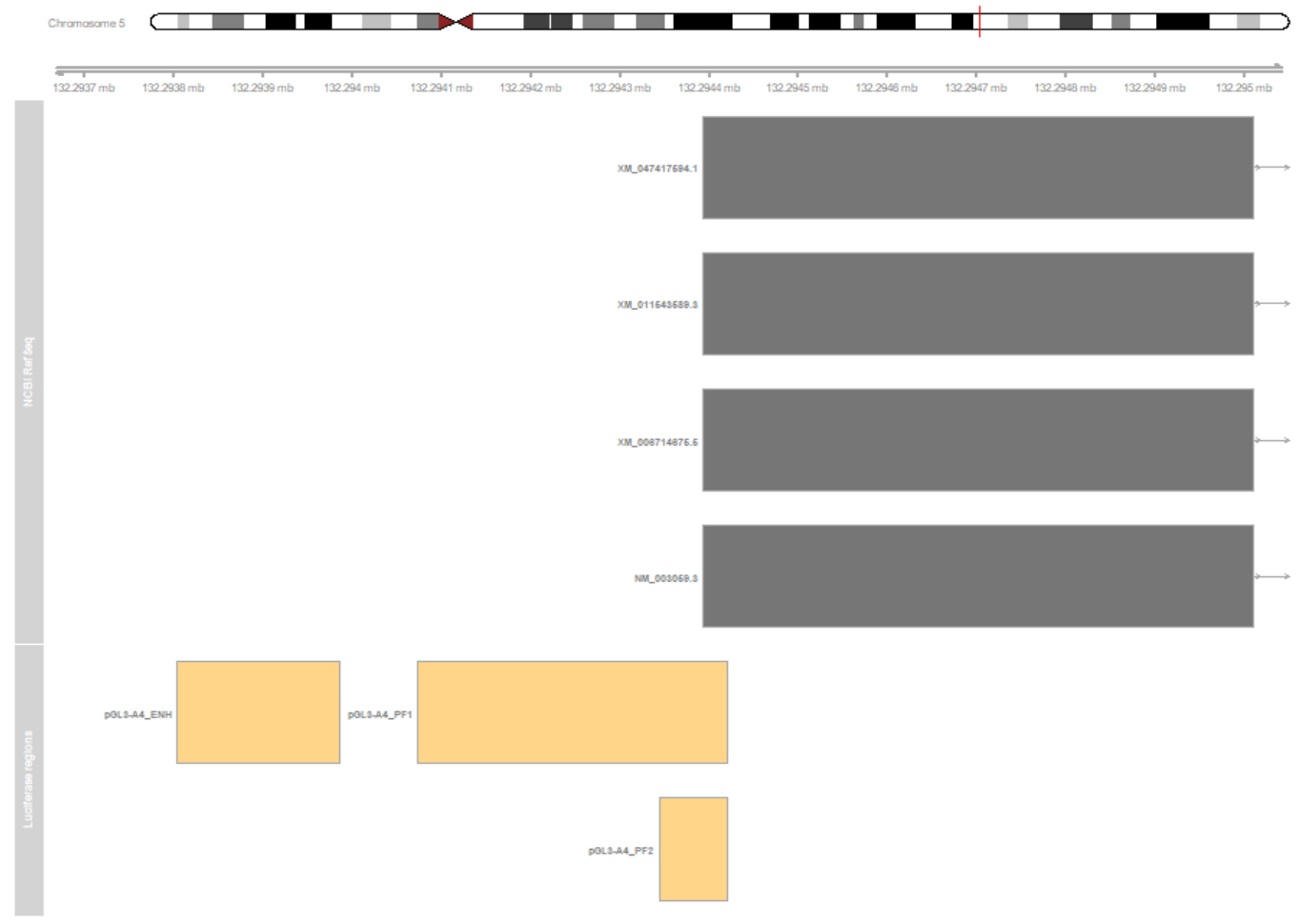


Figure 2-7. Illustration of the promoter regions that were inserted into each of the pGL3 plasmids for the SLC22A4 protein analysis. Exons are shown as boxes, grey lines between each exon indicate the intron.

Figure 2-8. Illustration of the promoter regions that were inserted into each of the pGL3 plasmids for the SLC22A5 protein analysis. Exons are shown as boxes, grey lines between each exon indicate the intron.

Table 2-11. List of primers used for the generation of pGL3 reporters for SLC22A1, SLC22A4 and SLC22A5. Bold letters highlight the restriction sites added on the 5' of the pGL3_A1_INT1 oligos for further sub-cloning into the pGL3 plasmids.

Gene	Construct	Oligos (5' – 3')	Length (bp)
SLC22A1			
	pGL3-A1_PF1	**FWD -** GGAGTTTAACGGCCCATCCCT **REV** –GTCTCCCTCAGAGATCTTTG	2558
	pGL3-A1_PF2	**FWD-** CCACTGACTCGCTCCCGGGC **REV** –GTCTCCCTCAGAGATCTTTG	182
	pGL3-A1_PF3	**FWD -** CCAGCACCATAGGGTAAAAGA **REV** –GTCTCCCTCAGAGATCTTTG	111
	pGL3-A1_PF4	**FWD -** GGTTGCCTTCCAGATGTTTC **REV** –GTCTCCCTCAGAGATCTTTG	78
	pGL3-A1_INT1	**FWD - GGTGCTAGC**CAGGAGCCTCTCTGCCTGCCCC **REV – GGTCTCGAG**CTCAGTCCCAGAGAGCAAACAC	1035
	pGL3-A1_INT2	**FWD -** GTTTAATGACAGCCCAAGGCCG **FWD** –TCTTAGCACCTAGCCTTGCAGTG	1522
	pGL3-A1_INT3	**FWD -** GGAGCCTTGGAGAGGCAGCCATG **REV** –TCTTAGCACCTAGCCTTGCAGTG	604
SLC22A4			
	pGL3-A4_ENH	**FWD -** GCTGTCCTGAGGACGCTGTCCG **REV** –TGTCAGTGGAGTCGGTTCCC	182
	pGL3-A4_PD1	**FWD -** CGGGGCCACGCGGCCCGA **REV -** TCTGAGCTTGATGCCACTGT	347
	pGL3-A4_PD2	**FWD -** GGGCCGCCCTCCCCTTCC **REV -** TCTGAGCTTGATGCCACTGT	76
SLC22A5			
	pGL3-A5_P3	**FWD -** TGTGGAAAGGGCATCTGGAC **REV -** TGTGTTGAAAACGCACCCCTC	218
	pGL3-A5_INT	**FWD -** TGGCAGGATGTTCTGACTTCA **REV -** TGTCATGTGGCCAAGGACTTG	1038

pGL3_A1_INT1 insert was excised from the pJet1.2 vector with NheI-HF® (NEB, cat # R3131S), and XhoI (NEB, cat # R0146S). pGL3-Basic plasmid was linearized using NheI and XhoI for complementarity with the insert.

Fragments were excised from pJet1.4 with XhoI and NcoI-HF® (NEB, cat # R3193S), and purified by agarose gel. PGL3-Basic vector was linearized with

XhoI and NcoI and purified by agarose gel. Respective digested vector and inserts were ligated to generate the different constructs listed on Table 2-10. Visual map of the pGL3-Basic vector can be found in Apendix.

Directed mutagenesis amplification

In order to generate a construct to evaluate the sequence upstream pGL3-A1_PF2, a pair of oligos were designed to delete the sequence from the pGL3-A1_PF1 plasmid.

Site-directed mutagenesis was performed using the pGL3-A1_PF1 as template were the 194 bp were deleted. Oligos were designed to have complementary nucleotides to the plasmid DNA, method was adapted from García-Nafría et al., 2016.

Table 2-12. Oligos for preparation of mutagenesis pGL3-PF1_DEL underlined nucleotides highlight the homologous sequence to the reverse oligo

Construct	**Oligo (5'-3')**
pGL3-PF1 _DEL	Forward- GGCTGAACTTCAATTCTCTTCTCGAGATCTGCGATCTAAG Reverse- AAGAGAATTGAAGTTCAGC

Briefly, PCR reaction was performed using 1 ng template DNAs (pGL3-PF1), according to the following protocol: 30 sec at 95 °C, 18 cycles of 10 sec at 95 °C, 30 sec at 60 °C, 4 min at 72 °C, and a final 5 min extension at 72 °C. Addition of 1 µl FastDigest DpnI enzyme (ThermoFisher, cat # FD1703) was followed by 40 min incubation at 37 °C prior to transformation. Plasmid was sent for sequencing to confirm deletion.

2.2.7. Transfection protocols

Calcium phosphate-mediated transfection of HEK293 cells

HEK293 cells were seeded at a density of 1 x 10^6 cells into a 10-cm diameter cell culture dish. Cells were allowed to adhere for 24 hours after seeding and medium was replaced 2 hours before transfection.

For each dish to be transfected, a mixture was set up according to Table 2-13, a single positive control of GFP expression plasmid (pX461) was used to evaluate transfection.

Table 2-13. Composition of transfection mixtures for calcium phosphate-mediated transfection of HEK293 cells

Component	Volume / Concentration
2 M $CaCl_2$	100 µl
Plasmid DNA	14 µg
H_2O	to a total volume of 800 µl

Transfection mixtures were added separately in a dropwise fashion with gentle agitation to equal volumes of 2X HBS. Solutions were allowed to precipitate at room temperature for 15-20 min before being directly added to cells by dripping them slowly and evenly into the medium. Cells were incubated at 37 ºC with 5% CO_2 for 16 h, and washed three times with PBS. Cells were then replenished with fresh medium. At 48 post-transfection, cells were screened against antibiotic resistance with medium containing 0.5 µg/ml puromycin. HEK293 cells were incubated under puromycin selection for two generations and expanded for genome editing analysis and CRISPR-editing effects.

Lipid-Mediated Transfection of A549 Epithelial Lung Cell Line

A549 cells were transfected with plasmids listed in Table 2-7 using Lipofectamine™ 3000 Transfection Reagent (ThermoFisher, cat # L3000008). Prior to transfection, cells were seeded at a density of 5 x 10^4 in a 24-well plate. On the next day, cells were transfected using Invitrogen Lipofectamine 3000 Transfection Reagent, mixture was prepared according to Table 2-14.

Table 2-14. Composition of transfection mixture for Lipid-Mediated Transfection of A549 cells. Volumes are listed for a single well of a 24-well plate.

Tube	Component	Volume/Amount
1	Opti-MEM médium	25 µl
	DNA	1 µg
	P3000 reagent	1 µL
2	Opti-MEM medium	25 µl
	Lipofectamine 3000 reagent	1.5 µl

Tube 1 solution was added to tube 2 and mixture was incubated for 15 min at room temperature. DNA complexes were added to wells and plate was gently swirled to ensure homogeneous distribution. Cells were incubated in a humidified 5% CO_2 environment at 37°C.

A549 cells transfected with the nickase plasmids (pX461) were sorted according to protocol in Section 2.2.9 and further expanded for evaluation of genome editing.

Transfection of Calu-3 by Nucleofection

Calu-3 cells were transfected using the 4D-Nucleofector™ using the SE Cell Line 4D-Nucleofector™ X Kit (Lonza, cat #V4XC-1024), with program EO-120, according to manufacturer's protocol. Calu-3 cells were rinsed with PBS

and harvested by trypsinization. 1 x 10^6 cells were aliquoted, spun down and resuspended in 100 µl of Nucleofection Solution supplemented with 5 µg of the respective vector.

After nucleofection, cells were carefully resuspended in 400 µl of pre-warmed medium and transferred into a 12-well plate filled with 500 µl of pre-warmed medium. Cells were incubated overnight in a humidified incubator at 37°C/ 5% CO_2 for 48h before antibiotic selection.

Transfection of RNP complexes by Nucleofection

Cells were transfected using the strip format for the 4D-Nucleofector™ X Unit. The SF Cell Line 4D-Nucleofector™ X Kit (Lonza, cat # V4XC-2024), was used for HEK293 and A549, whereas the kit SE Cell Line 4D-Nucleofector™ X Kit (Lonza) was used for transfection of Calu-3 cells. Method was adapted from Xu et al., 2021 and carried-out in parallel, as leaving cells in Nucleofector™ Solution for extended periods of time might reduce transfection efficiency.

sgRNAs were designed according to Section 2.2.3. Guides were diluted to a working concentration of 100 µM in 10 mM Tris-HCl (pH ~7.4 - 7.5). 2.5 µl of Cas9 (40 µM) (Cas9 Nuclease Protein NLS (Horizon™, cat #CAS12206).) was added to 1.5 µl of diluted sgRNA, separately. Complexes were incubated for 10 min at room temperature.

Cells were seeded at a high density, two days before transfection. On the day of transfection, cells were trypsinised and 2 x 10^5 cells were spun down and resuspended in 20 µl of supplemented nucleofector solution. Cells were added

into the tube containing the Cas9/gRNA complexes and gently mixed by pipetting. Cell mixture was transferred into the Nucleocuvette™ Strips and electroporated according to manufacturer's protocol. After run completion, cells were resuspended with 175 µl of pre-warmed medium and plated into a 24-well plate. For analysis of genome editing, genomic DNA was harvested 48 h after transfection, using GenElute™ Mammalian Genomic DNA Miniprep kit (Merck, cat #G1N350).

Transfection of luciferase plasmids

pGL3 plasmids were co-transfected with *Renilla* luciferase (Rluc) plasmid, which functions as a control for transfection efficiency. Luciferase reporters is measured using the Dual-Luciferase Reporter (DLR) Assay System (Promega, E19100).

2.5 x 10^4 HEK293 and A549 cells were seeded into a 24-well plate, separately. Whilst Calu-3 cells were seeded at a density of 0.5 x 10^5 per well. Plates were incubated overnight at 37 °C/5% CO_2. On the next day, cells were transfected according to Table 2 14, volumes were calculated per well.

Table 2-15. Lipofectamine transfection. Reaction volumes are listed per well.

Tube	Component	Volume / Concentration
1	Opti-MEM Medium	25 uL
	Renilla Vector	100 ng
	pGL3 Vector	200 ng
	P3000™ Reagent	1 uL
2	Opti-MEM Medium	25 uL
	Lipofectamine™ 3000 Reagent	1.5 uL

Mixture from tube 1 was added to tube 2 and was then incubated for 15 minutes at room temperature using Gibco Opti-MEM I Reduced Serum Medium No

Phenol Red (ThermoFisher, cat # 11520386). DNA-lipid complexes were added to cells in a dropwise manner, wells were gently swirled to ensure even distribution over the entire plate. Cells were incubated overnight at 37°C/ 5% CO_2. Experiment was performed in three biological and technical replicates.

Measuring of Luciferase Activity

24 hours after transfection, medium was aspirated and cells were then carefully washed once with PBS. 50 µl of freshly-prepared 1X Passive Lysis Buffer was added to each well. Plate was placed on an orbital shaker with gentle movement for at least 20 min to ensure complete lysis. Complete lysis was confirmed under inverted microscope and 10 µl of lysate was placed on a white 96-well plate with flat clear bottom.

Activity of firefly luciferase is measured by adding Luciferase Assay Reagent (LARIII) to generate a luminescent signal. Afterwards, the reaction is quenched, and the Renilla luciferase reaction is simultaneously initiated by adding the Stop & Glo® reagent to the same sample.

Dual-luciferase assays were performed using a Glomax 96 microplate luminometer equipped with dual robotic auto-injectors (Promega E6521). Both injectors were washed before run, injector 1 was primed with LARII while Stop & Glo® reagent was primed in injector 2. Measurements of luminescence were done by injecting 25 µl of LARII followed by 25 µl of Stop & Glo, with a 0.4 sec delay between injections and measurement and a 10s for integration time.

For data analysis, the *luciferase* luminescence reading was divided by the *Renilla* luminescence reading. This provides a signal value for each well that is normalized by transfection efficiency.

2.2.8. Gene Expression protocols

To detect variation in mRNA expression of SLC22A1, SLC22A4 and SLC22A5, quantitative real-time PCR (qPCR) was performed using primers designed to span introns of the target gene. Primers were designed to according to their target gene and validated by a serial dilution.

RNA extraction

Cells were rinsed with ice cold PBS, trypsinised and pelleted using standard methods. mRNA isolation was performed using), GenElute™ Mammalian Total RNA Miniprep Kit (Merck, cat #RTN70), according to the manufacturer's protocol.

Briefly, samples were lysed and homogenized in guanidine thiocyanate and 2-mercaptoethanol to release RNA and inactivate RNases. Lysates were spun through a filtration column to remove cellular debris and shear DNA. The filtrate was then applied to silica column to bind total RNA, followed by washing and elution with RNAse-free water.

cDNA synthesis

SuperScript™ III Reverse Transcriptase (ThermoFisher, cat# 18080093), was used to synthesize cDNA from total RNA using random primers. RNA concentration was measured with Nanodrop™ and 200 ng of total RNA per

sample was used to perform cDNA synthesis. Conditions on thermocycler were as follow, a 25°C for 5 minutes, 50 °C for 60 minutes followed by 70°C for 15 minutes. qPCR was performed on a Rotor-Gene Q (QIAGEN).

qPCR validation and assay

All primers were designed to span intron junction to avoid amplification of genomic DNA. Reactions were prepared using GoTaq® qPCR Master Mix (Promega, A6102). Oligos were designed to target different transcripts for every specific gene. The positions and directions of the designed oligos (reverse and forward) for qPCR experiments are depicted in Table 2-9. Black arrows indicate their location relative to the corresponding exon structure.

Table 2-16. qPCR oligos for gene expression analysis.

Gene	Sequence (5' – 3')	Product size (bp)
SLC22A1	FWD – CCTGTTTGAATGCGGGCTTCTT REV - GAGACACAGCTTACGGCCAAAC	80
SLC22A4	FWD – CCACCTCCCTGTTCTTCGTAGG REV – TACAGCCATGGTTGCGAAGAGA	104
SLC22A5	FWD – CTCCTTCATTTCAGGGCAGC REV- TGCCAAGAATTTCTGTCCCCA	203
GAPDH	FWD - AGGTGAAGGTCGGAGTCAAC REV - GATGACAAGCTTCCCGTTCT	200

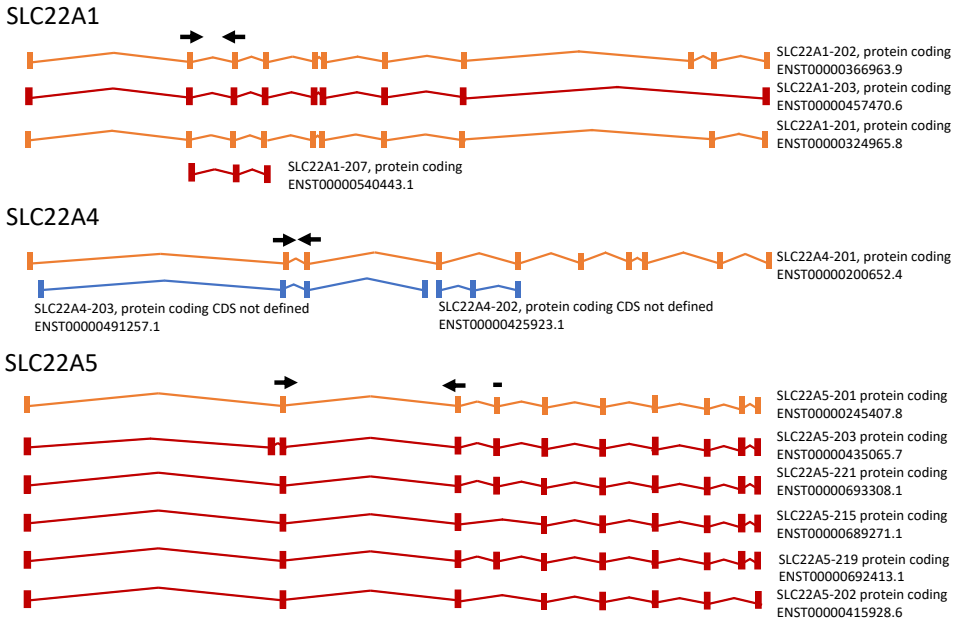


Figure 2-9. location of oligos (reverse and forward) designed for quantitative PCR (qPCR) experiments. Black arrows indicate the position and direction of the oligos relative to the corresponding exon structure. Exon information was retrieved from Ensembl transcripts (accessed March 25, 2024). Transcript colouring scheme: blue: processed transcript, orange: merged Ensembl/Havana transcript, red: Ensembl protein coding transcript

A standard curve was generated using 2-fold dilutions of cDNA as template for qPCR reactions (Figure 2-10). Primer specificity was determined by melting curve analysis. Melting curve analysis is an assessment of the dissociation characteristics of double-stranded DNA (product from the qPCR reaction) during heating. Primer specificity was confirmed from the separation of amplification products on a 2% agarose gel and melting curve analysis, Figure 2-11.

The thermal profile for all qPCR reactions was 10 minutes at 95°C as initial denaturalization, followed for 40 cycles of 10 sec at 95°C, 20 sec at 60°C and 72°C for 20 sec. Cycler was programmed to run a melting curve analysis at the end of the run by gradually increasing the temperature from 60°C to 95°C

(0.05˚C·s^{-1}), as shown in Figure 2-11. All samples were run in triplicates and the mean Ct values for each trial were calculated. Reactions with template free control were included for each set of primers on each run.

The reaction was run on Rotor-Gene Q System and Rotor-Gene Q Series Software version 2.3.5 was used to generate the efficiency of reactions, amplification plots, standard curves and dissociation curves (Figure 2-11).

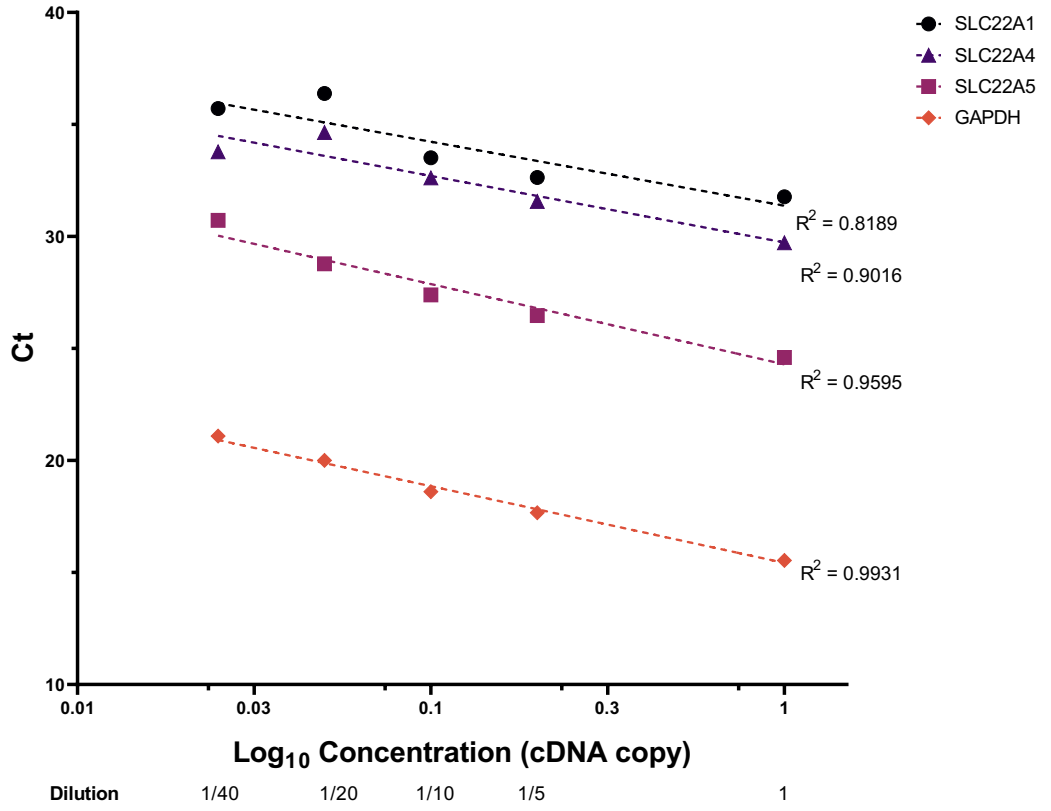


Figure 2-10. A standard curve was generated using a 2-fold dilution of cDNA as template for qPCR reactions. Resulting Ct values are plotted against the Log10 of the cDNA input. R2 values are within the acceptable range.

Data Analysis

Oligo efficiency was obtained by a 2-fold serial dilution of the cDNA. Relative gene expression was normalised with a stably expressed reference gene as internal control, known as the comparative Ct or the ΔΔCt method:

Equation 2-2. qPCR data normalisation, ΔΔCt method. GOI stands for gene of interest, RF stands for reference gene.

$$\Delta\Delta C_t = (\mathrm{Ct}_{\mathrm{GOI}}\text{ , query-Ct}_{\mathrm{GOI}}\text{, control}) - (\mathrm{Ct}_{\mathrm{RF}}\text{, query-Ct}_{\mathrm{RF}}\text{, control})$$

$$\text{Relative quantification} = 2^{-\Delta\Delta\mathrm{Ct}}$$

The $\Delta\Delta C_t$ method normalises the C_t values of a target gene to the internal reference gene before comparisons are made between samples. Quantification of the target mRNA was normalised using GAPDH as the reference mRNA. First, the difference between C_t (ΔC_t values) of the target gene and the mean of GAPDH is calculated for each sample, and then the difference in the ΔC_t ($\Delta\Delta C_t$) is calculated between two samples (e.g., control and treatment, or pre and post treatment). The fold-change in expression of the two samples is calculated as $2^{-\Delta\Delta Ct}$. The gene expression of HEK293-edited cells was normalised against expression in WT cells. Oligos are listed in Table 2-16.

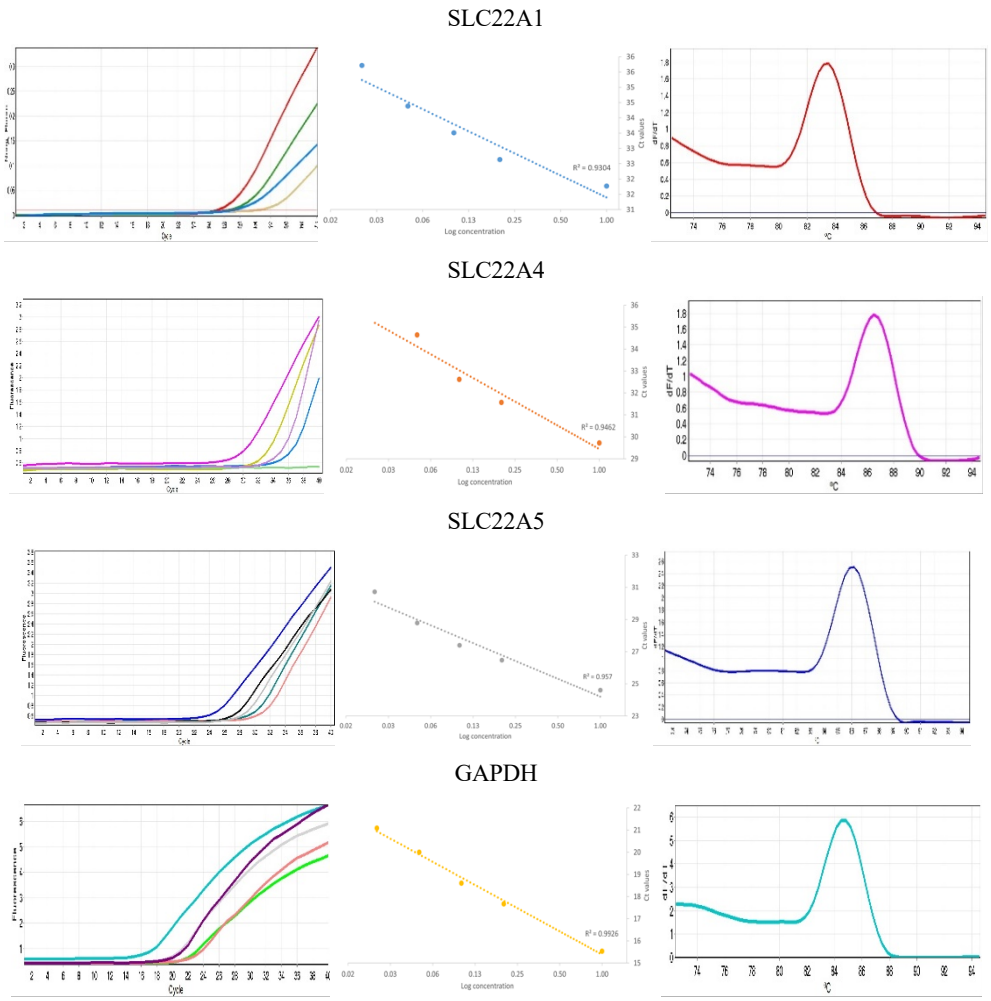


Figure 2-11. Amplification plots, standard curves, and melting curves for 4 genes. Expression was recorded using SYBR green based qPCR. The primers efficiency was obtained by achieving serial dilution of the primers. Amplification plots are presented on the left panel, the middle one represents the standard curves and the right panel shows the melting curves, per gene, respectively.

2.2.9. Flow cytometry / Cell-sorting

Flow cytometry is a technique that measures characteristics of a population of cells as they flow past single or multiple lasers. It is predominantly used to measure the expression of cell surface or intracellular molecules using fluorescent labelled antibodies or ligands for their detection.

To separate a population from a heterogenous sample (cell sorting) a vibration is introduced generating waves that break the fluid stream into regular droplets allowing to select individual cells and divert them into a collection tube.

The Astrios EQ cell sorter (Beckman Coulter) was used to isolate individual GFP-transduced cells from untransduced cells, sorting was conducted in the School of Medicine at the University of Nottingham. Cell debris was excluded in an FSC-A versus SSC-A dot plot, duplets were discriminated in an FSC-W vs FSC-A dot plot and the remaining cell population was analysed depending on the application. Flow cytometry data was analysed using Kaluza Analysis Software v 2.1.

2.2.10.Functional Assays

The impact of the CRISPR-editing of EMX1, SLC22A1, SLC22A4 and SLC22A5 in cell proliferation, migration and cell-adhesion was evaluated in HEK293-edited cells.

Cell Proliferation assay

Proliferation was assessed in a 5-ethynyl-2'deoxyuridine (EdU) incorporation assay, by using a Click-iT EdU Alexa Fluor 488 cell proliferation assay kit (Thermofisher). At least 200 nuclei were analysed.

Cells were seeded at 0.5 x 10^6 cells per well in a µ-Slide 8 Well (ibidi®) and allowed to recover overnight. Half of the media was replaced with 20 µM EdU labelling solution and the cells were incubated for two hours.

After incubation, cells were fixated with 3.7 % paraformaldehyde in PBS and kept for 15 minutes at room temperature. Cells were washed twice with 3% BSA in PBS and permeabilised with 0.5% Triton®X-100 in PBS. Click-iT reaction cocktail was prepared fresh, according to the manufacturer and immediately added to the chamber.

Chamber was placed on a rock plate to for evenly distribution. Reaction was incubated for 30 minutes, protected from light. Cells were washed twice with 3% BSA in PBS and counterstained with 1 µg/ml DAPI in PBS for 15 min at room temperature.

The percentage of cells in s-phase was determined by counting EdU stained cells as a percentage of DAPI stained cells (n=200) using the EVOS M5000 Imaging System mounted with lightcubes for DAPI (357/447 nm) and GFP (470/525 nm). Experiment was performed in triplicates. Images were analysed with ImageJ.

Wound Healing assay

Wound healing assay was adopted to detect cell migration. Cells were seeded in a 2 well culture-insert (ibidi GmbH) at a density of 3500 cells/well. On the next day, culture insert was removed and cells were rinsed with PBS and covered with fresh growth medium. Migration was documented for 72 h. For every time point, three pictures were taken using the Celestron HD Microscope Imager mounted on an inverted microscope.

The rate of migration was determined by measuring the entire wound area immediately after wounding and at the indicated time points, and by normalizing to the control condition. Migration analysis was performed with ImageJ.

Cell Adhesion

Cell adhesion was indirectly asessed using a colorimetric detection format based on toluidine blue staining. Cells were harvested by trypsinization, according to the General Tissue Culture protocols (Section 2.2.4). Cells were seeded at a density of 0.5 x 10^6 cells per well in 200 µl of supplemented media and incubated at 37°C for 16 hours. Prior to fixation, cells were washed twice with PBS. A 3.8% paraformaldehyde solution was used to fix the cells. They were further stained with a 1% toluidine blue/3% PFA and washed twice with PBS. Cells were lysed for 5 min with 2% SDS (sodium dodecyl sulfate), plates were placed on an orbital shaker to allow for uniform colour. Absorbance was measured at 590 nm in a BioTek Synergy Multi-mode Microplate Reader (Agilent). Experiment was performed in three biological and technical replicates.

2.2.11.Statistical Analysis

All the statistical analysis were performed with GraphPad Prism 9.0 (GraphPad Software, Inc., San Diego, CA), except for Differential Expression and Mehylation data, which were analysed using R software (version 4.2.2). $p < 0.05$ were considered significantly differentially expressed. Values are expressed as the arithmetic mean ± standard error of the mean (SEM).

CHAPTER 3. EXPRESSION PROFILES OF SLC22A1, SLC22A4 AND SLC22A5 TRANSPORTERS IN COPD, ASTHMA AND LUNG CANCER

CHAPTER 3. EXPRESSION PROFILES OF SLC22A1, SLC22A4 AND SLC22A5 TRANSPORTERS IN COPD, ASTHMA AND LUNG CANCER

The solute carrier (SLC) superfamily encompasses a large variety of membrane-bound transporters that facilitate the transport of a wide array of substrates across biological membranes. They have important roles in physiological processes ranging from nutrient uptake, ion transport and absorption of drugs.

Pharmacologically, SLC22 family members play a major role as determinants of the absorption and disposition of many prescription drugs (Lozano et al., 2013; Motohashi & Inui, 2013). SLC22 transporters have been associated with common human diseases, drug response and various phenotypic traits. Moreover, mutations in some of these transporters, (SLC22A5) lead to rare monogenic disorders .

SLC22A1 has been found to be significantly overexpressed in various cell lines and samples from lymphoma patients, making this transporter a prime candidate for transporter-specific tumour therapy with antineoplastic substances (irinotecan and paclitaxel) (S. Gupta et al., 2012). This transporter is then thought to contribute to susceptibility to antineoplastic drugs. Another study suggested that sensitivity to oxaliplatin was due to the role of SLC22A1 transporting oxaliplatin into cancer cells. Platinum-based therapies are used in the treatment of a variety of tumours, including testicular cancer, ovarian cancer, small cell lung cancer, and head and neck cancers. Even when cisplatin is often the drug of choice, oxaliplatin may be a better choice when OCT1 is expressed

in the tumours (Zhang et al., 2006). In addition, expression of SLC22A1 in chronic myeloid leukaemia cell lines and primary cells has suggested the influx of imatinib by SLC22A1, thus, differential expression of this transporter may be a critical determinant of drug resistance (J. Thomas et al., 2004).

Given the importance of the SLC transporters in many essential physiological processes, its dysregulation in human disease suggests a potential therapeutic targets. Accumulating evidence suggest that dysregulation of SLC transporters may play a pivotal role in the pathogenesis of human diseases. Upregulation of specific transporters might be due to the increased demand for nutrients in tumours which is met by increased availability of nutrients by these transporters, translating to a gain resistance to chemotherapeutic agents (Ganapathy et al., 2009).

A recent study by Whisenant & Nigam, 2022, revealed an association between overall survival and expression of multiple SLC22 transporter genes between two renal cell carcinoma (clear renal carcinoma and papillary cell renal cell carcinoma) and normal tissue. SLC22A6, SLC22A7, SLC22A8, SLC22A12, SLC22A13 were amongst the various differentially expressed (DE) SLC22 transporter genes. These are, interestingly, uric acid transporters, and altered uric acid levels have been associated with kidney cancer. Another study has linked SLC22A1 expression with advanced cholangiocellular carcinoma stages, where its downregulation was significantly associated with larger tumour diameters ($p=0.02$) and worse patient survival ($p<0.05$) (Lautem et al., 2013).

One comprehensive study on expression of SLC transporters in cancer tissue and clinical outcome (Edemir, 2020) identified that expression of SLC22A5, as

a favourable prognosis in endometrial, pancreatic and renal cancer. Whereas SLC22A1 expression was found to be favourable only in liver cancer. Although higher expression of OCT1 in proximal tubule cells might represent less differentiation toward a tumour cell.

Differential expression of SLC transporters in breast cancer was recently investigated (Sutherland et al., 2021), SLC22A1 and SLC22A4 showed higher expression in MDA-MB-231 ($p<0.05$) whereas SLC22A5 showed higher expression in MCF-7 ($p=0.001$). However, SLC transporters showed a variable expression in breast cancer samples. SLC22A5 was highly expressed across all samples, however, SLC22A1 was expressed in all tumours at different levels, whereas SLC22A4 showed no to normal expression between samples. These transporters could have a significance in the treatment or pathogenesis of breast cancer.

Asthma, Chronic Obstructive Pulmonary Disease (COPD), Lung Adenocarcinoma (LUAD) and Lung Squamous Cell Carcinoma (LUSC) are all serious lung illnesses that affect millions of people worldwide.

Asthma is a chronic respiratory disease characterized by inflammation and narrowing of the airways, leading to difficulty breathing, wheezing, chest tightness, and coughing. According to the World Health Organization (WHO), asthma affects an estimated 235 million people worldwide (*(*NEW) 2019 GINA Report*, n.d.). COPD is a chronic respiratory illness characterized by airflow obstruction, bronchitis, and emphysema. According to the WHO, COPD is the fourth leading cause of death globally and it is estimated that 384 million people have moderate to very severe COPD (*Chronic Obstructive Pulmonary Disease*

(COPD), n.d.). LUAD and LUSC are the two main subtypes of lung cancer. LUAD is the most common form of lung cancer accounting for around 40% of all lung cancer cases. LUSC represents around 25% of all lung cancer cases. Lung cancer is the leading cause of cancer-related deaths worldwide, with 1.8 million deaths annually (*Lung Cancer Guide*, n.d.).

Given the high prevalence and debilitating nature of these lung illnesses, it is important to understand the underlying mechanisms of the disease to develop effective interventions. One area of research that is becoming increasingly important in this regard is the expression and function of transporter proteins in drug absorption. Protein transporters play a crucial role in the transport of drugs across cell membranes and are a major determinant of drug efficacy and toxicity. Understanding the expression and function of protein transporters in these lung illnesses can lead to the development of targeted therapeutics that can improve drug absorption and efficacy.

Organic cation transmembrane transporters belonging to the SLC22 family have been increasingly recognised as involved in drug disposition in the respiratory tract. SLC family genes have been reported to be associated with cellular proliferation and survival in lung adenocarcinoma (LUAD) and they may be useful diagnostic and prognosis biomarkers (Guo et al., 2020; K. Hu et al., 2020; H. Zhou et al., 2021).

Public databases such as the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) are an excellent public source of whole transcriptomic profiles of multiple cancers. The Gene Expression Omnibus (GEO) database is a public data repository launched in 2000 to support public use of genomic

resources provided by the scientific communities. More than 24,000 platforms and 5, 450, 000 samples are currently stored. Datasets have been widely used for the study of dysregulation genes and potential therapeutic targets in many illness.

Approximately 40% of all pharmaceutical drugs are identified as organic cations (Neuhoff et al., 2003) The organic cation transporter 1 (OCT1) is a member of the solute carrier 22 (SLC22) family of transporters, and plays a vital role in the uptake and elimination of various endogenous compounds and xenobiotics, particularly in the liver and kidney (Koepsell, 2013). However, recent studies have also highlighted the importance of OCT1 in the lung epithelia, where it plays a crucial role in the pharmacokinetics of various clinically used drugs (Bosquillon, 2010) and as such, they have the potential to act as substrates and/or inhibitors for organic cation transporters (OCTS). Strategies based on targeting a specific transporter by designing prodrugs may be used to improve the amount of transportation of some drugs. Aerosol medications have become more and more important in treating bronchial asthma, although with unwanted systematic effects and poor local anti-inflammatory effects. In order to improve the uptake of prednisolone across human bronchial epithelial, (Bourgon et al., 2010) synthesized an L-carnitine ester prodrug form which was found to be actively translocated by the carnitine transporter OCTN2 (SLC22A5) in BEAS-2B. Study also indicated that the prodrug caused sustained reduction of the *in vitro* release of IL-6.

In this chapter, datasets available from the GEO and TCGA were used to evaluate the expression of Solute Carrier (SLC) family 22 genes in asthma,

chronic obstructive pulmonary disease (COPD), lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC).

3.1. Results

Gene expression profiles GSE67472 and GSE76925 from the GEO database were analysed to screen for Differential Expression (DE) of SLC22A1, SLC2A4 and SLC22A5 for asthma and chronic obstructive pulmonary disease (COPD), respectively. TCGA-LUSC and LUAD datasets from the Cancer Genome Atlas (TCGA) were analysed for DE of these transporters in lung adenocarcinoma and lung squamous cell adenocarcinoma, respectively.

3.1.1. Data pre-processing

Microarray data analysis requires normalization to account for technical variations that can arise during different stages of the experiment, such as sample preparation, labelling, and scanning. These variations can introduce artifacts into the data, potentially masking or skewing true biological differences between samples. In this study, the normalization process aimed to adjust the measured expression values for each gene across all samples, removing technical variations while preserving biological signal. This allows for reliable comparisons of gene expression patterns between samples and facilitates the identification of biologically relevant changes. By effectively removing technical variations through normalization, we ensured that subsequent analyses were based on data that accurately.

The neqc() function from the limma package was used to normalise data from Illumina HumanHT. This method leverages negative control probes for

background correction and employs quantile normalization to ensure consistent intensity distributions across samples.

Data from Affymetrix Human Genome was normaliseed using the affy package and the Robust Multi-array Average (RMA) method, which incorporates mismatch probes for background correction and utilizes quantile normalization to address intensity-dependent biases inherent in Affymetrix data, leading to more accurate expression level estimates.

As depicted in Figure 3-1, the boxplots reveal a significant improvement in data consistency after normalization. The spread of expression levels is visibly reduced, indicating the removal of technical variations. Additionally, the median values of gene expression per sample are well-represented by the horizontal lines, suggesting minimal alteration of the biological signal. These observations demonstrate the effectiveness of the chosen normalization methods in correcting for systematic biases and ensuring that subsequent analyses are based on reliable and comparable expression data.

The initial step of the gene expression analysis was to filter out low-transcription genes in any of the samples (Bourgon et al., 2010). This step was necessary to ensure that the downstream analyses were performed on genes that had a meaningful amount of expression data. In order to evaluate the effectiveness of this filtering step, a comparison of the gene expression distribution before and after filtering was performed. Datasets were analysed for differentially expressed genes (DEGs) and the screening criteria was $P<.05$ and $FC>2.0$.

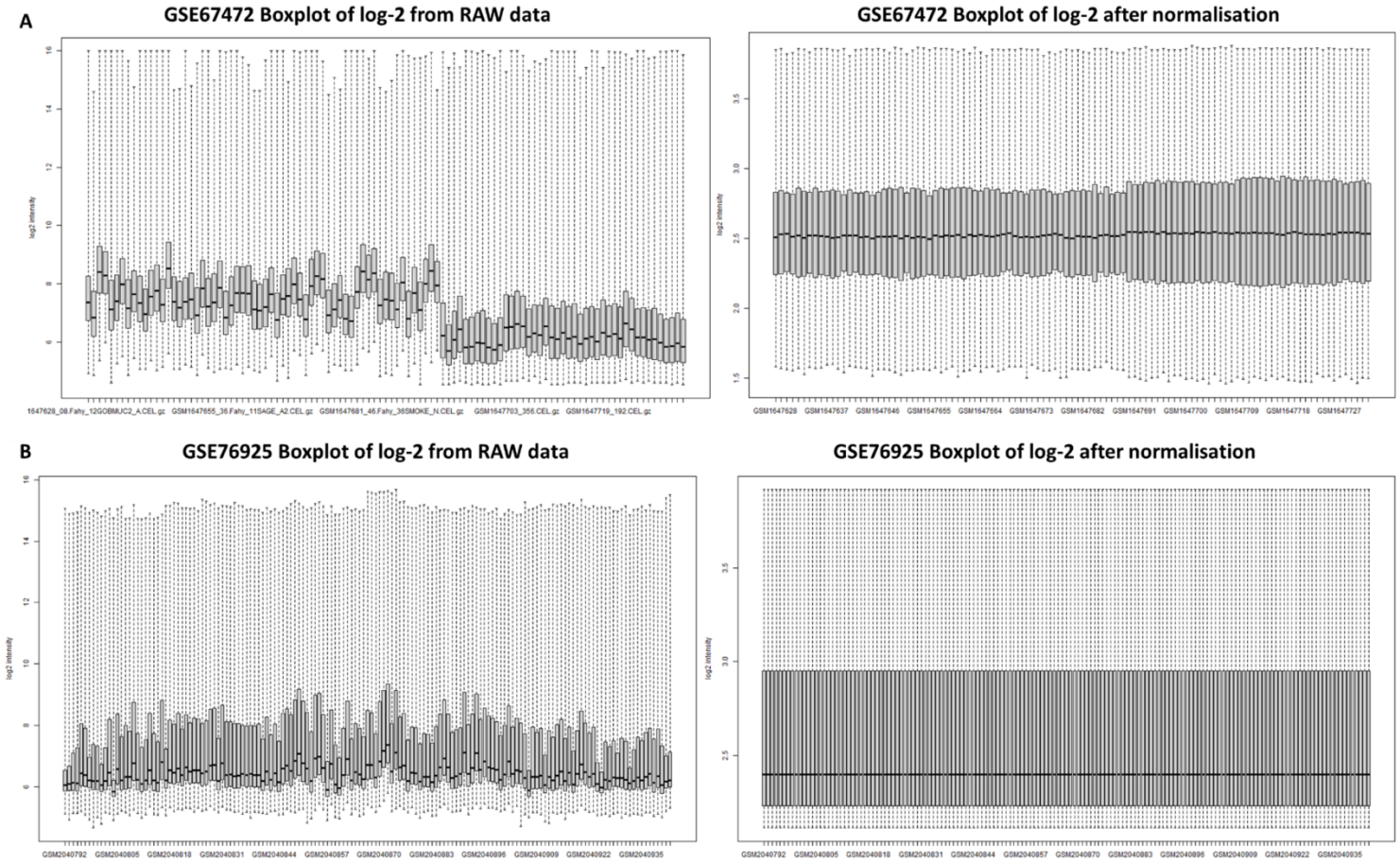


Figure 3-1. Box plots of the gene expression data before and after normalisation. Horizontal axis represents the samples and vertical axis represent the gene expression values. Black lines represent the median value of gene expression per sample (A) GSE67472, Asthma dataset (B) GSE76925, COPD dataset

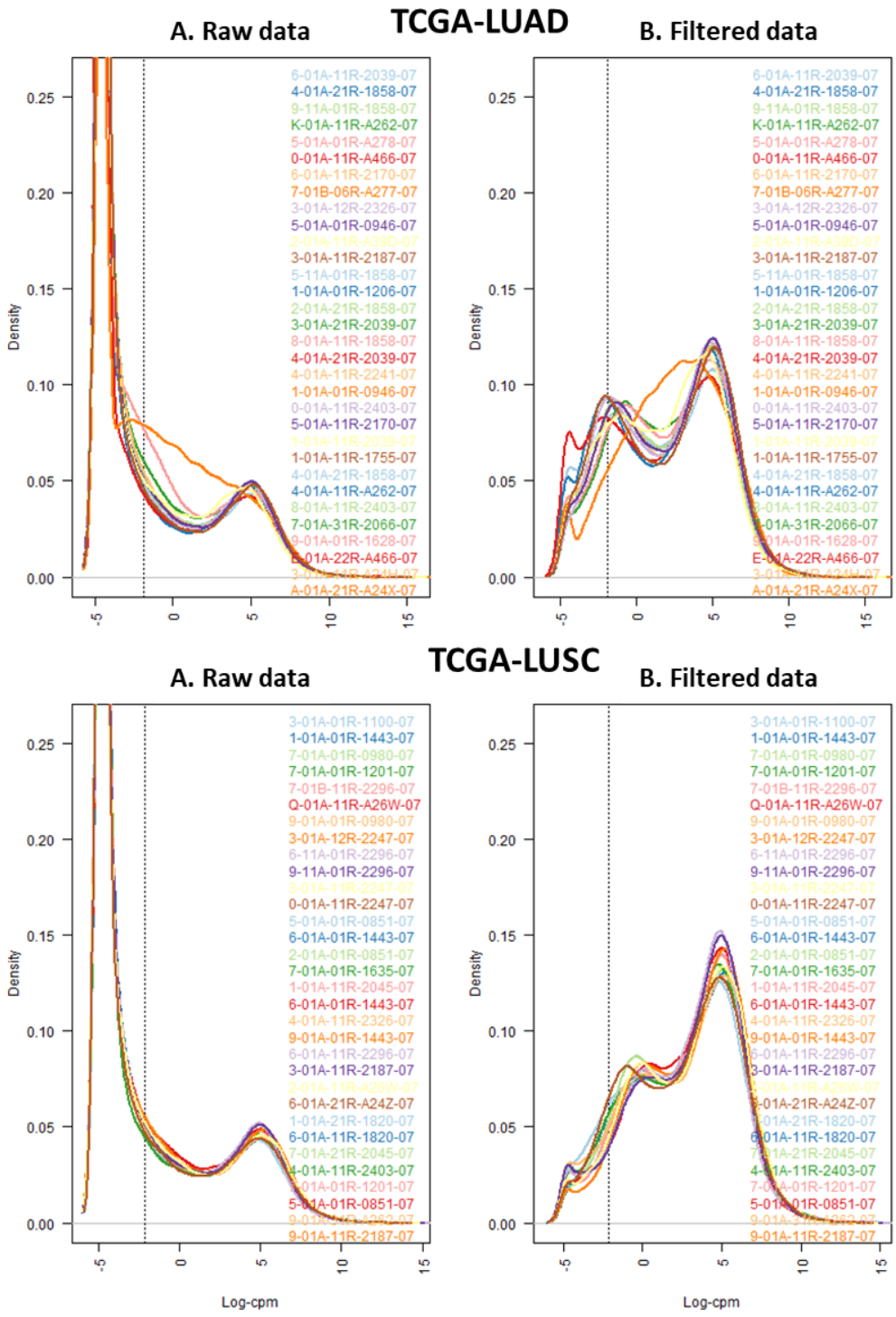


*Figure 3-2. Density of log-CPM values for **(A)** raw pre-filtered data and **(B)** post-filtered data are shown for each sample. Dotted vertical lines mark the log-CPM threshold used in the filtering step*

To visualize the distribution of normalized counts and determine an appropriate filtering threshold, a histogram was generated. Figure 3-2 shows the filtering of

TCGA datasets, demonstrate a clear shift in the distribution of expression levels after filtering, with a reduction in the number of genes with low expression levels and an increase in the number of genes with high expression levels. This indicates that the filtering step effectively removed genes that had a low level of expression and therefore did not have a meaningful contribution to the downstream analyses. Log-CPM values after filtering showed a nearly unimodal distribution for each sample. The filtering step, therefore, ensured genes to be expressed with a cpm >1.0 in at least 30 samples and it was critical for the proper interpretation of the results.

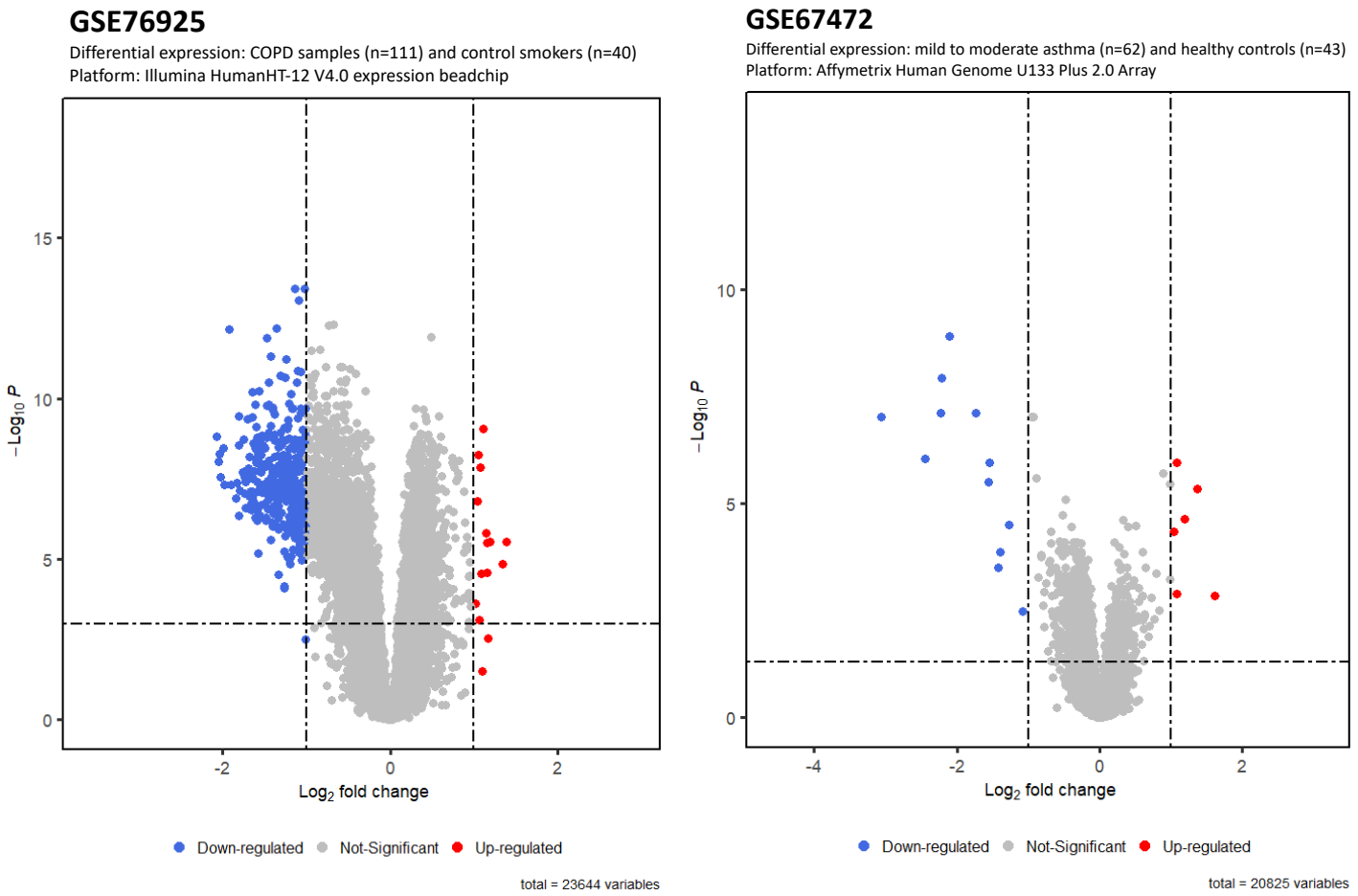


Figure 3-3. Volcano plots showing the differentially expressed genes in the two GEO datasets, GSE76925 (111 COPD samples and 40 control smokers, microarray) GSE67472 (63 asthma and 43 healthy controls, microarray). Down-regulated genes are coloured in blue, up-regulated genes are coloured red and not significant genes are coloured in grey.

3.1.2. Expression of SLC22A genes in COPD and asthma

Differential gene expression analysis was performed on each dataset to identify genes with statistically significant changes in expression between the disease and control groups. A volcano plot was generated for each dataset to visualize the log2 fold change (FC) versus the adjusted p-value (-log10(adjusted p-value)) for all genes (Figure 3). Genes with an adjusted p-value < 0.05 and a fold change threshold of ± 1.2 were considered differentially expressed. The COPD dataset (GSE76925) identified 15 upregulated and 355 downregulated genes, while the asthma dataset (GSE67472) revealed 12 downregulated and 6 upregulated genes. A complete list of significantly differentially expressed genes for both COPD and asthma can be found in Table 6-1 and Table 6-2, respectively.

To explore the roles of SLC22A genes expression in COPD, asthma, expression datasets from GEO database were analysed using R software. Table 3-1 summarises the differential expression of SLC22A1, SLC22A4 and SLC22A5 across different datasets.

Analysis of DEGs in GSE76925 showed in Figure 3-4 that expression of SLC22A15, SLC22A17, SLC22A18, SLC22A18AS, SLC22A5 and SLC22A7 was significantly higher in COPD tissue than in control lung tissue, however, expression of SLC22A1, SLC22A10, SLC22A12, SLC22A14, SLC22A16, SLC22A2, SLC22A20, SLC22A23, SLC22A3, SLC22A4 and SLC22A6 showed no significant difference.

Table 3-1. Differential expression of SLC22A1, SLC22A4 and SLC22A5 from the different datasets analysed. Log2FC<0: down-regulated, Log2FC>0: up-regulated. adj.P.Val: Benjamini-Hochberg false discovery rate adjusted p-value

Dataset	SLC22A1	SLC22A4	SLC22A5
GSE67472	0.163629284 (1.78E-03)	-1.026833773 (7.04E-10)	0.363580728 (3.76E-06)
GSE76925	0.040953475 (3.914475e-01)	-0.205366414 (4.366855e-01)	0.275712551 (5.757112e-04)
TCGA-LUSC	0.26632349 (1.76E-01)	-1.625400568 (1.09E-33)	-0.375142385 (1.88E-07)
TCGA-LUAD	-0.265514206 (1.15E-01)	-1.155433435 (4.94E-14)	-0.127042719 (1.36E-01)

A closer examination of SLC22A1, SLC22A4, and SLC22A5 expression revealed a statistically significant difference only in SLC22A5, Figure 3-5. Specifically, SLC22A5 expression was significantly higher in COPD samples compared to healthy subjects within the COPD dataset (GSE76925) ($p < 0.001$). In contrast, the expression of these three SLC genes did not show any significant difference in the asthma dataset (GSE67472).

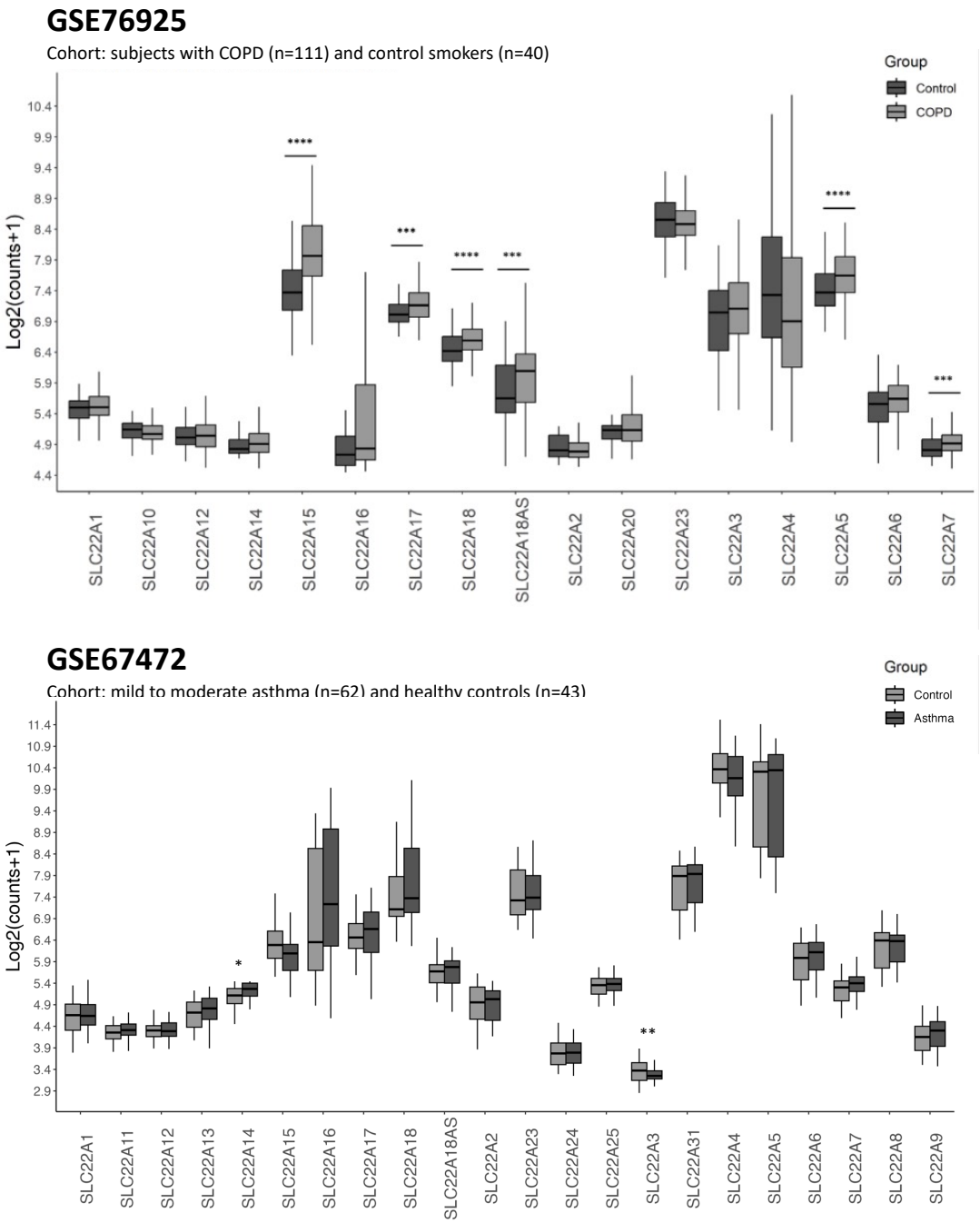


*Figure 3-4. mRNA expression profiles of SLC22A families in GSE76925 (COPD and control smokers) and GSE67472 (asthma and control samples). Control samples are coloured in dark grey. *** p <0.005, **** p<0.01*

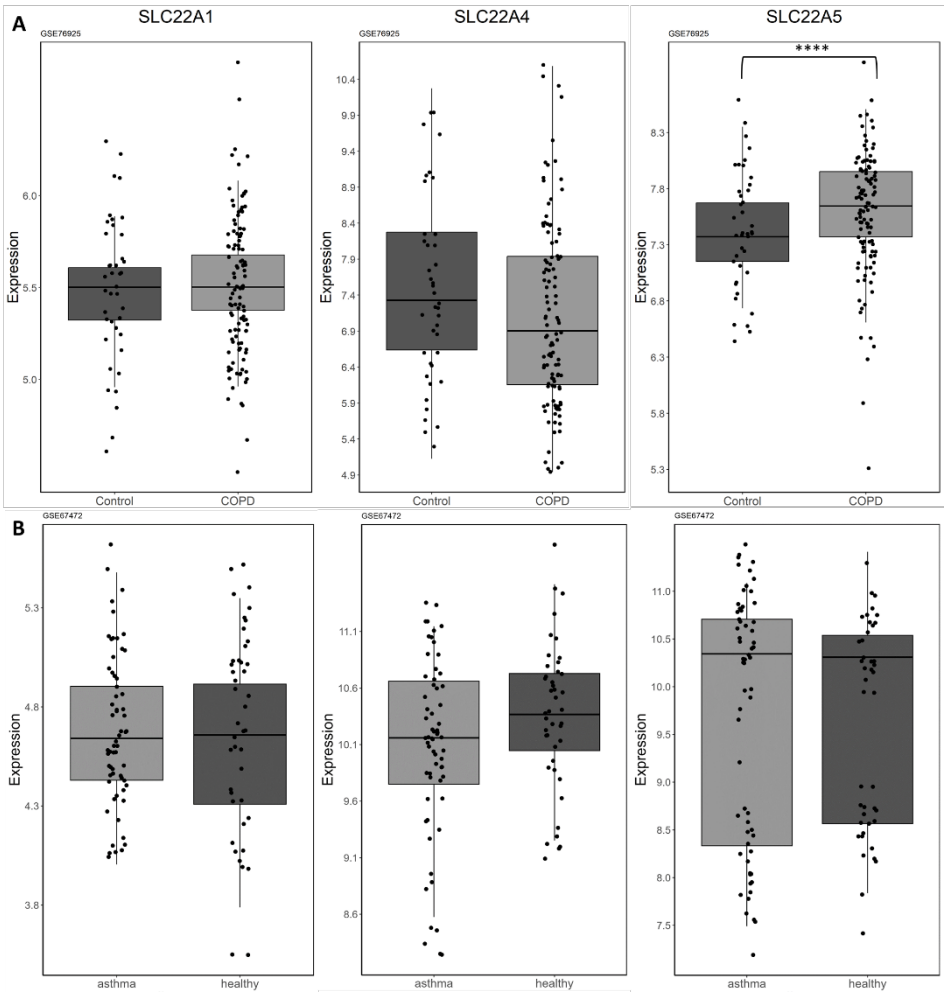


*Figure 3-5. Comparison of the expression levels of SLC22A1, SLC22A4 and SLC22A5 between A) COPD and healthy samples using the GSE76925 dataset B) asthma samples and control smokers using the GSE67472 dataset. **** $p<0.001$*

3.1.3. Expression of SLC22A genes in LUAD and LUSC cohort from TCGA

The mRNA expression profiles of SLC22A genes in lung adenocarcinoma (LUAD) and in lung squamous cell carcinoma (LUSC) were obtained from the TCGA portal and analysed using R software and package limma. Figure 3-6 (A) presents the analysis of expression of SLC22A genes in the LUAD cohort. As can be seen from the profiles, a significant downregulation ($p<0.001$) of

SLC22A15, SLC22A16 ($p<0.05$), SLC22A17 ($p<0.005$), SLC22A3, and SLC22A4 was found in adenocarcinoma tissue, while SLC22A11, SLC22A18, SLC22A18AS, SLC22A20P genes showed a significant ($p<0.001$) up-regulation in expression in lung adenocarcinoma tumour, in comparison to normal tissue.SLC22A1 and SLC22A5 showed a lower but not significant difference in expression profiles in tumour cells; SLC22A23 and SLC22A31 showed a broader range of expression ($p>0.05$).

Analysis of the TCGA LUAD dataset using limma revealed a total of 7,470 downregulated and 7,682 upregulated genes when compared to the normal tissue control group. A comprehensive list of the top 100 differentially expressed genes (up or downregulated) ranked by their adjusted p-value can be found in Appendix section, Table 6-3 and Table 6-4for LUAD and LUSC, respectively. Table 3-1 summarises the differential expression of SLC22A1, SLC22A4 and SLC22A5 in LUAD and LUSC.

Analysis of expression of SLC22A genes in LUSC patients is shown in Figure 3-6 (B). Analysis of differentially expressed genes revealed 7,579 to be upregulated and 7,759 downregulated genes in tumour samples, compared to control group. Tumour tissue showed a significant ($p<0.001$) downregulation of SLC22A15, SLC22A17, SLC22A3, SLC22A31, SLC22A4, SLC22A5, whereas expression of SLC22A18, SLC22A20P was significantly ($p<0.001$) higher in LUSC tissues than in normal lung tissues. SLC22A23, SLC22A1, SLC22A16, SLC22A3 genes showed a broader but no significant expression in cancer samples than in relative normal tissues.

Analysis of expression of SLC22A genes in LUSC patients is shown in Figure 3-6 (B). Tumour tissue showed a significant ($p<0.001$) downregulation of SLC22A15, SLC22A17, SLC22A3, SLC22A31, SLC22A4, SLC22A5, whereas expression of SLC22A18, SLC22A20P was significantly ($p<0.001$) higher in LUSC tissues than in normal lung tissues. SLC22A23, SLC22A1, SLC22A16, SLC22A3 genes showed a broader but no significant expression in cancer samples than in relative normal tissues.

Interestingly, the downregulation of SLC22A15, SLC22A17, SLC22A3, SLC22A4, and SLC22A5 was observed in both LUAD and LUSC cohorts, suggesting potential commonalities in the dysregulation of these genes across both lung cancer subtypes. Conversely, SLC22A18 and SLC22A20P displayed significant upregulation ($p < 0.001$) in both LUAD and LUSC tumor tissues. Whereas SLC22A16 showed no significant difference in expression in either LUAD or LUSC, suggesting its potential lack of involvement in the development of these specific cancer subtypes. Overall, this analysis identified both shared and distinct patterns of differential expression for SLC22A genes in LUAD and LUSC, suggesting potential similarities and differences in the underlying molecular mechanisms of these two lung cancer subtypes.

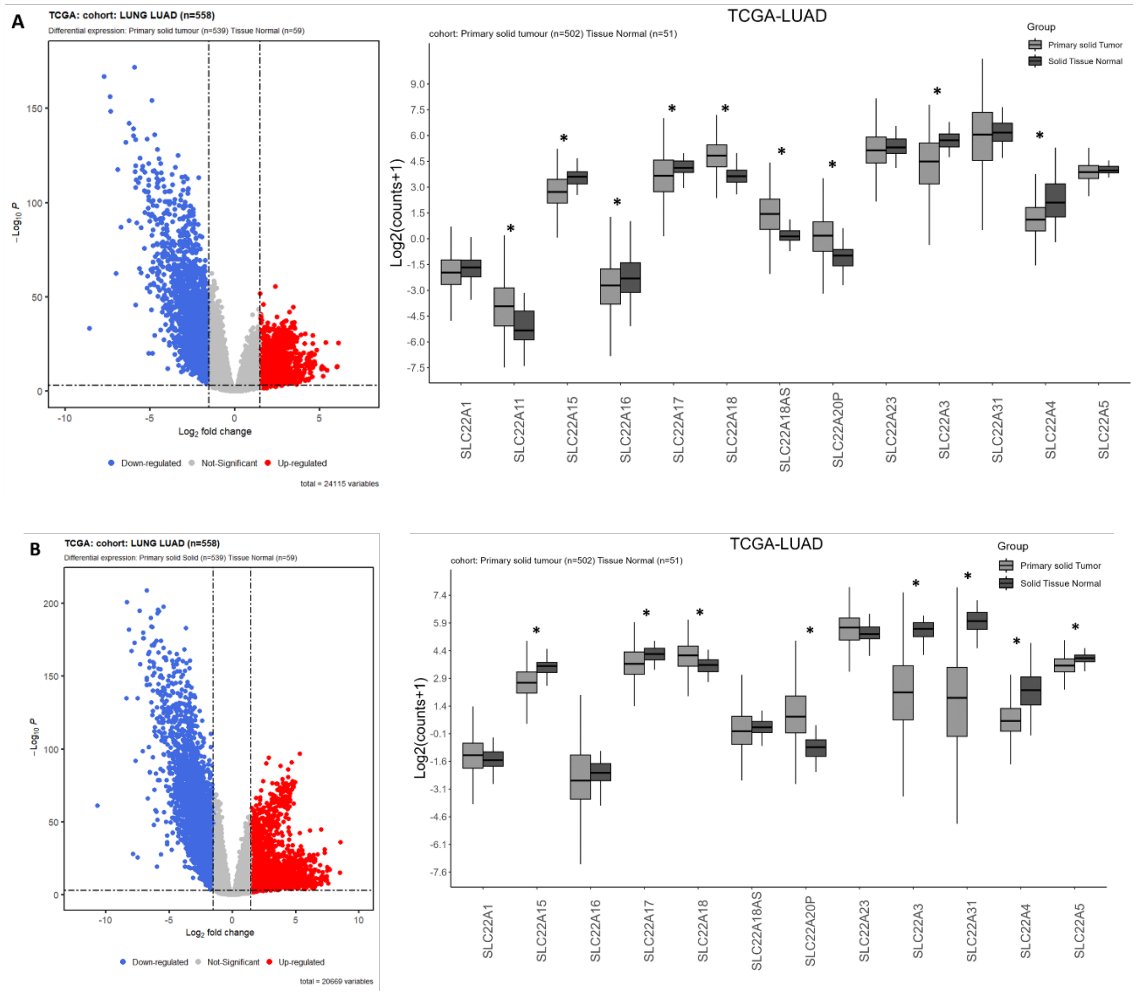


Figure 3-6. Identification of differentially expressed genes (DEGs) in lung cancer from the TCGA profiles. Right, Volcano plots of up-regulated and downregulated DEGs. Left, Expression profiles of SLC22A genes (A) Lung Adenocarcinoma (LUAD) (B) Lung Squamous Cell Carcinoma (LUSC).

3.1.1. Prognostic values of SLC22A1, SLC22A4 and SLC22A5 in LUAD and LUSC samples

While several organic cation transporters (OCTs) and organic carnitine transporters (OCTNs) are expressed in the lung epithelium, their roles remain under investigation (Koepsell et al., 2007). Previous studies have categorized SLC22A1, SLC22A4, and SLC22A5 as potential prognostic markers in cancers like liver and kidney. Prior analysis within the context of lung cancer also revealed a significant difference in SLC22A4 expression between lung squamous cell carcinoma (LUSC) and lung adenocarcinoma (LUAD) compared to healthy tissues. Based on this existing evidence and their established expression in the lung (as shown in Table 1.3), we focused on SLC22A1, SLC22A4, and SLC22A5 to investigate their potential as prognostic markers in lung cancer subtypes (LUAD and LUSC).

Kaplan-Meier survival plots were generated to investigate the association between SLC22A gene expression (SLC22A1, SLC22A4, and SLC22A5) and overall survival in patients diagnosed with lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) retrieved from The Cancer Genome Atlas (TCGA) database. Patients were classified into two groups ("down" or "low") based on the median expression as reference,

Results obtained from the overall survival analysis of adenocarcinoma samples (LUAD) are shown in Figure 3-7 (D-F), whereas plot for lung squamous adenocarcinoma (LUSC) are shown in Figure 3-8. In LUAD patients, the log-rank p-values for SLC22A1 ($p = 0.8$), SLC22A4 ($p = 0.91$), and SLC22A5 ($p =$

0.45) indicate no statistically significant differences in survival probabilities between the high and low expression groups for any of the three genes.

However, for LUSC patients, the p-value for SLC22A1 ($p = 0.091$) suggests a potential trend towards a difference in survival based on SLC22A1 expression. Further investigation with a larger sample size might be necessary to confirm this association. SLC22A4 ($p = 0.95$) and SLC22A5 ($p = 0.52$) expression levels still showed no statistically significant association with overall survival in LUSC patients.

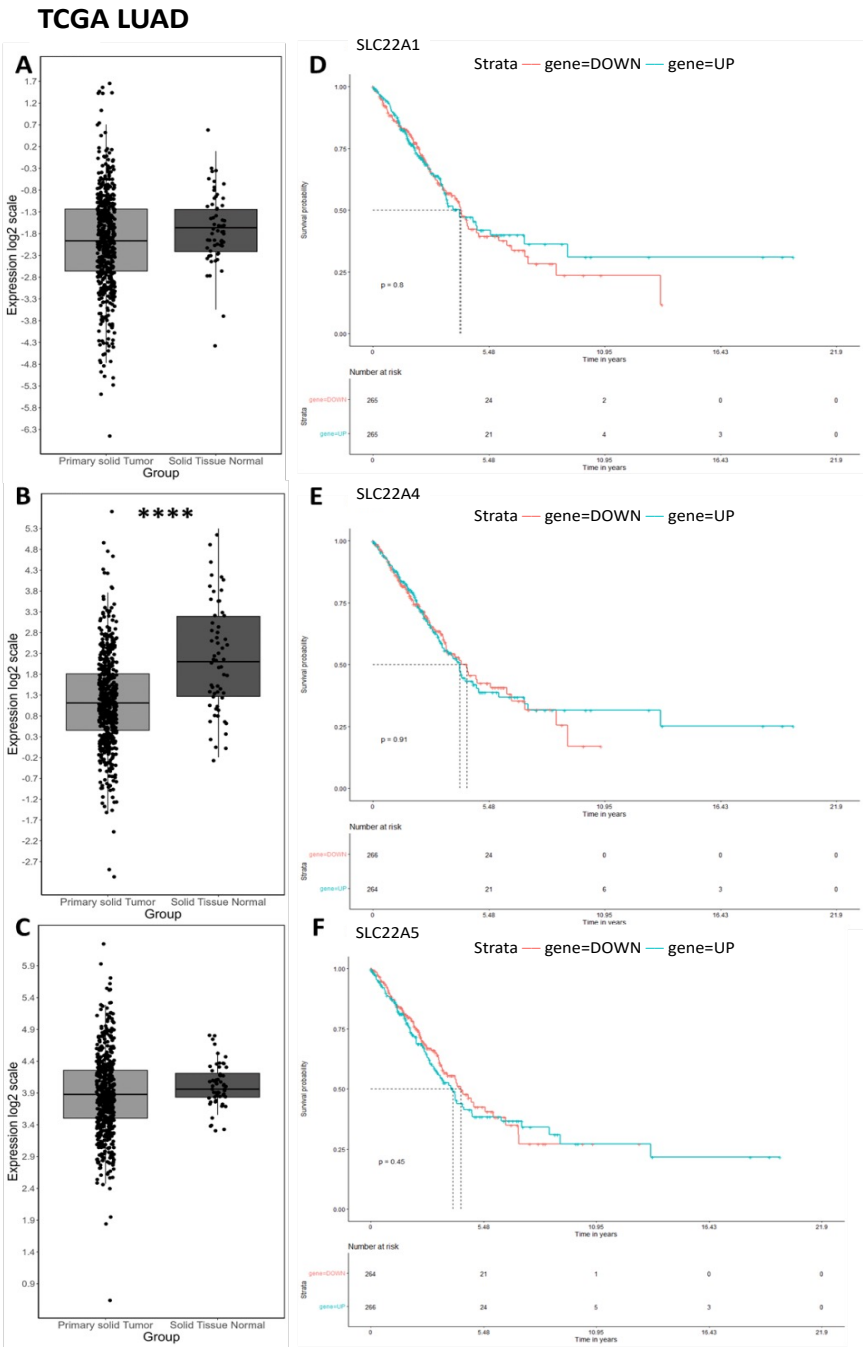


*Figure 3-7. Expression and survival analyses of the three SLC22A genes in lung adenocarcinoma (LUAD). (A-C). Expression of SLC22A1, SLC22A4 and SLC22A5 in LUAD tissues and adjacent non-tumour tissue. (D-F) Kaplan-Meier survival analysis of SLC22A1, SLC22A4 and SLC22A5 in the TCGA cohort. X-axis shows the time for survival (days) and y-axis shows the probability of survival, where 1.0 corresponds to 100 percent. P value for Kaplan-Meier plot showing results from analysis of correlation between mRNA expression level and patient survival. ****P < 0.0001.*

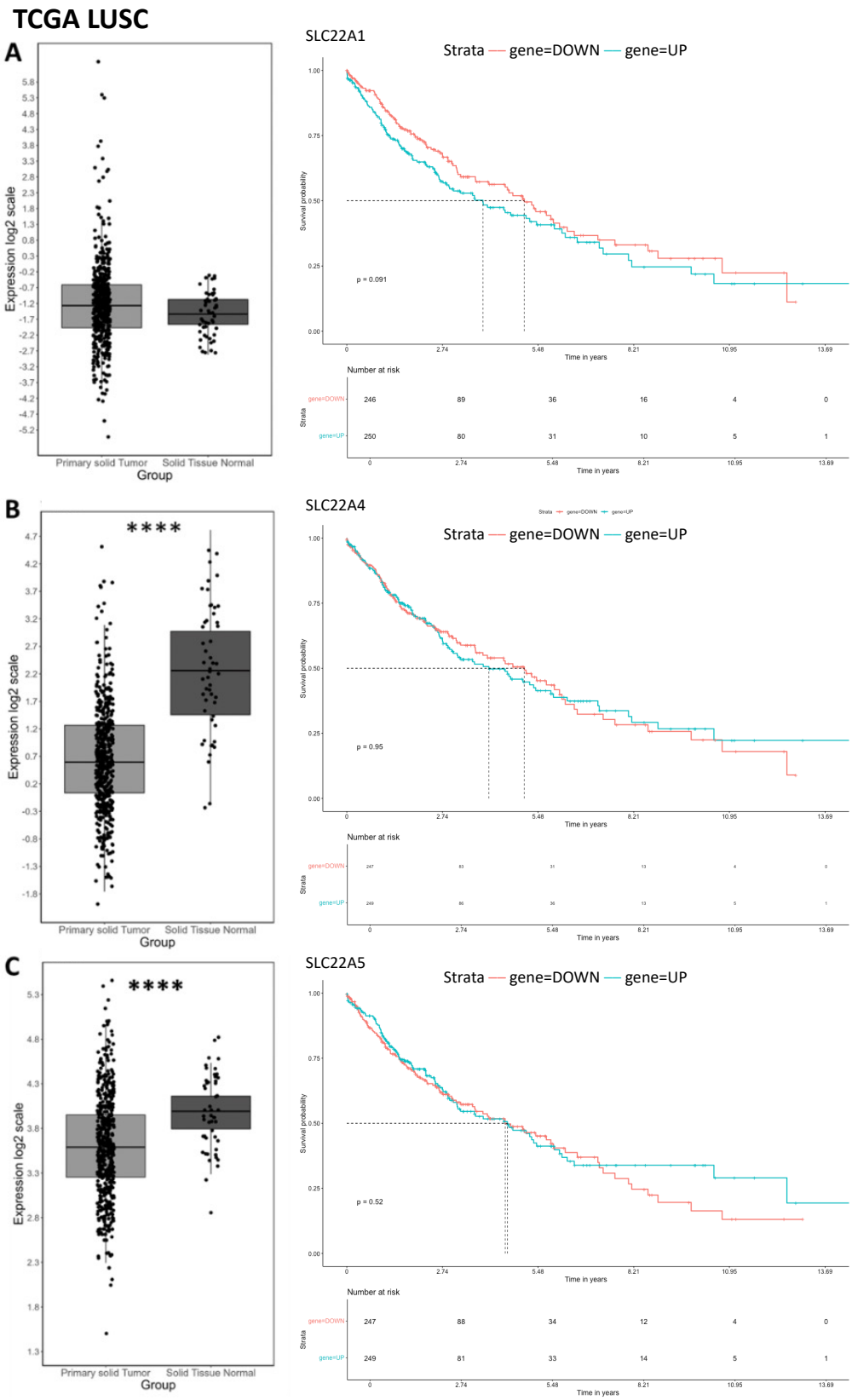


Figure 3-8. Expression and survival analyses of the three SLC22A genes in lung squamous cell carcinoma (LUSC). (A-C). Expression of SLC22A1, SLC22A4 and

*SLC22A5 in LUSC tissues and adjacent non-tumour tissue. (D-F) Kaplan-Meier survival analysis of SLC22A1, SLC22A4 and SLC22A5 in the TCGA cohort. X-axis shows the time for survival (days) and y-axis shows the probability of survival, where 1.0 corresponds to 100 percent. P value for Kaplan-Meier plot showing results from analysis of correlation between mRNA expression level and patient survival. ****P < 0.0001.*

3.2. Discussion

Development of high-throughput sequencing technology and its popularization in clinical research has led to an increase in prognostic biomarkers and identification of therapeutic targets in various malignancies through bioinformatics analyses. Nonetheless, limited studies have published connecting OCT/N expression to lung diseases. Publicly available gene expression data, stored in databases, contain vast amounts of data that can be re-analysed to identify gene profiles associated with diseases.

mRNA expression profiles of SLC22A1, SLC22A4, ad SLC22A5 was investigated in samples of asthma, chronic obstructive pulmonary disease (COPD), lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). Results showed that while not statistically significant, expression of SLC22A1in COPD samples showed a slight increase in expression. In contrast, expression of SLC22A1 in LUSC samples showed a broader range. Additionally, samples from the LUAD datasets revealed a lower expression profile of SLC22A1 in patients with adenocarcinoma. It has been demonstrated that downregulation of SLC22A1 has been associated with a significant reduction ($p<0.05$) in overall patient survival in cholangiocellular carcinoma (CCA). In addition, low expression was further linked ($p=0.02$) with advanced CCA stages and their enlarged diameter (Lautem et al., 2013).

Based on the Kaplan-Meier curves and the log-rank test results, SLC22A1, SLC22A4, and SLC22A5 gene expression levels do not appear to be strong prognostic markers for overall survival in LUAD patients within this TCGA dataset. However, SLC22A1 expression in LUSC patients warrants further exploration due to the borderline significant p-value ($p = 0.091$).

Low expression of SLC22A1 in hepatocellular carcinoma (HCC) tumour samples was associated with a worse patient survival ($p<0.05$), with advanced cancer stages ($p=0.025$), larger tumours diameter ($p=0.035$) and worse differentiation ($p=0.001$) (Heise et al., 2012). In accordance, the Human Protein Atlas, categorised high expression of SLC22A1 gene ($p<0.001$) as an association with favourable prognosis in liver cancer according to data analysed from the TCGA portal (*The Human Protein Atlas*, n.d.).

Genetic variants of the SLC22A1 have previously been linked with an increased risk for chronic inflammatory diseases such as Crohn's disease and ulcerative colitis (Jung et al., 2017). However, expression analysis of SLC22A4 showed no significant difference in COPD and asthma datasets when compared to healthy samples. A previous study carried out in an in-vitro asthmatic-like simulation with lipopolysaccharide (LPS) and house dust mite (HDM) in the human broncho-epithelial cell line Calu-3 cells showed a significant upregulation in SLC22A1, SLC22A3, SLC22A4 and SLC22A5 on mRNA and protein levels when exposed to LPS , while HDM showed similar effects only on SLC22A1, SLCC22A3 and SLC22A5 expression (Mukherjee et al., 2017). However, this outcome is contrary to that of Rotoli et al., (2020) where inflammatory stimuli with LPS and tumour necrosis factor-α (TNFα) showed no

significant changes in mRNA levels of SLC22A1, SLC22A3 or SLC22A5. Nonetheless, it is important to bear in mind the possible limitation of the use of airway epithelial brushing rather than single-cell analysis since bronchial and alveolar epithelial cells have demonstrated different responses to LPS-induced inflammation (Schulz et al., 2002). Protein and mRNA levels of SLC22A1/2 were downregulated in alveolar epithelial cells (A549) after treatment with LPS (D. Li et al., 2020). Likewise, challenging with cigarette smoke extract (CSE) induced a significant reduction in SLC22A4 and SLC22A5 mRNA levels (Qi et al., 2020).

In addition, a previous study between ex-smokers with severe COPD and healthy subjects reported no difference in mRNA levels of SLC22A1, SLC22A3, SLC22A4 and SLC22A5 models (Berg et al., 2014). Current analysis found no significant difference in expression profiles of SLC22A1, SLC22A4 and SLC22A5 in asthmatic patients. Although, SLC22A4 showed a slight decrease in expression in asthma samples. A previous genome wide analysis, however, revealed a lower expression in both mild-moderate ($p < 0.01$) and severe asthma ($p<0.0001$) subjects compared to controls (Portelli et al., 2021)

A previous study has reported association between SLC22A4 variants with cancer progression in sporadic colorectal cancer in early age and in ulcerative colitis (UC) (Q. Wang et al., 2020). This relationship needs further exploration, since, as an acetylcholine transporter, makes it a constituent of the non-neuronal cholinergic system. It is plausible to consider SLC22A1 involvement in the pathophysiology of lung cancer (Qian et al., 2019). The role of acetylcholine, acting via nicotinic or muscarinic receptors, adds further significance to this

connection, as it influences pivotal cellular processes such as proliferation, Epithelial- Mesenchymal Transition (EMT) induction, migration, and invasion of human lung cancer cells. It has been documented that acetylcholine (ACh) exerts a detrimental impact on the progression of human lung cancer, functioning as a growth factor in this context (Friedman et al., 2019). These findings underscore the potential implications of SLC22A4 in the context of lung cancer and warrant further investigation in this important area of research.

L-carnitine is an essential nutrient involved in fatty acid transport into mitochondria and oxidation to produce energy and loss of OCTN2 function causes systemic carnitine deficiency. Results revealed that while SLC22A5 expression in LUAD samples showed a broader expression, it was not found to be statistically significant. In contrast, the expression of SLC22A5 in LUSC samples was found to be significantly lower when compared to healthy lung samples. Berg et al., 2018 showed no significant difference in expression of SLC22A1, SLC22A4 and SLC22A5 genes between COPD patients and healthy controls. Moreover, two large-scale GWAS with physician-diagnosed asthma found a significant association between disease and single-nucleotide polymorphisms in SLC22A5 (Moffatt et al., 2010; Ober & Nicolae, 2011). This differ from the findings presented here, where SLC22A5 expression was significantly lower ($p<0.001$) than control samples. This discrepancy could be attributed to the sampling size used in the previous study, where they analysed biopsy material from a small cohort of only three healthy subjects and seven patients with severe COPD.

Microarray and qPCR results may not directly correlate as they are averaged in a different manner (Morey et al., 2006), although high-throughput analysis provides a higher discovery power and higher sensitivity to quantify transcripts. Expression of SLC22A5 has been generally found to be higher in the upper lung, where a previous study by Sakamoto et al., (2015b) found trachea-bronchial cells (BEAS2-B) to have the highest protein levels (fmol/µg) among immortalised and primary cultured cells.

Previous studies have investigated the role of SLC22A5 in various types of cancer, including glioblastoma (GBM). These studies have shown that overexpression of SLC22A5 in GBM is associated with poor overall patient survival. Additionally, administration of L-carnitine to GBM cells has been found to increase their tolerance towards cytotoxicity (Fink et al., 2019). Furthermore, it has been demonstrated that glioma cells survival depends on both fatty acid oxidation and functional carnitine transport by OCTN2 (Juraszek et al., 2021). Additionally, it was found that overexpression of SLC22A5 in skin fibroblasts does not increase energy metabolism, but rather increases the expression of gene clusters connected to mitochondria (Hlobilová, 2017).

While the current study has shed light on the expression profiles of SLC22A1, SLC22A4, and SLC22A5 in the context of lung diseases, several limitations acknowledge further investigation. One limitation is the lack of functional validation to complement the in-silico analysis. The availability of mutation data within the same TCGA cohorts presents an opportunity to explore the potential functional consequences of these mutations and their potential links to the observed expression patterns.

Furthermore, the current study primarily focused on gene expression levels without delving into the spatial distribution of these proteins within the tissue. Immunofluorescent staining on tissue samples could provide valuable insights into the cellular and subcellular localization of these transporters, potentially revealing distinct expression patterns not captured by bulk RNA-sequencing data. Additionally, manipulating gene expression levels in relevant cell lines and observing resulting phenotypic changes would provide further evidence for the functional roles of these genes in lung disease.

This analysis revealed a significant overlap in the expression patterns of several SLC22A genes between LUAD and LUSC, suggesting potential common regulatory pathways involving these genes in lung cancer development. However, the presence of partially overlapping and subtype-specific expression patterns for some genes highlights the need for further investigation into their distinct roles in each cancer subtype. Further research is needed to understand the implications in the lung epithelium and its potential role in drug therapy for lung diseases.

3.3. Conclusion

In conclusion, the functional role of transporters in lung epithelia, specifically SLC22A1, SLC22A4 and SLC22A5, is an important area of research in the field of drug disposition, toxicity and cancer prognosis. While analyses were conducted in retrospective, using public databases, results provided valuable insights into the potential impact of transporter regulation on pulmonary drug disposition and toxicity. SLC transporters showed no difference in expression in asthma compared to healthy samples. The study identified SLC22A5 expression

to be significantly higher in COPD samples. In cancer, however, there are dramatic changes in expression patterns of SLC transporter. SLC22A4 expression was downregulated in LUAD and LUSC, whereas SLC22A5 expression was lower in LUSC samples. These findings have significant implications for the understanding of expression of OCTNs transporters in lung cancer and possible implication in cell physiology in the lung epithelia. Further research is required to establish the relationship on the expression profiles of SLC transporters in lung malignancies as biomarkers and therapeutic targets.

CHAPTER 4. GENERATION OF OCT/OCTN EDITED CELL LINES USING CRISPR/CAS9 SYSTEM

CHAPTER 4. GENERATION OF OCT/OCTN EDITED CELL LINES USING CRISPR/CAS9 SYSTEM

4.1. Introduction

OCT1 is involved in the transport of a diverse range of substrates, encompassing drugs, toxins, and endogenous compounds. Inactivation of OCT1, as demonstrated by studies employing Oct1−/− mice, leads to a significant reduction in both hepatic uptake and intestinal excretion of organic cations (Jonker et al., 2001). Further investigation utilizing Oct2 single-knockout and Oct1/2 double-knockout mice revealed the cooperative role of OCT1 and OCT2 in the renal secretion of small organic cations (Jonker et al., 2003). While previous studies have focused on their roles in renal and hepatic secretion these transporters are also expressed in lung epithelia [mention a relevant reference here]. Their function in this tissue remains less understood.

The role of SLC22A1, SLC22A4, and SLC22A5 in lung epithelia has been previously examined using a variety of inhibitors and knockdown approaches with RNA interference (Mukherjee, 2012). Drawbacks of this methodologies have been associated with certain limitations, including incomplete knockdown of the mRNA level and a silencing duration of up to five days, which necessitates re-transfection for an extended duration, thereby restricting the scope of the experiments Moreover, the use of inhibitors in vivo may be restricted by toxicity, off-target effects, and variable pharmacokinetics, warranting a cautious interpretation of results and complementation with additional techniques, such as genetic knockout approaches, to validate the findings. In light of these limitations, the present study explores the Cas9-CRISPR-mediated disruption of

these transporters as an alternative technique to examine the functionality in different cell types.

CRISPR-Cas9 technology is a microbial adaptative immune system that has revolutionised the field of gene editing. CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats, these repeat elements allow for genome editing. The bacterial system can be reconstituted in mammalian cell using minimal components: the CRISPR-associated nuclease Cas9 (SpCas9), a specificity-determining CRISPR RNA (crRNA) composed of a 20-nt sequence, and an auxiliary trans-activating crRNA (tracrRNA) (Ran et al., 2013). Cas9 recognises a protospacer adjacent motif (PAM, 5'-NGG) that is immediately preceded by the target DNA that is complementary to the spacer sequence of crRNA. PAM recognition induces a structural alteration in Cas, resulting in unwinding of target DNA and the generation of an R-loop between gRNA and target DNA.

Recognition of PAM induces a structural alteration in Cas9, where the nicking of HNH and RuvC domains catalyses the cleavage of target and nontarget strands, respectively. Upon cleavage by Cas9, the target locus typically undergoes one of two major pathways for DNA damage repair: error-prone non-homologous end joining (NHEJ) and template-dependent homology-directed repair (HDR). Both of which can be used to achieve a desired editing outcome. The NHEJ pathway facilitates the repair of DSBs by re-joining the two broken ends, which can lead to the introduction of random insertions or deletions at the DSB site, ultimately disrupting the gene sequence. Conversely, the HDR pathway utilizes homologous recombination when a donor template with a

matching sequence is accessible, thereby allowing for the insertion of desired nucleotides into the targeted DNA region. NHEJ can be harnessed to mediate CRISPR-edited, as indels occurring within a coding exon can lead to frameshift mutations and premature stop codons. Multiple DSBs can additionally be exploited to mediate larger deletions in the genome.

A simpler, fast and efficient version can be achieved by fusing the crRNA and tracrRNA into a chimeric single guide RNA (sgRNA) using an already prepared expression cassette.

4.1.1. Ribonucleoprotein complexes

Alternative approach consists of the Cas9 protein in complex with the sgRNA. Cas9 RNP are capable of cleaving genomic targets with similar efficiency as compared to plasmid-based expression of Cas9/gRNA and can be used for most of the current genome engineering applications of CRISPR.

The delivery of the complex of Cas9 protein and sgRNA has emerged as a powerful and widespread method for genome editing due to reduced off-target effects. When pre-formed RNPs are introduced by nucleofection into cells, the enzyme rapidly starts cutting targeted genomic DNA. sgRNAs can be purchased from a commercial vendor or *in-vitro* transcribed from ssDNA, which can be generated by PCR.

4.1.2. CRISPR nickase system

An aspartate-to-alanine (D10A) mutation in the RuvC catalytic domain allows the Cas9 nickase mutant (Cas9n) induces single-strand nicks rather than double-strand breaks, and then subsequent preferential repair through HDR can

potentially decrease the frequency of unwanted indel mutations from the off-target DSBs. Paired sgRNAs can guide Cas9n to simultaneously nick on opposite strands both strands of the target locus to mediate a DSB, increasing the specificity of target recognition. Thus, DSBs would only occur if both sgRNAs are able to locate target sequences within a defined space (up to 100 bp). nD10A has been used to generate paired DNA nicks and efficiently disrupts genes in *Drosophila* and cell culture (Amo et al., 2022).

Due to the action of the endogenous base-excision repair pathway on nicked genomic DNA, it is anticipated that Cas9 nickases would impart minimal or negligible damage to the genome (Shen et al., 2014).

The most important consideration for double-nicking sgRNA design is the spacing between the two targets. D10A creates 5' overhangs, therefore sgRNAs must be designed such that 5′ overhangs are generated upon nicking. PAM-out design of the sgRNAs has been previously found to have a higher editing profile, compared to a configuration with closer PAMs to the targeted region (PAM-in) (Amo et al., 2022; Labun et al., 2021; Schubert et al., 2021). The target loci for the sgRNA pairs must also have an optimal distance of 40–70 nt.

To test the ability of the vectors to disrupt genes for SLC22A1, SLC22A4 and SLC22A5 transporters, human embryonic kidney cells (HEK293) were used as a primary approach. They have been extensively used for pharmaceutical purposes, cell biology research and protein production (Y.-C. Lin et al., 2014; P. Thomas & Smart, 2005). Among the principal attributes of the cell lines are its easy growth and maintenance as well as known to be amenable to modification by Cas9-CRISPR.

Different CRISPR strategies were designed to target the SLC22A1, SLC22A4 and SLC22A5 genes. Kidney cells HEK293, lung epithelial A459 and Calu-3 cells were transfected and clones were further expanded and analysed for gene edition by T7E1 and Sanger sequencing. Cell proliferation, adhesion and migration of CRISPR-edited cells were evaluated using Click-iT EdU, toluidine blue and scratch wound, respectively.

4.2. Results

4.2.1. Construction of the pX459 sgRNA vector

All-in-one plasmids have been previously used for the simple and rapid construction and simple delivery into the cells. To target the proteins SLC22A1, SLC22A4, and SLC22A5, two pairs of sgRNAs were designed, and a set of guides targeting the human EMX1 gene were used as control, sequence was obtained from literature (Ran et al., 2013). RNA guides were inserted into the pX459 plasmid, which contains resistance to puromycin, enabling the selection of clones. The plasmid contains the Cas9 protein, a sgRNA scaffold and a BbsI cloning site for insertion of the guide sequence (Ran et al., 2013).

Figure 4-1 shows the gene structure of Solute Carrier Family 22 (Organic Cation Transporter), Member 1 located on chromosome 6q25.3, two sgRNAs were designed to target the gene on Exon4 and Exon5, according to the predominant transcript (NM_003057.3).

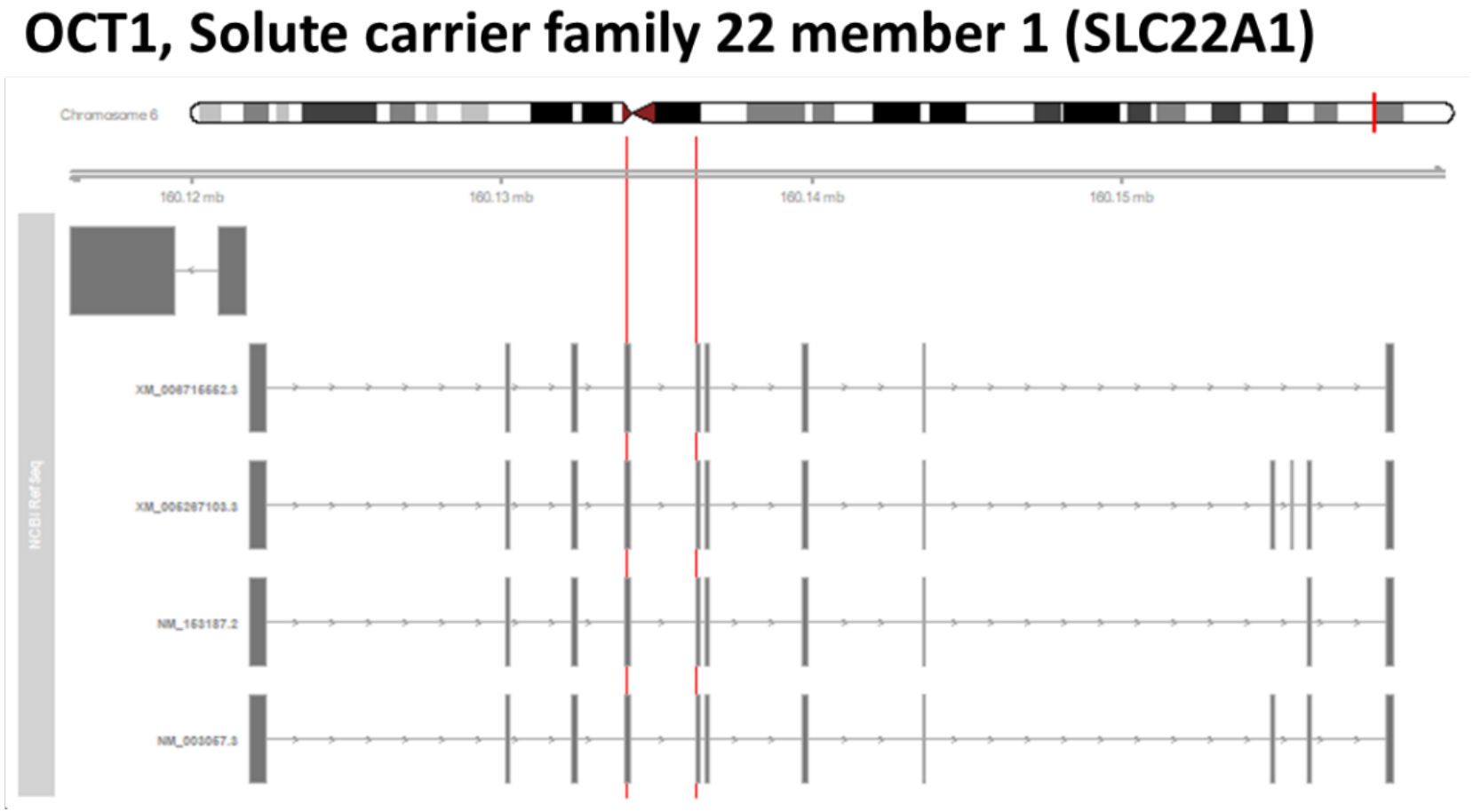


Figure 4-1. Gene structure of the Solute Carrier Family 22 (Organic Cation Transporter), Member 1 located on chromosome 6q25.3. Gene consists of 11 exons and comprises 2 splice variants. two sgRNAs were designed to target the gene on Exon4 and Exon5, according to the predominant transcript (NM_003057.3). Accessions with XM_ prefix are model RefSeqs produced either by NCBI's genome annotation pipeline or copied from computationally annotated submissions. Red lines highlight the location of exons for designed sgRNAs

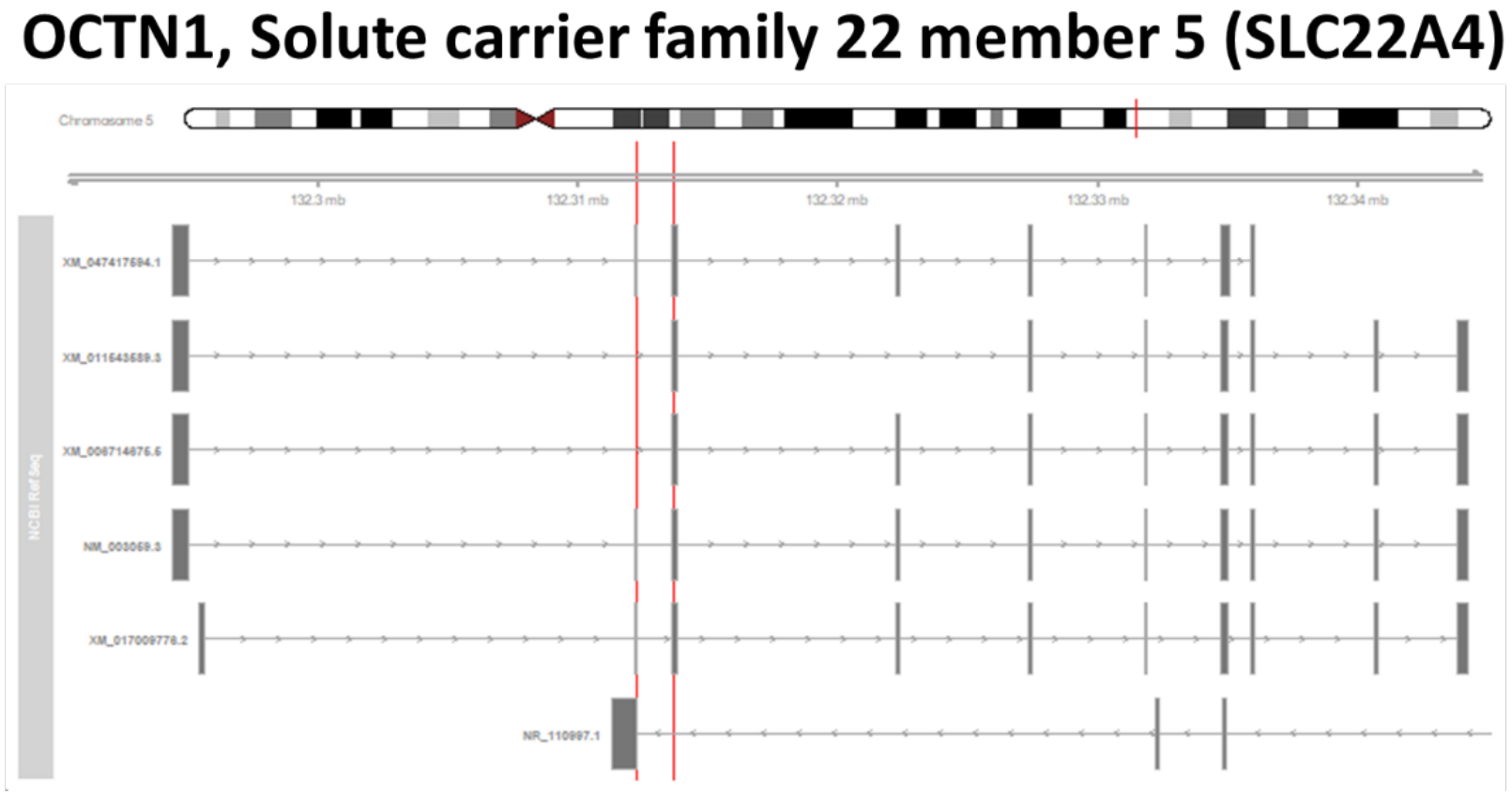


Figure 4-2. Gene structure of the Solute Carrier Family 22 (Organic Cation Transporter) Member 4 located on chromosome: 5q31.1, Gene is encoded in 10 exons. Two sgRNAs were designed for CRISPR-mediated disruption of SLC22A4 on Exon 2 and Exon3, according to the main transcript NM_003059.3. Accessions with XM_ prefix are model RefSeqs produced either by NCBI's genome annotation pipeline or copied from computationally annotated submissions. NR_110997.1 represents a long non-coding RNA (lncRNA). Red lines highlight the location of exons for designed sgRNAs.

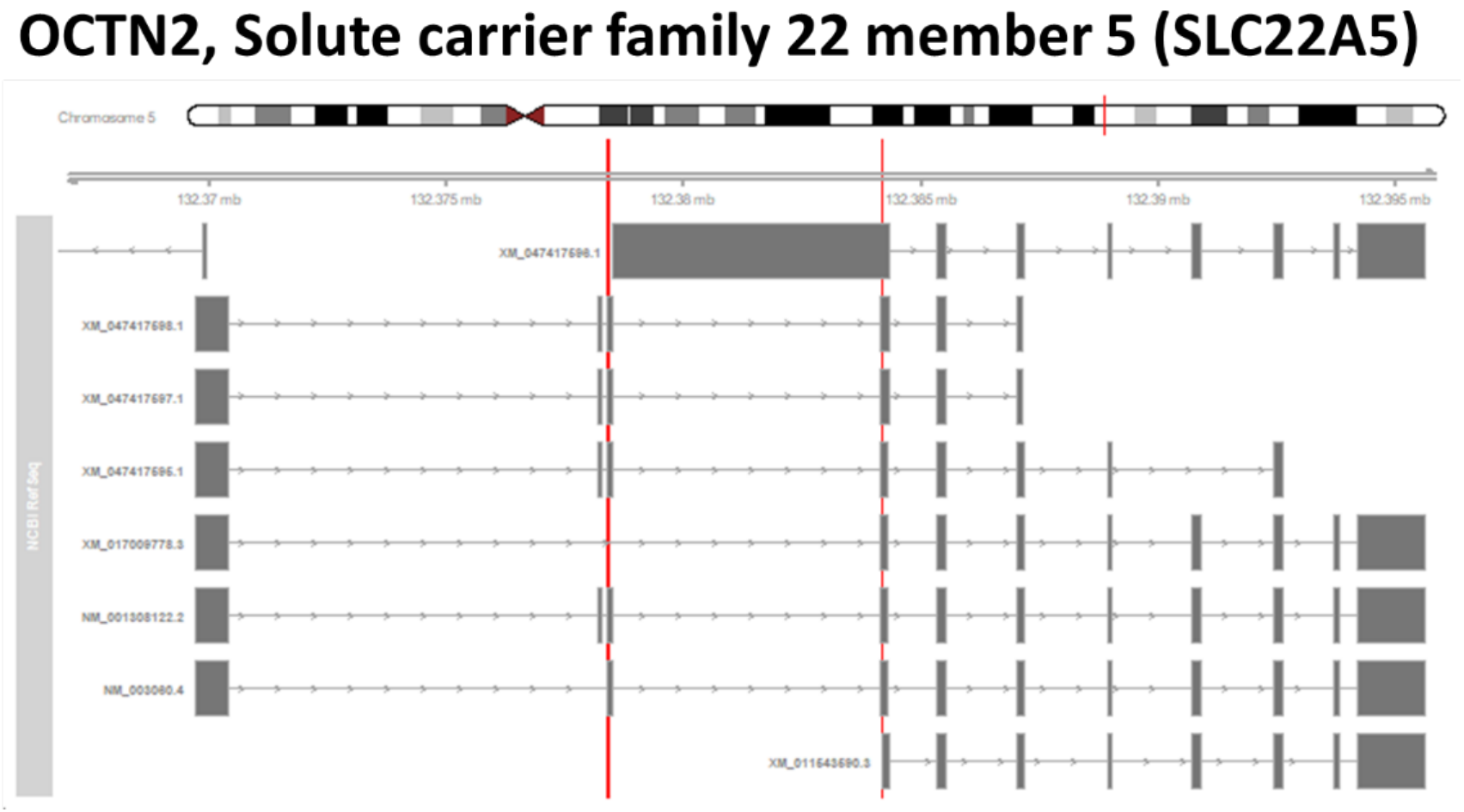


Figure 4-3. Gene structure of Solute Carrier Family 22 (Organic Cation Transporter) Member 5 located on chromosome: 5q31.1. Gene is encoded in 11 exons and has 2 splice variants. Two sgRNAs were designed to target Exon 3 and Exon 4, respectively, according to the canonical transcript NM_001308122.2. Accessions with XM_ prefix are model RefSeqs produced either by NCBI's genome annotation pipeline or copied from computationally annotated submissions. Red lines highlight the location of exons for designed sgRNAs

For the Solute Carrier Family 22 (Organic Cation Transporter) Member 4 located on chromosome: 5q31.1, two sgRNAs were designed for CRISPR-targeting of SLC22A4 on Exon 2 and Exon3, according to the main transcript NM_003059.3 (Figure 4-2). Lastly, two sgRNAs were designed to target the Solute Carrier Family 22 (Organic Cation Transporter) Member 5 located on chromosome: 5q31.1, on Exon 3 and Exon 4, respectively, according to the transcript NM_001308122.2 (Figure 4-3).

Figure 4-4 provides a visual representation of the process of inserting guide sequence oligos into a plasmid using BbsI enzyme-mediated digestion and ligation. The process starts by using oligos, which contain overhangs that are necessary for ligation into the pair of BbsI sites in the vector. These overhangs match the corresponding sites in the plasmid, with the top oligo representing the 20-bp sequence preceding the 5'-NGG in genomic DNA. The next step involves the digestion of the vector with the enzyme BbsI, which results in a sequence compatible with the overhang created with the oligos. The final product is the plasmid with the guide sequence inserted.

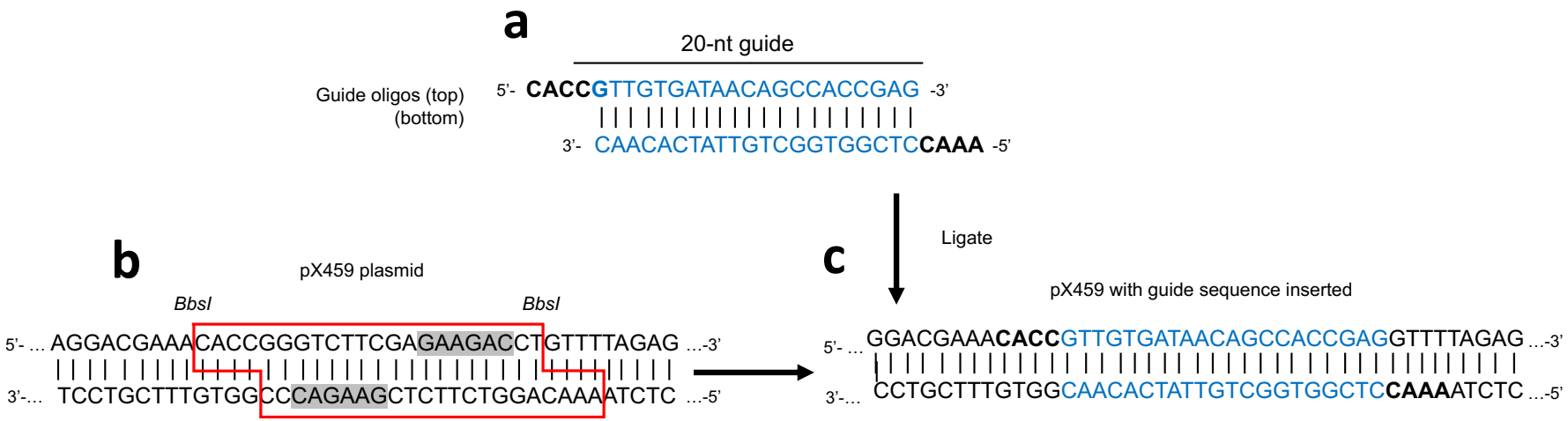


Figure 4-4. Schematic representation of the insertion of guide sequence oligos into the plasmid. a) The oligos (blue) contains the overhangs (bold) for ligation into the pair of BbsI sites in the vector, both matching the ones in the plasmid (the top oligo is the 20-bp sequence preceding the 5'-NGG in genomic DNA). b) Digestion of pX459 with BbsI allows the replacement of the restriction sites (red outline) with a direction insertion of annealed oligos. Recognition sites are marked with grey rectangles c) pX459 plasmid with guide sequence (blue).

Insertion of specific guides was confirmed through Sanger sequencing (Figure 4-5). BbsI restriction sites allows for scarless cloning. Sequencing results indicate the successful insertion of sgRNAs between the U6 promoter and the remainder of the scaffold.

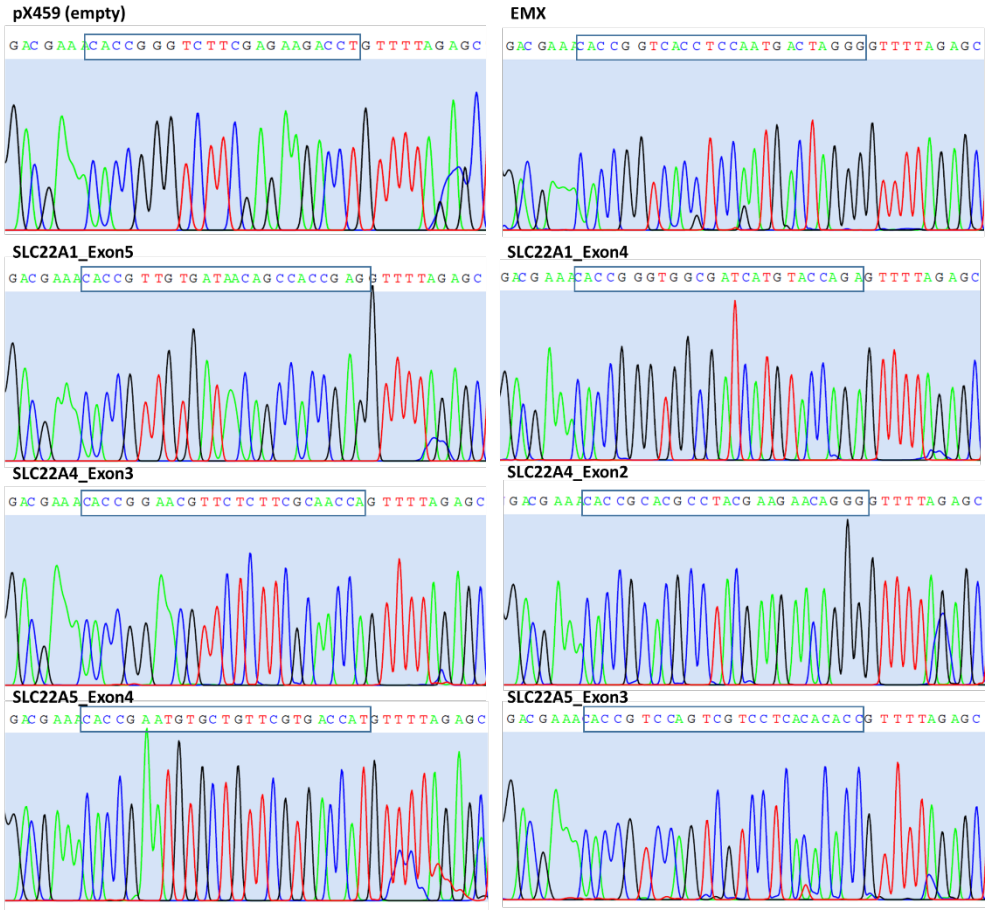


Figure 4-5. Sequencing results of the sgRNA cloned into the pX459 CRISPR/Cas9- vectors. Plasmids were sequenced using the U6-Fwd primer. Blue box highlights the insertion site for the guides, accordingly.

4.2.2. Antibiotic concentration for puromycin selection

The optimal concentration of puromycin for cell selection was determined by evaluating the dose-dependent effect on cell viability. The lowest concentration that resulted in 90-99% cell death within 7 days was considered as the optimal concentration. As shown in section 2.2.4 General Tissue Culture a concentration

of 0.5 µg/ml of puromycin resulted in optimal cell death in HEK293 and A549 cells, while a concentration of 1.5 µg/ml was optimal for Calu-3 cells.

4.2.3. Validation of sgRNAs in HEK293 cells

In order to determine the effectiveness of transfection in HEK293 cells, an assessment was conducted using an empty Green Fluorescent Protein (GFP) expression vector. The GFP expression serves as a marker for successful transfection, and its detection can be used to quantify the efficiency of the transfection process. The results of this assessment are presented in Figure 4-6, where the expression of GFP at 488nm is reflected by a fluorescence filter with a 488 nm excitation. Cells transfected with empty pX461 plasmid yield a transfection of about 30%, based on fluorescence from GFP. The data obtained from this evaluation provides valuable information on the transfection efficiency of HEK293 cells and can be used to optimize the transfection protocol for future experiments.

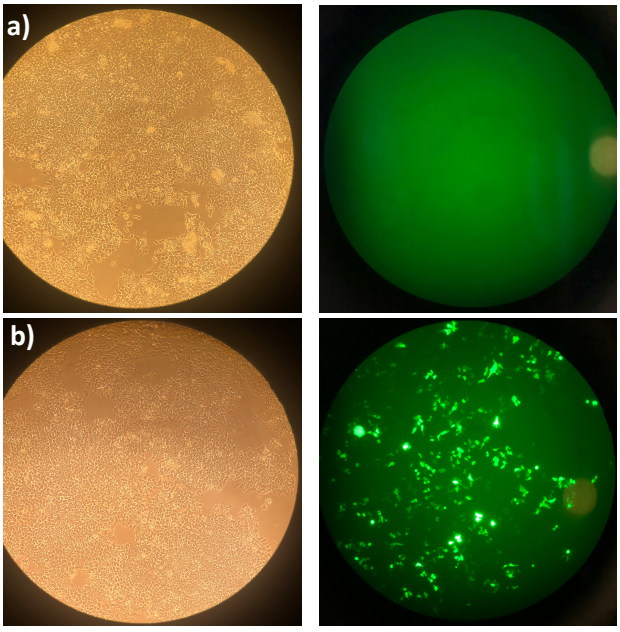


Figure 4-6. HEK293 cells that were transfected with a Green Fluorescent Protein (GFP) expression vector and plated in a 6-well. The GFP fluorescence at 488nm is shown on the right side a) Negative Control, transfection method without plasmid. b) Transfection with empty GFP plasmid (pX461).

The validation of pX459 plasmids was performed prior to transfecting HEK293 cells with CRISPR vectors and an empty vector as a control. The cells were then subjected to a 7-day screening process under puromycin selection (0.5 μg/ml), to enrich for cells that successfully express. After screening, bulk population of HEK293 cells were analysed for gene editing. CRISPR activity was determined using the T7 endonuclease 1 (T7E1) mismatch detection assay, which is based on the ability of T7E1 to distinguish between homo- and heteroduplex DNA. Notably**,** HEK293 cells, a hypotriploid human embryonic kidney cell line with a modal chromosome number of 64 (30% of cells) and exhibiting instances of higher ploidy (4.2% of cells), were chosen for this analysis.

The EMX1 gene was used as a control to assess the activity of the CRISPR system (Ran et al., 2013). Sequencing results were obtained from the bulk population of HEK 293 cells that were transfected with the pX459 plasmid, targeting the EMX1 gene in Exon3. Figure 4-7 presents the results of this analysis, with the blue line highlighting the position of the sgRNA and the orange line representing the PAM site. The sequencing results showed that the CRISPR system was able to effectively target the desired region of the EMX1 gene.

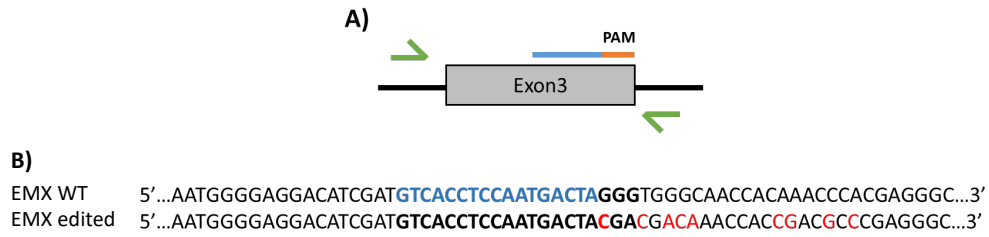


Figure 4-7. Sequencing results from the transfected HEK293 with pX459 plasmid targeting EMX1 in Exon3. Blue line highlights the position of the sgRNA, orange line represents the PAM site. A) Representation of targeted sequence in EMX1, B) Sequence alignment for EMX1 Exon3, Blue font highlights the sgRNA sequence, bold letters highlight the PAM site

Overall, these results provide important evidence for the effectiveness of the CRISPR system in targeting specific regions of the EMX1 gene and highlight its potential as a powerful tool for genome editing.

T7 endonuclease I (T7E1) assay was employed to analyse the presence of CRISPR-mediated edits in the targeted regions of the SLC22A genes within a bulk population of HEK293 cells. This assay detects mismatched DNA duplexes, and successful editing would introduce mismatches at the target site. Primers flanking the edited regions were designed for PCR amplification. The PCR products from wild-type and CRISPR-edited cells were then digested with T7E1. The digested fragments were resolved on a 1.5% agarose gel (Figure 4-8). Band after T7 assay sample was compared to that of the PCR sample, and any differences were used to confirm the presence of specific mutations. Figure 4-8A shows the analysis of Cas9 activity in HEK293 cells. The image in Figure 4-8B serves as a reference for the results obtained from the T7 analysis, which confirms the presence of specific mutations in the CRISPR-edited cells in comparison to wild-type cells.

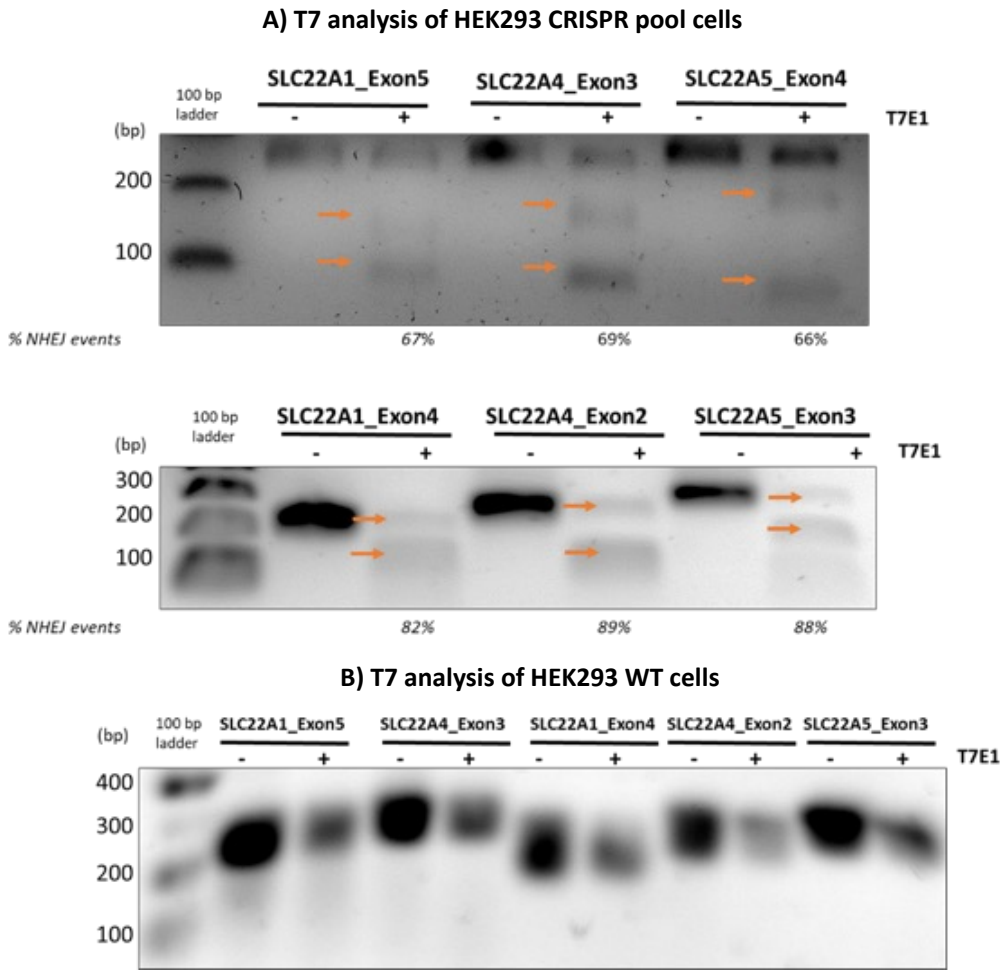


Figure 4-8. T7 endonuclease 1(T7E1) assay of HEK293 cells on a 1.5% agarose gel. PCR was performed using the oligos designed for genotyping, according to the target region. A molecular weight marker is located on the left side of the gel. A) T7 assay .in HEK293 CRISPR cells transfected with pX459 plasmids, cells were selected with puromycin and genomic PCR products was analysed by T7E1. B) T7 assay in HEK293 WT cells.

According to Figure 4-8A, the results of the assay indicated the presence of mismatched DNA in the SLC22A1_Exon5 target sequence, as smaller size products of approximately 90 and 150bp were observed. Similarly, SLC22A4_Exon3 showed smaller size products of approximately 70 and 180bp after digestion with T7E1, indicating a successful editing on target gene. Furthermore, SLC22A5_Exon4 also demonstrated a successful gene editing of the gene, as smaller size products of approximately 200 and 70bp were observed

after digestion with T7E1. Analysis of CRISPR activity in SLC22A1_Exon4 was performed using T7E1 assays and resulted in the identification of two distinct fragments with approximate sizes of 200bp and 150bp. This observation suggests the presence of mismatches and the successful operation of CRISPR in the target sequence. In a similar analysis, T7E1 digestion of SLC22A4_Exon2 revealed smaller size fragments, with approximate sizes of 250bp and 150bp. The presence of a smaller fragment alongside the full-length PCR product suggests the existence of a population of edited cells harbouring mismatches within the targeted region. Whereas analysis of WT cells as control showed no difference after digestion with T7E1.

Amplicons of the targeted regions were generated and subsequently sent for sequencing to confirm the presence of specific mutations introduced by CRISPR editing. The sequencing results from the transfected HEK 293 cells with pX459 plasmids were analysed to determine the effect of CRISPR-Cas9 editing on the gene. Figure 4-9 illustrates the sequencing results for the SLC22A1 CRISPR pool clones targeted at Exon5, Exon4 and wild-type HEK293 cells. The wild type sequence serves as the reference and is shown on the top of the alignment. The CRISPR-edited clone sequence is shown below, aligned with the wild type sequence at corresponding positions. The results of the sequence alignment for SLC22A1 exon 5 and exon 4 are shown in Figure 4-9B, respectively. Sequencing of this pool of amplicons provided a single, dominant sequence for each targeted region. This dominant sequence likely represents the most prevalent edited allele within the analysed pool.

It is important to acknowledge that HEK293 cells exhibit inherent genetic variability, including a reported modal chromosome number of 64 with a range spanning from 56 to 78 (Bylund et al., 2004; Stavropoulou et al., 2005; Stepanenko & Dmitrenko, 2015). This variability is further evidenced by documented fluctuations in the copy number of specific chromosomes, including chromosome 6, where SLC22A1 is located (Binz et al., 2019).These observations highlight the potential influence of such inherent genetic heterogeneity on the functional consequences observed following CRISPR editing in HEK293 cells.

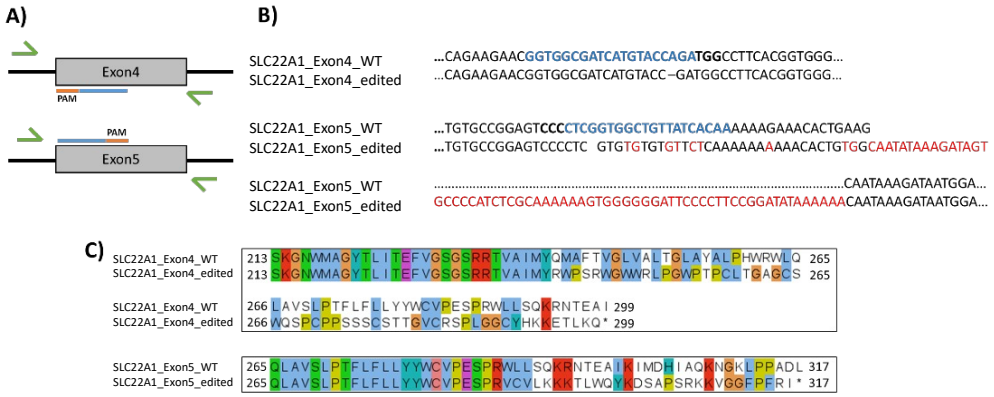


*Figure 4-9. Sequencing results from the transfected HEK293 with pX459 plasmids targeting SLC22A1 in Exon4 and Exon5. DNA sequence alignment compares the wild type sequence with the sequence from a CRISPR-edited pool. Blue line highlights the position of the sgRNA, orange line represents the PAM site. A) Representation of targeted sequence in SLC22A1, B) Sequence alignment for SLC22A1 Exon4 and Exon5, Blue font highlights the sgRNA sequence, bold letters highlight the PAM site C) Predicted Aminoacid Alignment of wild-type and CRISPR clones, *: stop codon*

In the sequence alignment of SLC22A1 exon 4 (Figure 4-9B), a single base pair deletion was observed in the CRIPSR-edited clone, compared to the wild-type sequence. This deletion resulted in a frameshift mutation in the coding region and altered the amino acid sequence of the protein, as shown in Figure 4-9C.

The blue line in the Figure 4-9A highlights the position of the sgRNA, which was designed to target the exon 4 region of SLC22A1.

Similarly, in the sequence alignment of SLC22A1 exon 5 (Figure 4-9B), point-mutations can be seen within the sgRNA sequence, followed by the insertion of 64 base pairs was also observed in the CRISPR-edited clone. This resulted in a frameshift mutation and altered the amino acid sequence of the protein, as shown in Figure 4-9C.

Overall, the sequencing results showed that the CRISPR-Cas9 system effectively edited the SLC22A1 gene, resulting in frameshift mutations in exons 4 and 5. These mutations had a significant effect on the amino acid sequence of the protein, which may affect its function and contribute to the phenotype of the CRISPR-edited cells.

Figure 4-10 presents the prediction of protein structures of SLC22A1 obtained from the AlphaFold database, red spheres in each panel indicate the stop codon location according to the multi-aligned amino acid sequence.

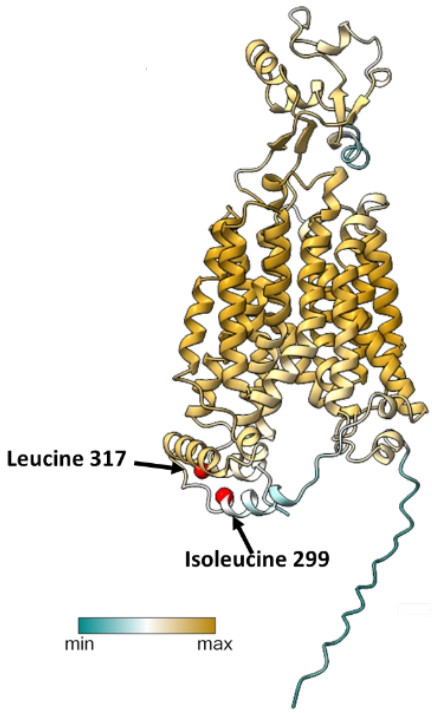


Figure 4-10. Predicted protein structures of SLC22A1 obtained from AlphaFold. Red spheres show the stop codon location according to the multiple amino acid sequence alignment.

Figure 4-11 displays the sequencing results from the transfected HEK 293 cells with pX459 plasmids targeting SLC22A4. The sequencing results were compared to the wild-type sequence to determine the effect of CRISPR-Cas9 editing on the gene. Figure 4-11B shows the DNA sequence alignment for SLC22A4 exon2 and exon3. The sequencing results indicate that the CRISPR-Cas9 system was able to introduce a point mutation in the SLC22A4 gene at the exon2. While the Exon3 showed modification of 4 nucleotides within the sgRNA sequence followed by mutations downstream of the location of the protospacer adjacent motif PAM site.

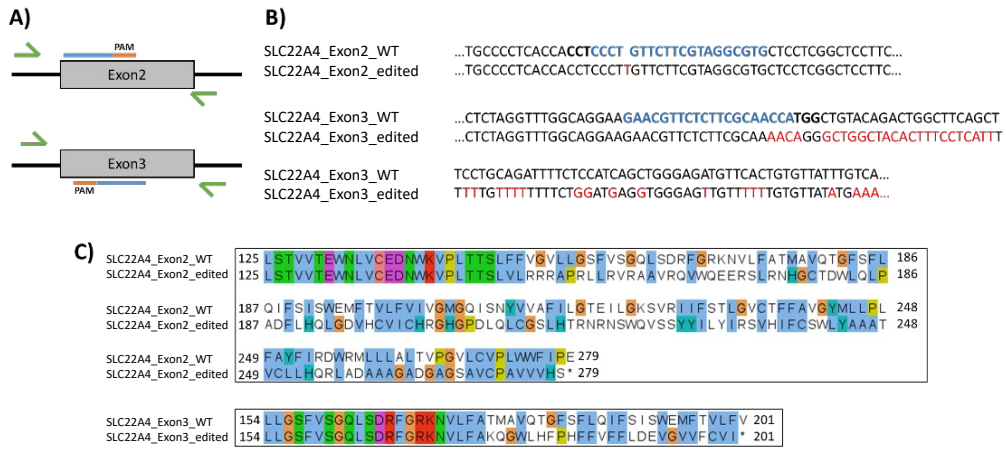


*Figure 4-11. Sequencing results from the transfected HEK293 with pX459 plasmids targeting SLC22A4 Exon2 and Exon3. DNA sequence alignment compares the wild type sequence with the sequence from a CRISPR-edited pool. Blue line highlights the position of the sgRNA, orange line represents the PAM site. A) Representation of targeted sequence in SLC22A4, B) Sequence alignment for SLC22A4 Exon2 and Exon3, Blue font highlights the sgRNA sequence, bold letters highlight the PAM site C) Predicted amino acid Alignment of wild-type and CRISPR clones, *: stop codon*

The amino acid alignment for SLC22A4 exon3 and exon4 is shown in Figure 4-11C. The alignment was conducted to evaluate the functional effects of the CRISPR-Cas9 system on the target gene. The results of the amino acid alignment indicated that the CRISPR-Cas9 system introduced a stop codon at position of aminoacid 279 for the target site in Exon2, a stop codon at position 201 for the target in Exon3.

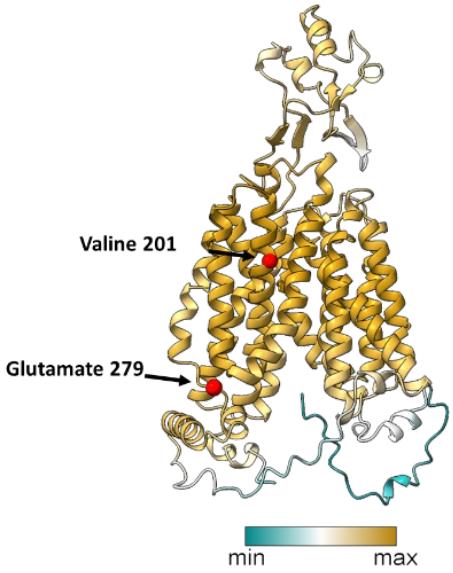


Figure 4-12. Predicted protein structures of SLC22A4 obtained from AlphaFold database. Red spheres show the stop codon location according to the multi-aligned amino acid sequence. – CHANGE STRUCTURE TO PLDDT AND ADD NUMBERS TO THE Transmembrane domains

CRISPR-Cas9 editing was employed to target SLC22A4 in HEK293 cells. The introduced stop codons by the Cas9 nuclease are predicted to result in premature termination of translation. This would lead to the production of truncated SLC22A4 protein products compared to the wild-type protein.

The prediction of the protein structure of SLC22A4 was obtained from the AlphaFold database. Figure 4-12 depicts the predicted structure of the wild-type SLC22A4 protein obtained from the AlphaFold database. The positions of the introduced stop codons within the wild-type structure are highlighted in red spheres. While the precise functional consequences of these truncations remain to be determined, the structural modelling suggests that they could significantly alter the overall protein conformation.

Sequencing results from the transfected HEK 293 cells with pX459 plasmids targeting SLC22A5 exons 3 and 4 are presented in Figure 4-13. The DNA sequence alignment shown in panel B compares the wild-type sequence with the sequence from a CRISPR-edited pool. Where deletions were found within the sgRNA target site for both targeted exons, individually. The blue line in panel A highlights the position of the sgRNA, while the orange line indicates the PAM site.

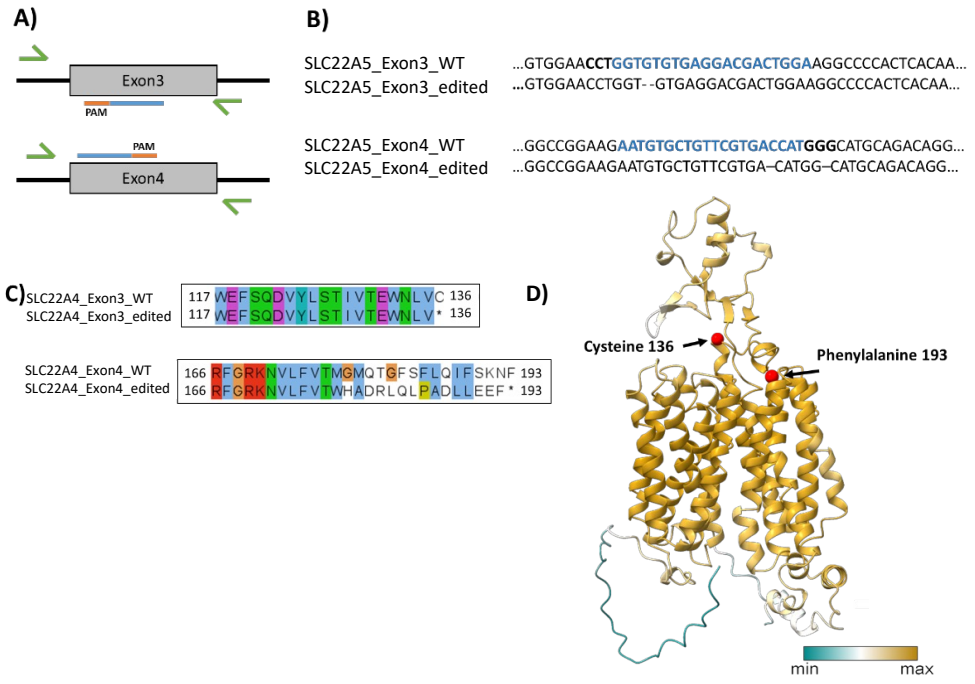


*Figure 4-13. Sequencing results from the transfected HEK293 with pX459 plasmids targeting SLC22A5 Exon3 and Exon34 DNA sequence alignment compares the wild type sequence with the sequence from a CRISPR-edited pool. Blue line highlights the position of the sgRNA, orange line represents the PAM site. A) Representation of targeted sequence in SLC22A5, B) Sequence alignment for SLC22A5 Exon3 and Exon4, Blue font highlights the sgRNA sequence, bold letters highlight the PAM site C) Predicted Amino acid Alignment of wild-type and CRISPR clones, *: stop codon D) Predicted protein structures of SLC22A5 obtained from the AlphaFold database. Red spheres show the stop codon location according to the amino acid sequence.*

CRISPR-Cas9 editing successfully introduced stop codons within exon3 and exon4 of the SLC22A5, which is indicated by the asterisk (Figure 4-13C). A stop codon at position 136 in Exon3, or a stop codon at position 193 in Exon4. These stop codons are predicted to prematurely terminate protein translation, potentially resulting in truncated protein products.

To gain further insight into the potential consequences of these edits, the predicted structure of the wild-type SLC22A4 protein was obtained from the AlphaFold database (Figure 4-13D). The figure depicts the wild-type protein

structure with red spheres highlighting the locations where the CRISPR edits would introduce stop codons. While protein functionality cannot be definitively confirmed without experimental validation, this structural analysis suggests that the stop codons introduced by CRISPR editing could significantly alter the overall structure of the SLC22A5 protein compared to the wild-type form. These structural changes may potentially affect protein function.

CRISPR-Cas9 editing was successfully employed to target SLC22A1, SLC22A4, and SLC22A5 genes in HEK293 cells. This approach resulted in the generation of cell lines harbouring edited alleles for each transporter. Enabling further functional studies to investigate the roles of SLC22A transporters in HEK293 cells. Subsequent sections will detail the utilization of these cell lines to explore the functional consequences of SLC22A disruption on various cellular processes.

4.2.4. Transfection of Calu-3 and A549 cells

Calu-3 cells were transfected with pX459 plasmids and cells were treated with 1.5 µg/ml of puromycin for selection. Transfection was carried out with the 4D nucleofector, as nucleofection is supposed to have a better yield than chemical or lipid-based transfections. Cells did not survive after transfection and selection with puromycin.

A549 cells are an additional suitable lung cell model that were utilized in this study. Transfection of these cells with pX459 plasmids was performed using lipofectamine and the cells were then selected using a puromycin concentration of 0.5 µg/ml. The results of this selection process showed that there was no cell

survival following exposure to puromycin. Figure 4-14 shows results from the T7 analysis of A549 cells under 2 days of puromycin selection. The assay showed no disruption of the gene as no band were visible on the T7 lanes.

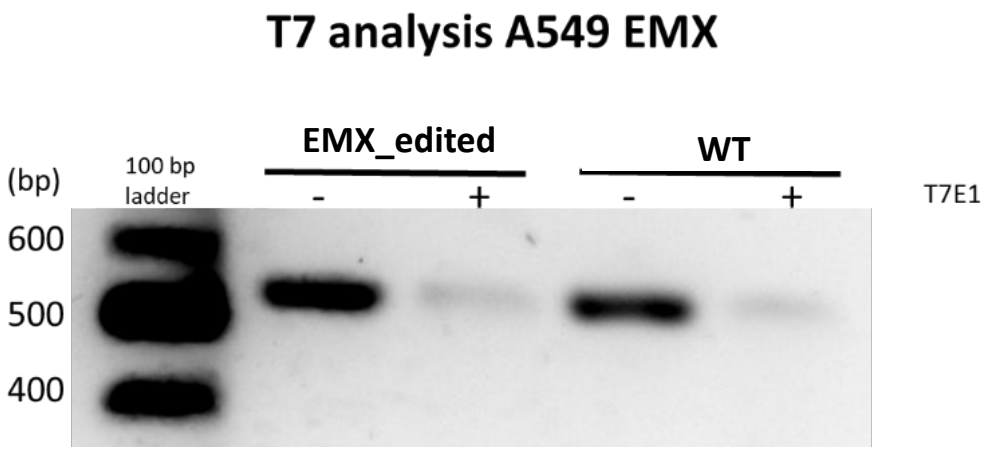


Figure 4-14. T7 Analysis. A549 cells were transfected with pX459 plasmid targeting EMX1_Exon3 by Lipofectamine 3000. Untransfected cells (WT) were used as controls. Genomic PCR was prepared to amplify the region flanking the CRISPR site (538 bp amplicon). T7 assay was carried out to identify CRISPR/Cas9 induced mutations. Left lane: DNA ladder (100 bp)

In the experimentation process, the cells were transfected and assessed for the efficiency of the CRISPR activity. Based on the results obtained from Figure 4-14, it was concluded that a different approach was required in order to achieve complete allele editing. The decision to utilize a different strategy, involving the use of nickase CRISPR with GFP as a selection marker, was based on the results obtained from Figure 4-14.The rationale behind this choice was to avoid the additional stress on the cells that could be caused by the use of selection with antibiotics.

4.2.5. Transfection of Ribonucleoprotein complexes

Direct delivery of CRISPR/Cas9 system as a ribonucleoprotein (RNP) complex has emerged as a powerful method for genome editing. This can accomplish a higher efficiency, by limiting the potential for off-target effects, when compared to traditional Cas9 approaches (DeWitt et al., 2017; S. Zhang et al., 2021). RNP complexes are active immediately, alleviating protein expression, transfection via nucleofection protocol to yield higher transfection efficiency, and avoid stress due to lipofectamine (Fiszer-Kierzkowska et al., 2011).

Due to its advantages, RNP complexes were prepared using sgRNAs from the respective pX459 plasmid as template. Calu-3 and A549 cells were transfected with RNP complexes by nucleofection.

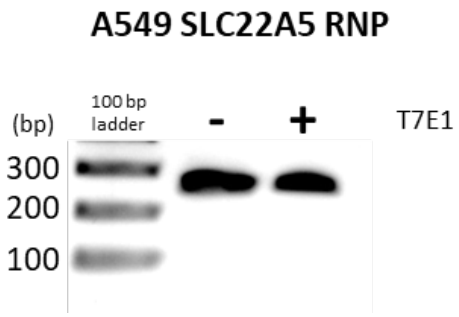


Figure 4-15. T7 endonuclease I assay in transfected A549 cells with RNP complexes. gDNA was purified 2 days after transfection and fragment was amplified for determination of gene editing. PCR lane shows the band amplified after while T7 shows the fragment after incubation with T7E1 enzyme. Left lane: DNA ladder (100 bp)

Transfected A549 with RNP complexes, cells were allowed to recover for 48 h and gene disruption effect was analysed with T7 assay. Figure 4-15 shows the results from the genotyping of SLC22A5 transfection in A549 cells. T7 endonuclease cleaves the DNA after recognition of a mismatch sequence;

however, amplified products showed no activity of gene edition, as no cleavage was observed.

4.2.6. CRISPR-nCas9 generation

For CRISPR nickase, guides were designed with PAM-out conformation, as previous studies have concluded that this orientation yields a higher rate of indel formations when used with D10A nickase (Amo et al., 2022; Kim et al., 2022; Schubert et al., 2021). This approach allows for the simultaneous targeting of both strands of the targeted gene, using GFP selection. This double-nicking strategy is a more efficient way of targeting the gene and results in successful gene disruption with lower off-target effect. The results of the sequencing of the pX461 plasmids with their corresponding sgRNA are illustrated in Figure 4-16.

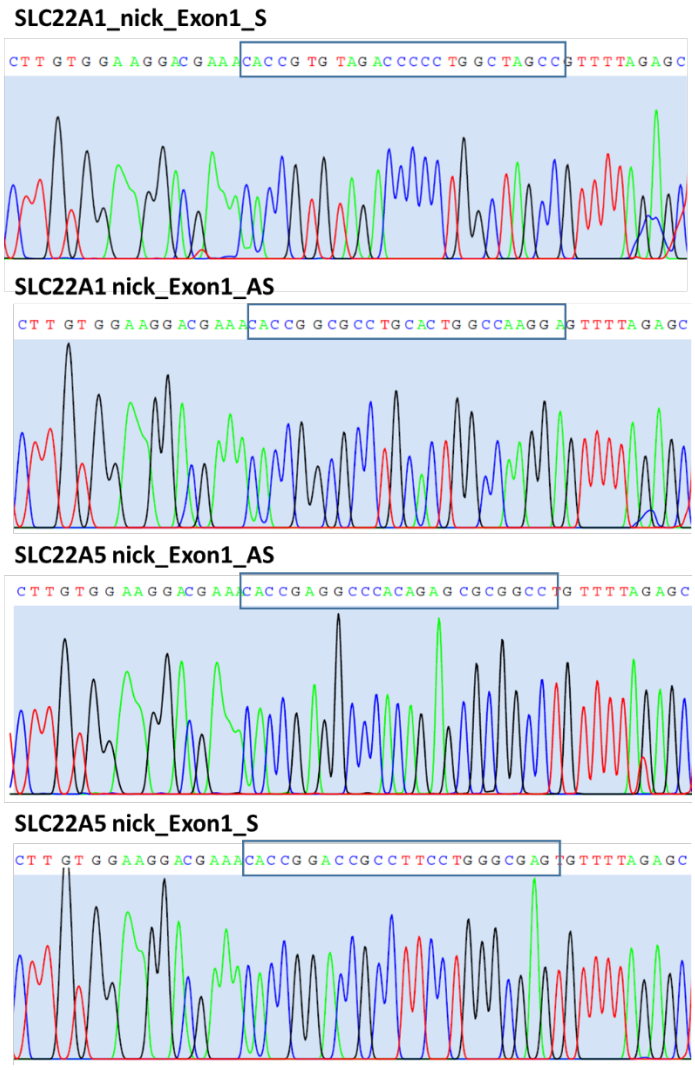


Figure 4-16. Chromatogram Sanger sequencing for pX461 Cas9n plasmids. Plasmids were sequenced with the U6-FWD primer. Blue boxes highlight the sequence of the respective sgRNA inserted, respectively.

Human embryonic kidney (HEK293) cells were transfected with a combination of plasmids at a 1:1 ratio, to ensure efficient transfection, HEK293 cells were transfected with lipofectamine 3000. After a recovery period of two days, genomic DNA was harvested for T7 endonuclease I assays to evaluate the efficiency of gene targeting. Gene editing was detected using the T7 assay and DNA sequencing, as seen in Figure 4-20.

4.2.7. CRISPR-nCas9 transfection in lung cells

After validation of Cas9 nickase activity in HEK293, Calu-3 cells were transfected using nucleofection. A ratio of 1:1 of the respective sense (S) and antisense (AS) Cas9n plasmid was used for the transfection of Calu-3 cells. After the cells were allowed to recover for two days, GFP expression was not observed, suggesting unsuccessful transfection.

A549 cells were then transfected using lipofectamine 3000 with pX461 plasmids targeting SLC22A1_nick_Exon1 and SLC22A5_nick_Exon1, conducted separately. After a 48-hour recovery period, GFP expression was assessed using flow cytometry. Forward scatter (FSC) and side scatter (SSC) were first employed to exclude doublets and dead/damaged cells. A specific laser wavelength (488 nm) then excited GFP molecules, and fluorescence intensity (FL1) was measured at 530 nm to identify GFP-positive cells. GFP expression was quantified based on the GFP-Area on a logarithmic scale (GFP-Area-Log) to improve sensitivity. Cells with a GFP-Area-Log value exceeding a threshold of 10^2 were considered GFP-positive. This approach ensured isolation of a population with robust GFP expression, as illustrated in Figure 4-17., were applied based on this gating strategy. The analysis revealed that approximately 1.03% of transfected A549 cells (102,301 cells) expressed GFP protein. This confirms successful expression of the GFP reporter in the transfected cells, which were then sorted into a 96-well plate for further expansion.

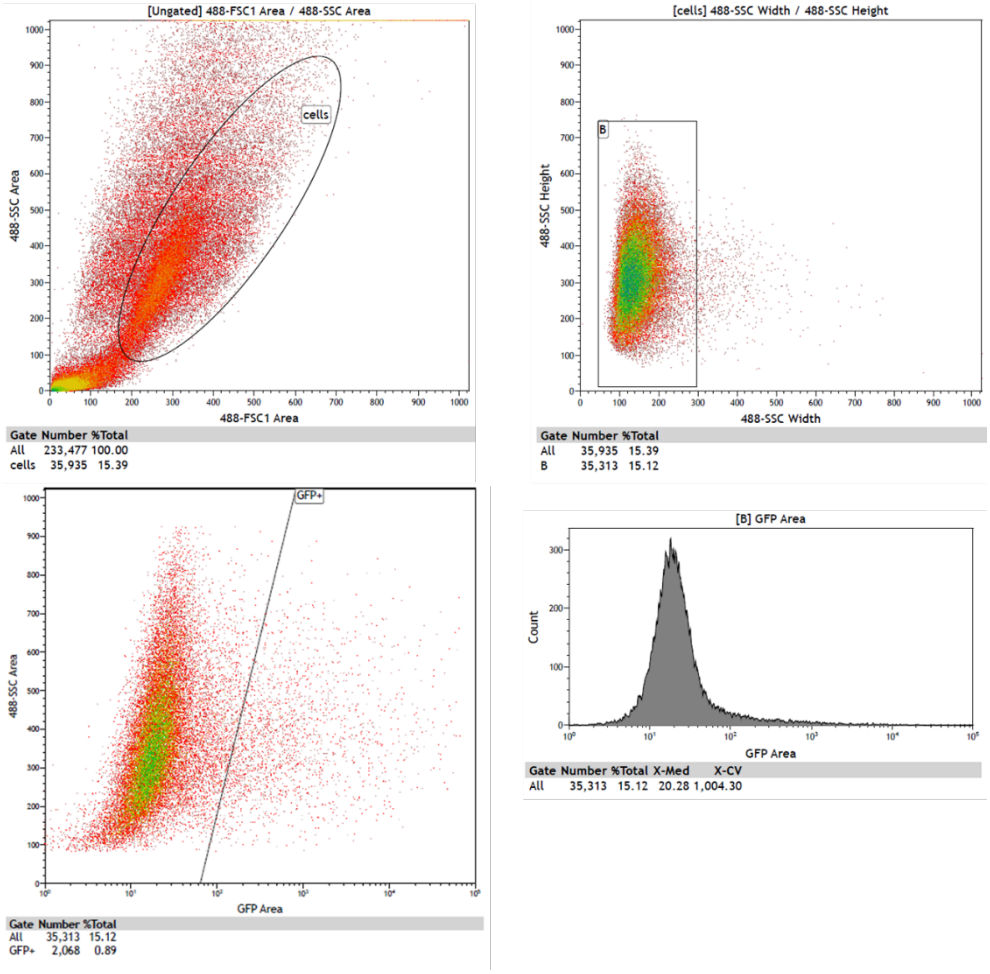


Figure 4-17. Cell sorting Transfected A549 cells with pX461_SLC22A1_nick_Exon1 by lipofectamine 3000 were analysed on a Beckman Coulter Astrios EQ. GFP was detected on the argon laser 530/30nm channel. Cells were seeded at 1 cell per well in a 96-well plate.

Following the transfection and expansion of colonies, further analysis was conducted to confirm the presence of gene disruption using the T7E1 assay. Figure 4-19 presents the results of the agarose gel analysis of 25 cell colonies that were transfected with the CRISPR system and further expanded. The right-most lane in the figure represents the wild-type (WT) or untransfected cells, which served as a control group. All amplicons were subjected to T7E1 analysis to assess the CRISPR activity.

Similarly, A549 cells were transfected with Lipofectamine 3000 and allowed to recover for 48 hours. Flow cytometry evaluation and sorting of GFP expression in A549 cells transfected with pX461_SLC22A1_nick_Exon1 plasmids showed 0.89% of cells expressing GFP, as depicted in Figure 4-18. These results indicate that the GFP protein was successfully expressed in the transfected cells and they were subsequently sorted into a 96-well plate for expansion.

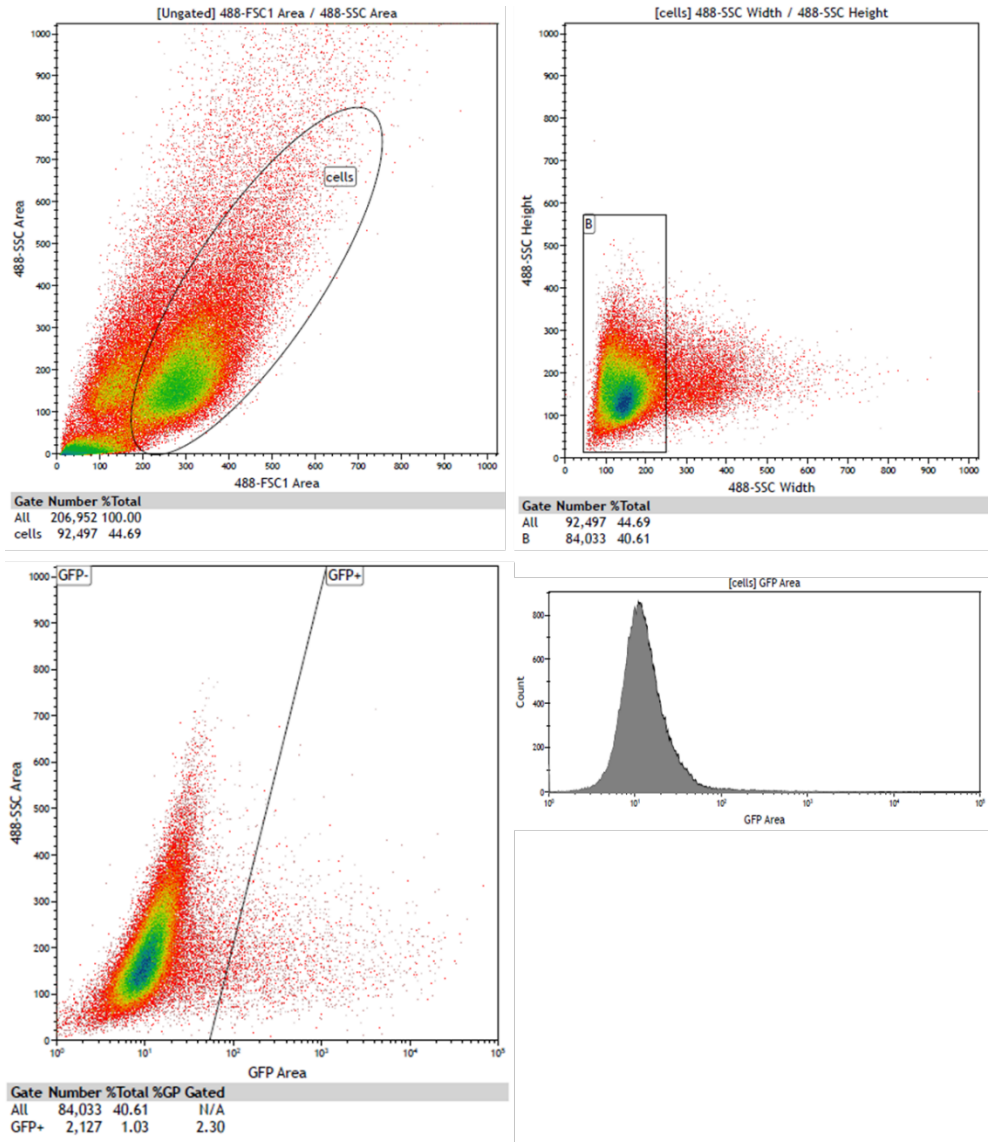


Figure 4-18. Cell sorting. Transfected A549 cells with pX461_SLC22A5_nick_Exon1 by lipofectamine 3000 were analysed on a Beckman Coulter Astrios EQ. GFP was detected on the argon laser 530/30nm channel. Cells were seeded at 1 cell per well in a 96-well plate.

Results show no notable difference between the WT and transfected cells. This suggests that the CRISPR system failed to disrupt the target gene, and that the T7E1 assay did not detect any cleavage products in the amplified sequences.

Flow cytometry evaluation and sorting of GFP expression in A549 cells transfected with pX461_SLC22A5_nick_Exon1 plasmids showed 10.7% of cells expressing GFP. These results indicate that the GFP protein was

successfully expressed in the transfected cells, and were subsequently sorted into a 96-well plate for expansion. Figure 4-18 presents the analysis of transfected A549 GFP+ cells. Cells were seeded in a 96-well plate with a single cell per well, and the GFP signal was detected using the argon laser 530/30nm channel.

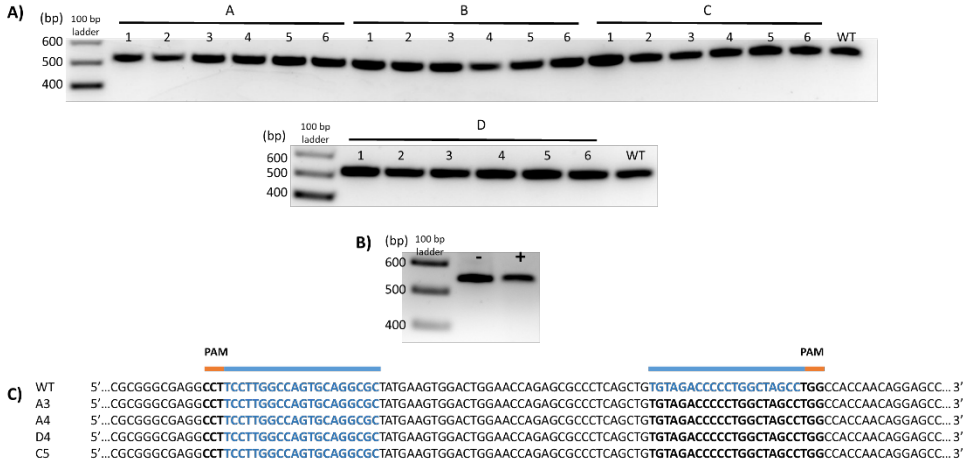


Figure 4-19. T7 analysis for sorted A549 cells transfected with SLC22A1_nick_Exon1 by Lipofectamine 3000. Clones were allowed to form colonies. Untransfected cells (WT) were used as controls. Genomic PCR was prepared to amplify the region flanking the CRISPR site (538 bp amplicon). T7 assay was carried out to identify CRISPR/Cas9 induced mutations. Left lane: DNA ladder (100 bp)

DNA was sent for sequencing and analysed for indels. Figure 4-19 shows the sequencing results and alignment for sequence comparison. Results showed that the CRISPR nickase samples exhibited the same sequence as the wild-type (WT) samples on SLC22A1 and SLC22A5. This is illustrated in Figure 4-19, where the comparison of the sequencing results between the A549 SLC22A1_nick_Exon1 CRISPR and A549 WT samples is shown. The data indicates that the CRISPR nickase treatment did not result in any significant changes in the sequence of the samples, as indicated by the high degree of sequence identity between the two groups.

4.2.8. Expression of SLC22A1, SLC22A4, SLC22A5 in edited HEK293 cells

To evaluate the expression levels of Solute Carrier 22A1 (SLC22A1), Solute Carrier 22A4 (SLC22A4) and Solute Carrier 22A5 (SLC22A5) in kidney cells (HEK293), quantitative reverse transcriptase polymerase chain reaction (qPCR) was performed. Oligo specificity was evaluated, and results were shown in Methods section.

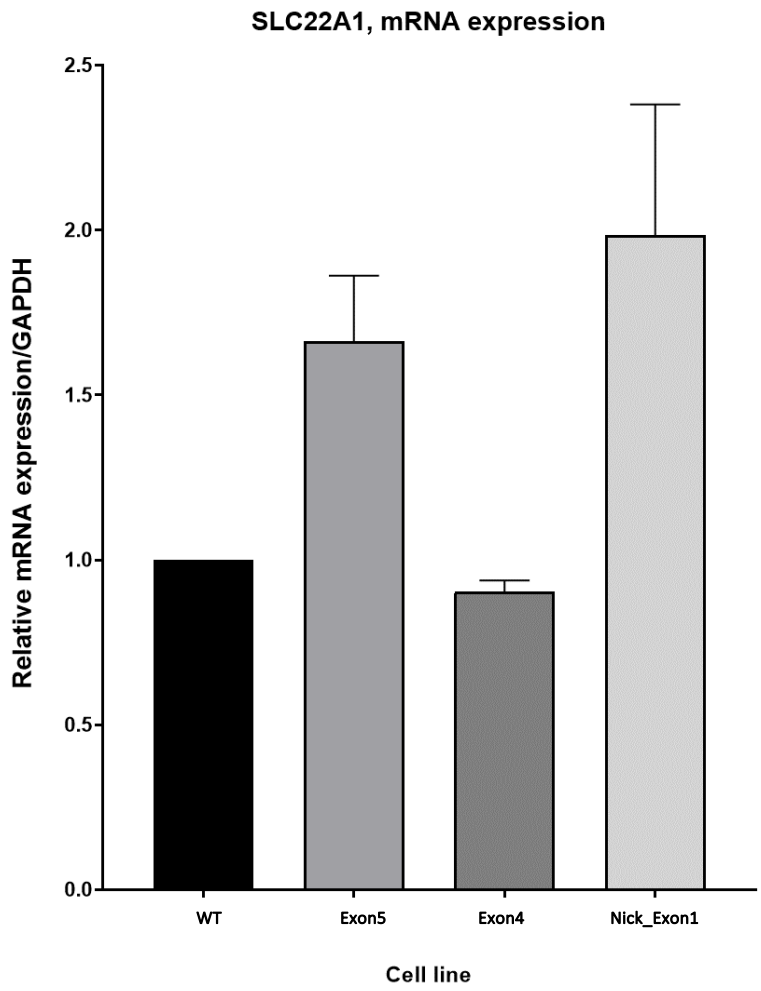


Figure 4-20. mRNA expression of SLC22A1 in HEK293 WT and HEK293 CRISPR edited cell lines. cDNA was prepared with 200 ng of total RNA. Expression was normalised with individual levels of GAPDH and wild type cells were used as control for ΔΔCt. Data was analysed with a two-way ANOVA and presented as Mean ± SEM.

The expression profiles of Solute Carrier 22A1 (SLC22A1) mRNA are summarised in Figure 4-20. The results obtained from the analysis of the wild-type and CRISPR-edited cell lines indicate that no significant differences in gene expression were found. However, when targeted in Exon5, an increase in expression levels was observed, with a fold change of 1.7 ± 0.2 when compared to the wild-type.

On the other hand, Exon4 showed a slight decrease in mRNA expression, with a fold change of 0.9030 ± 0.03445 when compared to the wild-type. Additionally, CRISPR nickase targeting in Exon1 exhibited an increase in expression levels with a fold change of 1.986 ± 0.3955. These results indicate that the specific targeting of different exons of SLC22A1 gene can have different effects on the expression levels of this gene.

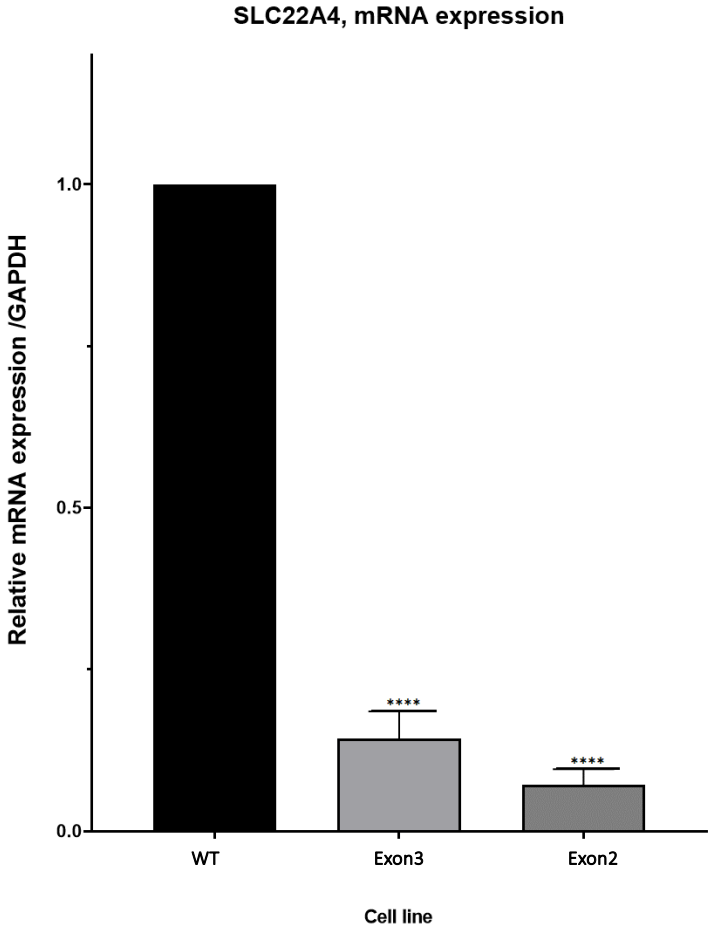


Figure 4-21. mRNA expression of SLC22A4 in HEK293 WT and HEK293 CRISPR edited cell lines. cDNa was prepared with 200 ng of total RNA. Expression was normalised with individual levels of GAPDH and wild type cells were used as control for ΔΔCt. Data was analysed with a two-way ANOVA and presented as Mean ± SEM.

Changes in Solute Carrier 22A4 (SLC22A4) mRNA levels are illustrated in Figure 4-21. A one-way analysis of variance (ANOVA) was performed to assess

the significance of the changes in gene expression. The results of the ANOVA revealed a significant reduction in gene expression when the target was located in Exon 3, with a fold change of 0.143 ± 0.042 ($p<0.001$).

Similarly, when the gene was targeted in Exon2, a significant decrease in mRNA levels was observed in comparison to the wild-type, with a fold change of 0.072 ± 0.024 ($p=0.001$). These results indicate that the specific targeting of different exons of the SLC22A4 gene can lead to significant reductions in gene expression levels.

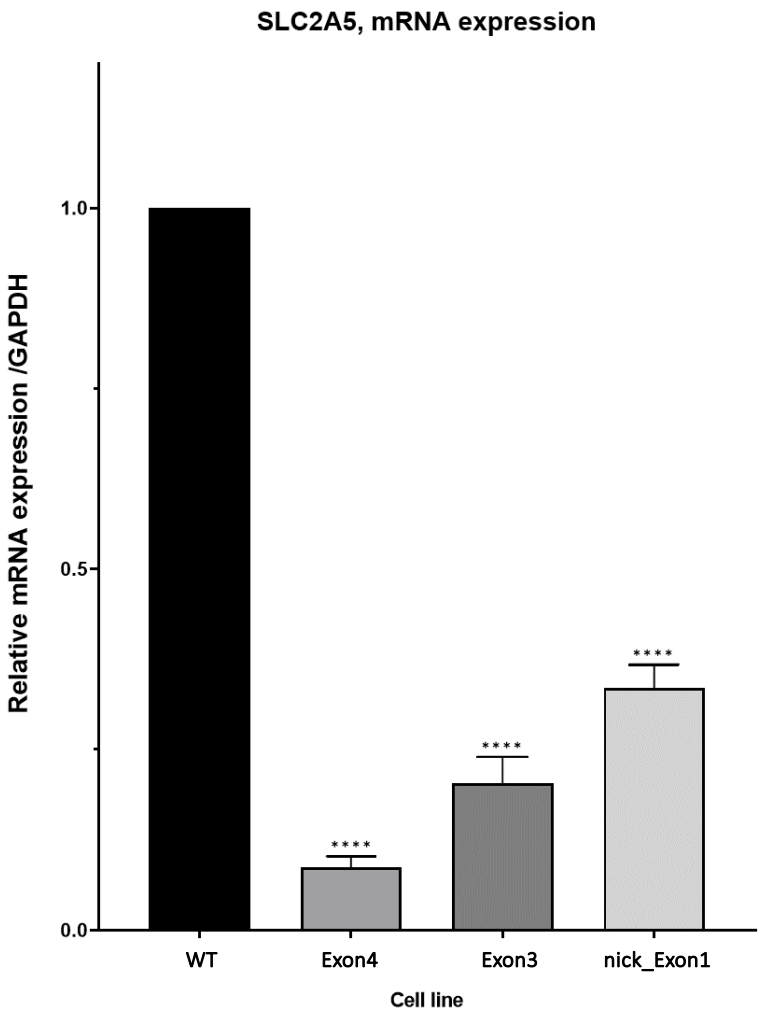


Figure 4-22. mRNA expression of SLC22A5 in HEK293 WT and HEK293 CRISPR edited cell lines. cDNa was prepared with 200 ng of total RNA. Expression was normalised with individual levels of GAPDH and wild type cells were used as control for ΔΔCt. Data was analysed with a two-way ANOVA and presented as Mean ± SEM.

Expression profiles of Solute Carrier 22A5 (SLC22A5) between wild-types and CRISPR-edited cells are shown in Figure 4-22. On average, the CRISPR-edited cells for SLC22A5 exhibited lower expression levels of mRNA compared to the wild-type cells. The target site in Exon3 was found to demonstrate a decrease in gene expression with a fold change of 0.0865 ± 0.0158 (p=0.001).

Similarly, the expression levels for the target site in Exon4 were found to be significantly lower (0.2038 ± 0.0362, p<0.001) compared to the other genetic

modifications in SLC22A5 via CRISPR. Additionally, the CRISPR nickase targeting in Exon1 showed a fold change of 0.3358 ± 0.03244 in expression levels, and this change was statistically significant ($p<0.001$).

The observed changes in mRNA expression levels following CRISPR-Cas9 editing of the target genes may be attributed, in part, to a cellular process known as nonsense-mediated decay (NMD) (Kurosaki & Maquat, 2016). NMD is a quality control mechanism that targets and degrades mRNAs harbouring premature termination codons (PTCs) introduced by frameshift mutations or insertions/deletions (indels) commonly generated during CRISPR-Cas9 editing (Sharpe & Cooper, 2017; Tuladhar et al., 2019). These indels can disrupt the reading frame, leading to the creation of a PTC upstream of the normal stop codon. NMD machinery recognizes such transcripts as aberrant and triggers their degradation, potentially explaining the observed reductions in mRNA levels. However, residual variation in expression levels could also be due to compensatory changes in gene expression (Lindeboom et al., 2016). For example, the expression level of a gene containing an NMD-triggering PTC could be partially rescued by compensatory up-regulation of the non-mutated allele.

4.2.9. mRNA expression of SLC22A1, SLC22A4 and SLC222A5 in kidney and lung cells

Quantitative polymerase chain reaction (qPCR) was performed to determine the normalized mean expression levels of SLC22A1, SLC22A4, and SLC22A5 mRNA in the kidney epithelium cell line HEK293 and the lung cell lines A549 and Calu-3. Primer specificity was confirmed using methods described in the

Methods section, including separation of amplification products on a 2% agarose gel and melting curve analysis.

The results presented in Figure 4-24 represent the normalised mean expression levels of target genes relative to the housekeeping gene GAPDH. Normalization is crucial to account for potential variations in RNA extraction, cDNA synthesis, and overall qPCR efficiency. The normalized mean expression level is calculated by dividing the target gene's Ct value by the Ct value of the reference gene (GAPDH) and then calculating the anti-logarithm (2^-ΔCt).

The normalised mean expression levels of SLC22A1 mRNA revealed no statistically significant differences between HEK293 (16.11 ± 0.253) and A549 cells (17.21 ± 0.253). However, Calu-3 cells displayed a significantly lower SLC22A1 mRNA level (14.67 ± 0.460) compared to HEK293 cells (p-value < 0.05).

Similarly, no significant differences were observed in SLC22A4 mRNA expression between HEK293 (13.466 ± 0.263), A549 (11.976 ± 0.173), and Calu-3 cells (13.27 ± 0.518).

SLC22A5 exhibited the highest mRNA expression among the three transporters analysed, with a trend towards increased levels in all cell lines compared to SLC22A1 and SLC22A4. However, these differences did not reach statistical significance. HEK293 cells showed a normalized mean expression level of 7.513 ± 0.094, while A549 and Calu-3 cells had levels of 9.146 ± 1.217 and 8.60 ± 0.532, respectively.

Based on its consistent expression of SLC22A1, SLC22A4, and SLC22A5, CRISPR edited HEK293 cells were used for subsequent functional studies.

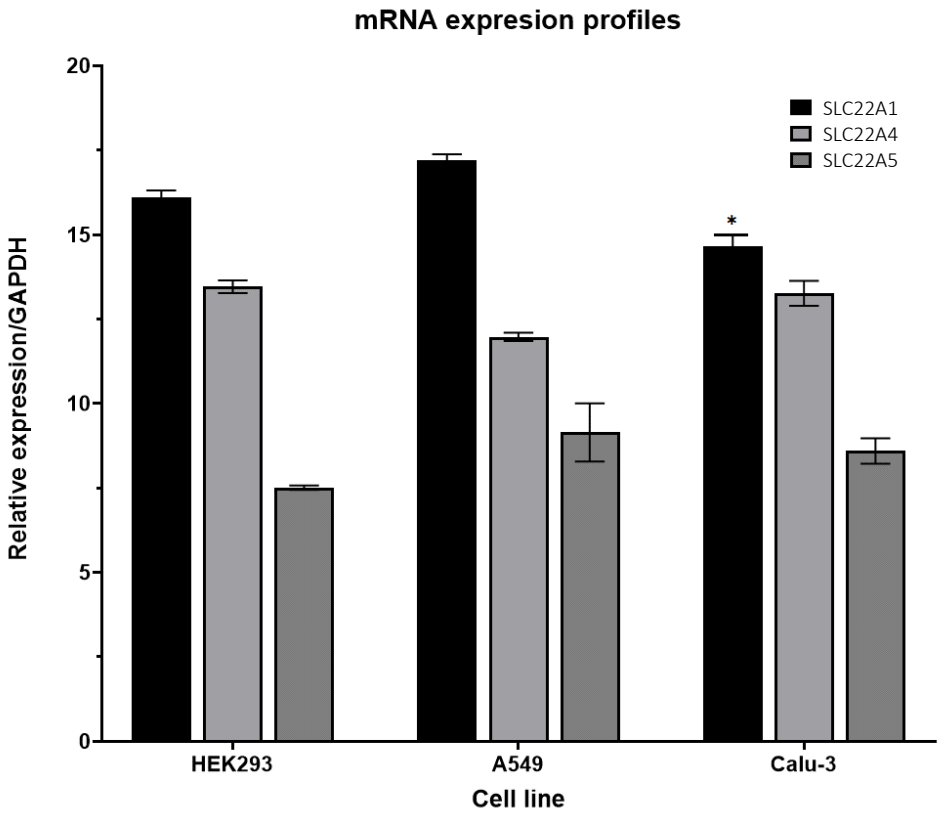


Figure 4-23. Relative expression of SLC22A1, SLC22A4 and SLC22A5 in HEK293, A549 and Calu-3. cDNA was prepared with 200 ng of total RNA. Expression was normalised with individual expression of GAPDH. Data was analysed with a two-way ANOVA and presented as Mean ± SEM.

4.2.10.Functional Analysis in gene edited HEK293

To complement the analysis of gene expression, functional consequences of the SLC22A genes were investigated in HEK293 cells. This involved a series of assays, including cell migration, cell proliferation, and cell adhesion, to elucidate how gene manipulation of OCT/Ns might affect cellular functions. The results obtained from these assays are crucial for understanding the role of these

transporters in the regulation of cell behaviour and will provide insight into the potential implications of the loss of these proteins in disease progression.

However, it is important to acknowledge that protein level determination was not feasible within the scope of this study. While attempts were made to assess protein expression using Western blotting, these attempts were unsuccessful.

Evaluation of cell proliferation in HEK293 Cells after CRISPR Targeting of SLC22A genes

The method used to assess cell proliferation in this study was the 5-ethynyl-2'deoxyuridine (EdU) incorporation assay, which utilizes the Click-iT EdU Alexa Fluor 488 cell proliferation assay kit from Thermofisher. The percentage of cells in S-phase was determined by counting EdU stained cells as a percentage of DAPI stained cells using the EVOS M5000 Imaging System and ImageJ software. At least 200 nuclei were analysed for each condition.

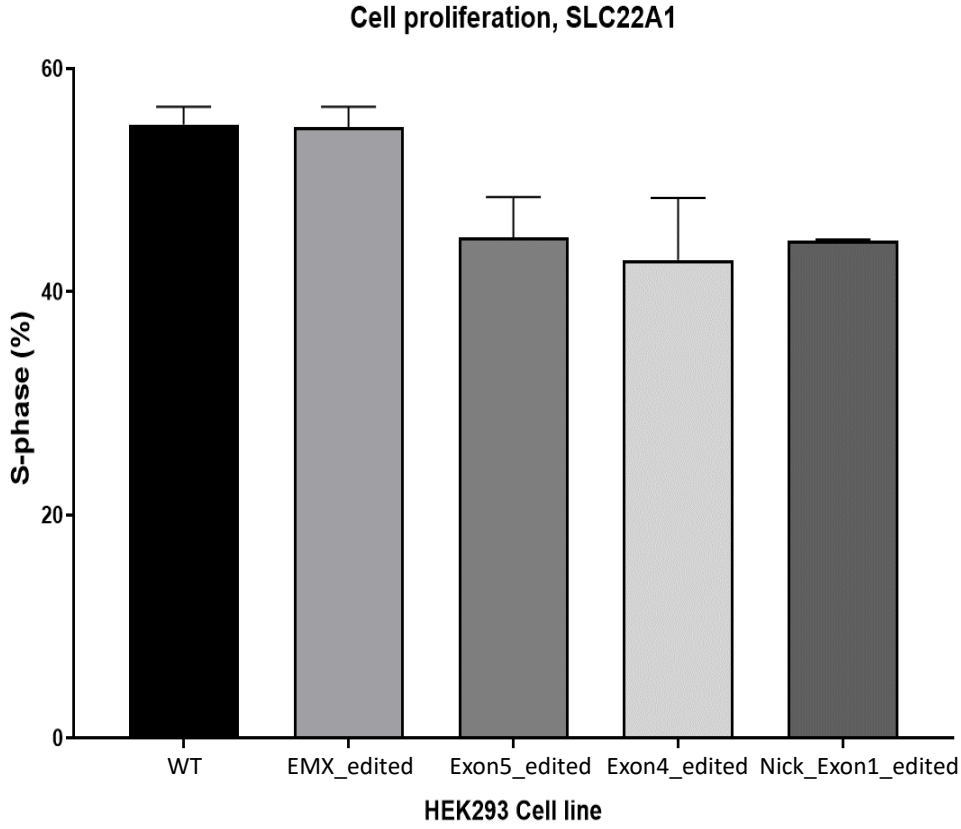


*Figure 4-24. Effect of SCLC22A1 gene editing on cell proliferation in HEK293 wild- type and edited cells. Cells were seeded at a density of 0.5 x 106 cells per well and cell proliferation was analysed with EdU. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001.*

Cell proliferation of pX459_EMX1_Exon3 edited cells were found to be unchanged when compared to the control cells. pX459_EMX1 cells and wild-type cells showed no significant difference in cell proliferation (54.77% ± 1.830% vs 55.01% ± 1.595%). Cell cycle analysis revealed that the number of cells in the S-phase decreased in the pX459_Exon5 group in comparison to the wild-type cells (44.85 % ± 3.650 vs 55.01 ± 1.595). However, genetic modification in pX459_Exon4 showed lower proliferation than control cells and the results were not significant (42.88% ± 5.565%).

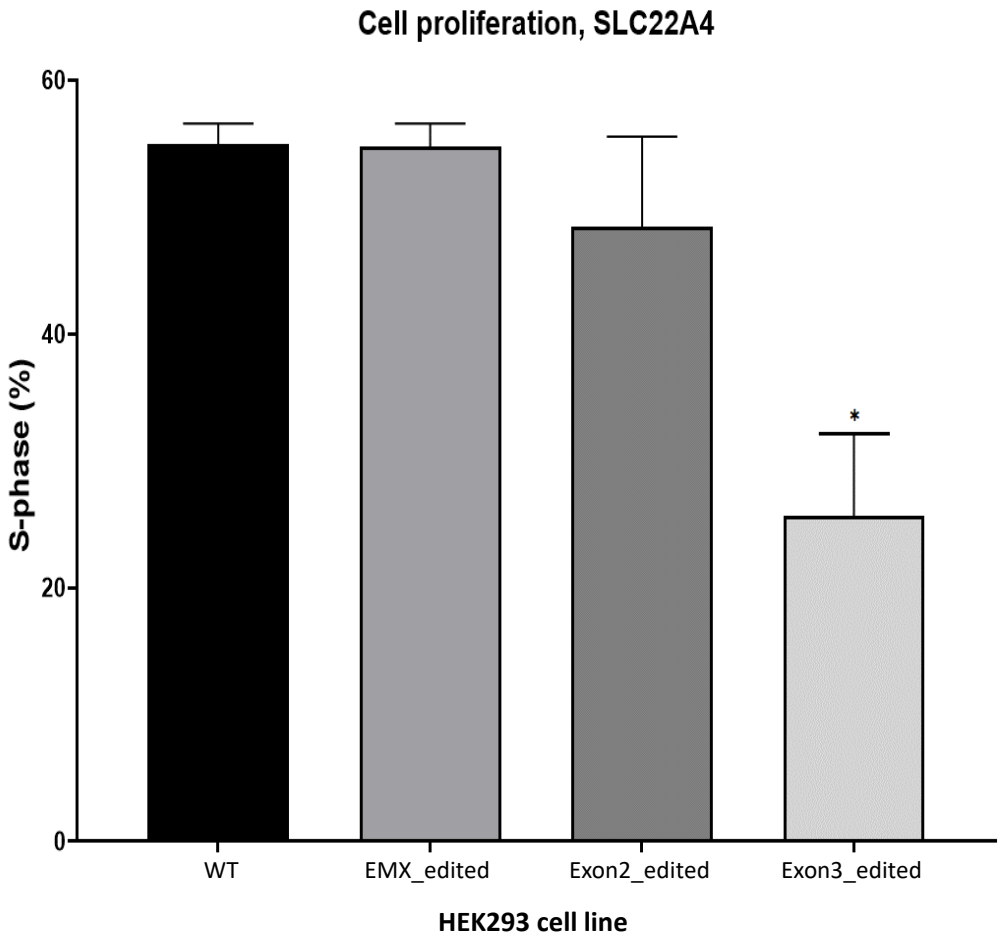


*Figure 4-25. Effect of SCLC22A4 gene editing on cell proliferation in HEK293 wild-type and edited cells. Cells were seeded at a density of 0.5 x 106 cells per well and cell proliferation was analysed with EdU. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001.*

Editing of SLC22A4 in HEK293 revealed a significant reduction ($p<0.05$) in cell proliferation in cells with a gene edition in pX459_Exon3, with values of 25.63% ± 6.51%, compared to wild-type cells, with a value of 48.48% ± 7.080, Figure 4-26.

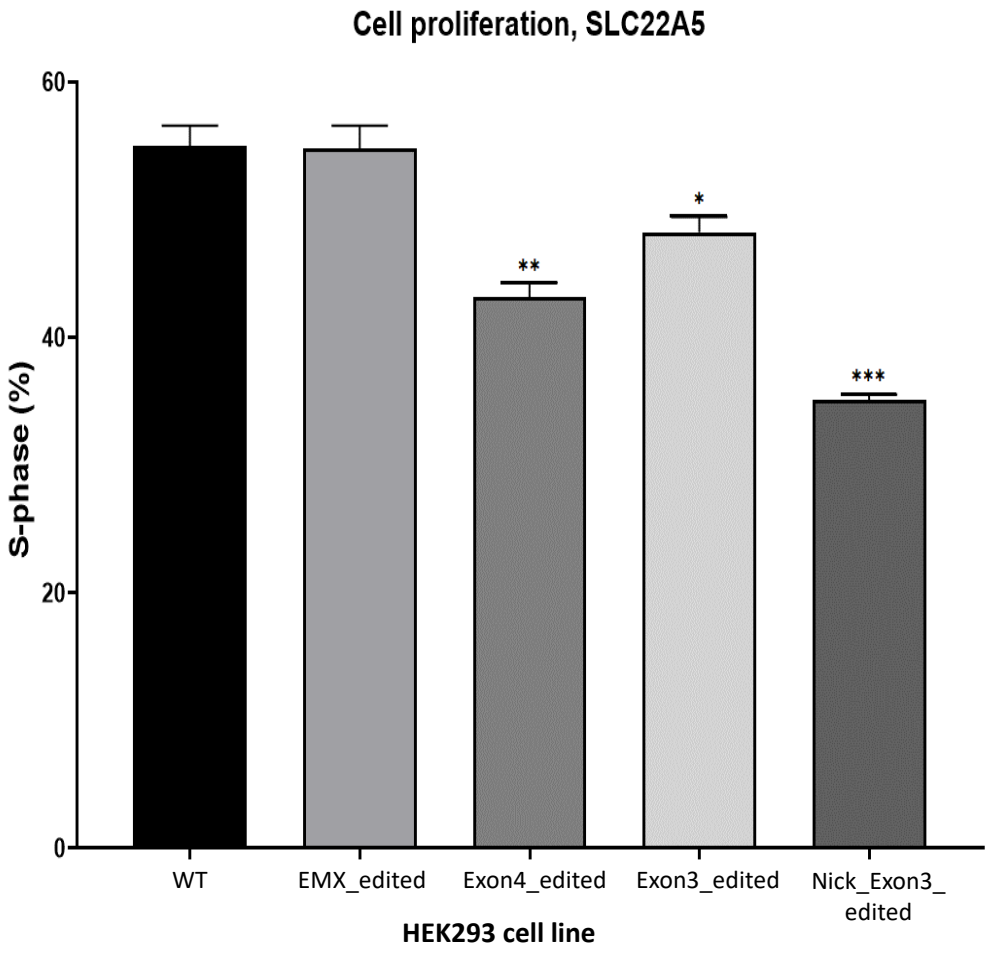


*Figure 4-26. Effect of SCLC22A5 gene edited on cell proliferation in HEK293 wild-type and edited cells. Cells were seeded at a density of 0.5 x 106 cells per well and cell proliferation was analysed with EdU. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001.*

EdU assay showed a significant decrease in all SLC22A5 edited cells, as shown in Figure 4-27. Genetic modification in Exon4 showed a significant reduction ($p<0.01$) decrease in cell proliferation compared to control cells.

Functional impact of SLC22A gene editing on cell adhesion in HEK293 Cells

Cell adhesion was determined by fixing and staining cells with a 1% toluidine blue/3% PFA solution, followed by lysis and measuring absorbance at 590 nm using a BioTek Synergy Multi-mode Microplate Reader (Agilent).

Adhesion analysis revealed no significant difference between HEK293 wild-type cells (99.2% ± 1.2%) and cells targeted with CRISPR at the EMX1 locus (93.8% ± 3.51%). This suggests that the CRISPR targeting approach itself did not significantly affect cell adhesion in this context. In contrast, HEK293 cells targeted with CRISPR at Exon 5 and Exon 4 of SLC22A1 exhibited a marked decrease in cell adhesion compared to the wild-type cells. Exon 5-targeted cells displayed a significant reduction in adhesion (38.71% ± 2.229%, $p < 0.001$), while Exon 4-targeted cells also showed a significant decrease (88.16% ± 2.767%, $p < 0.05$). These findings suggest that SLC22A1 may play a crucial role in cell adhesion, and its targeted disruption through CRISPR editing significantly impacts this cellular process.

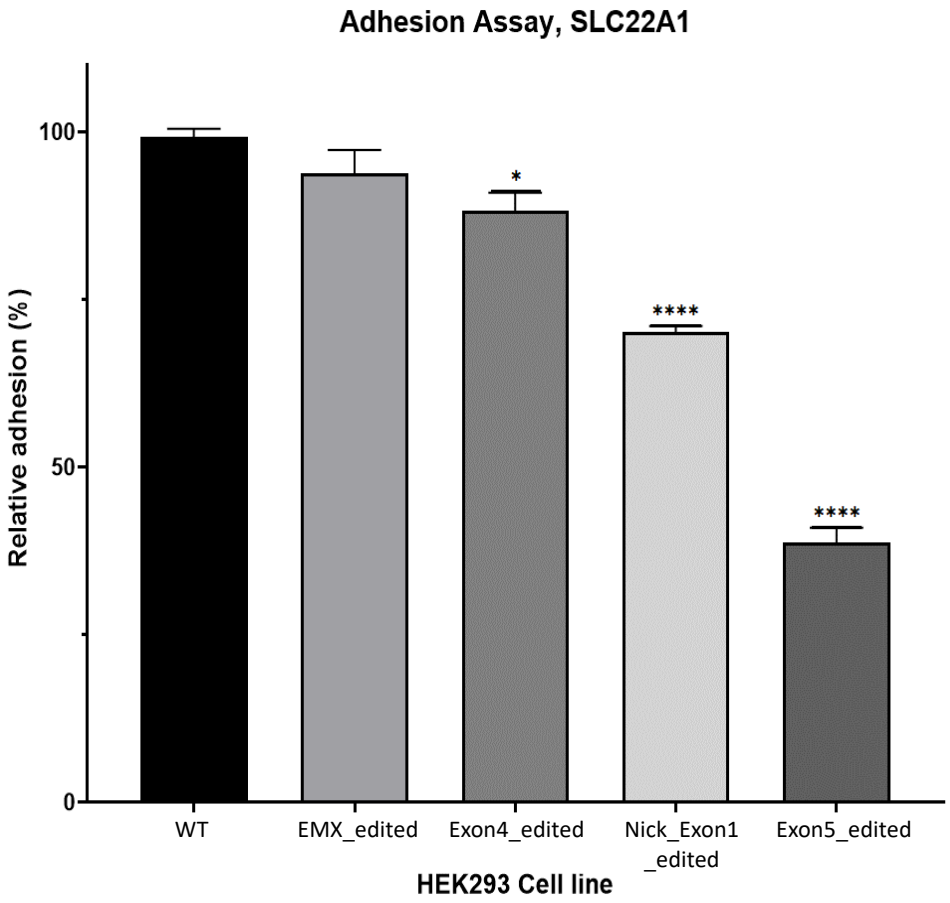


*Figure 4-27. Effect of cell adhesion in HEK293 wild-type and SLC22A1 edited cells. Cell adhesion was analysed using 1% toluidine blue/3% PFA. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001*

In addition, gene edits of SLC22A1 in HEK293 showed a significantly reduction in cell adhesion when transfected with nick_Exon1 for edits with Cas9, with 70.20% ± 0.8272% ($p<0.001$). These results indicate that edited alleles of SLC22A1 in either exons had a clear effect on cell adhesion with Exon5 showing the most significant changes.

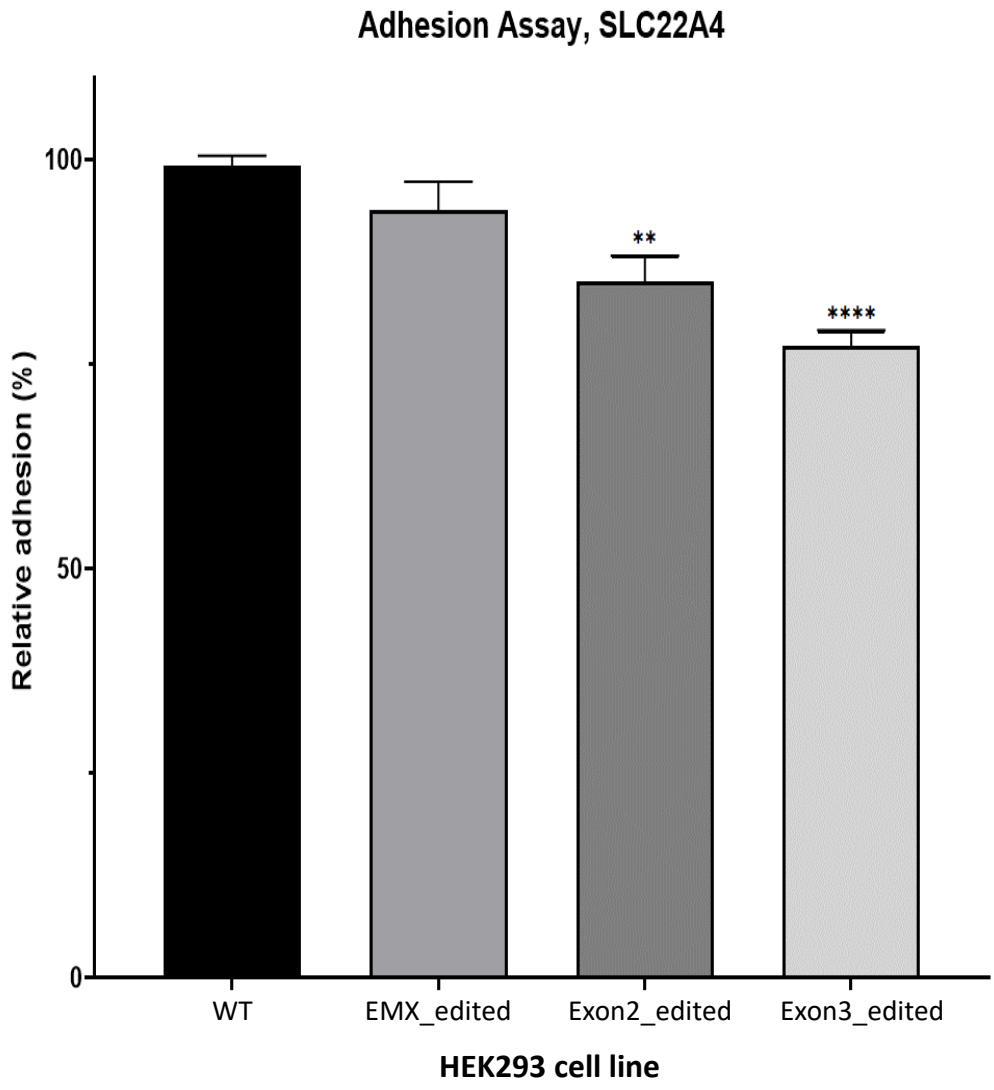


*Figure 4-28. Effect of gene edits within SLC22A4 on cell adhesion in HEK293 wild-type and edited cells. Cell adhesion was analysed using 1% toluidine blue/3% PFA. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001*

An analysis of the effect of SLC22A4 gene edits on cell proliferation and adhesion in HEK293 cells was conducted, as depicted in Figure 4-28. Results revealed a significant reduction ($p<0.001$) in cell adhesion for cells with a gene targeting in Exon3, accompanied by a decline in cell proliferation ($p<0.05$) with values of 77.15% ± 1.980% and 25.63% ± 6.510%, respectively. Additionally,

targeted mutations in Exon2 resulted in a significant reduction ($p<0.01$) in cell adhesion, with a value of 85.04% ± 3.198%.

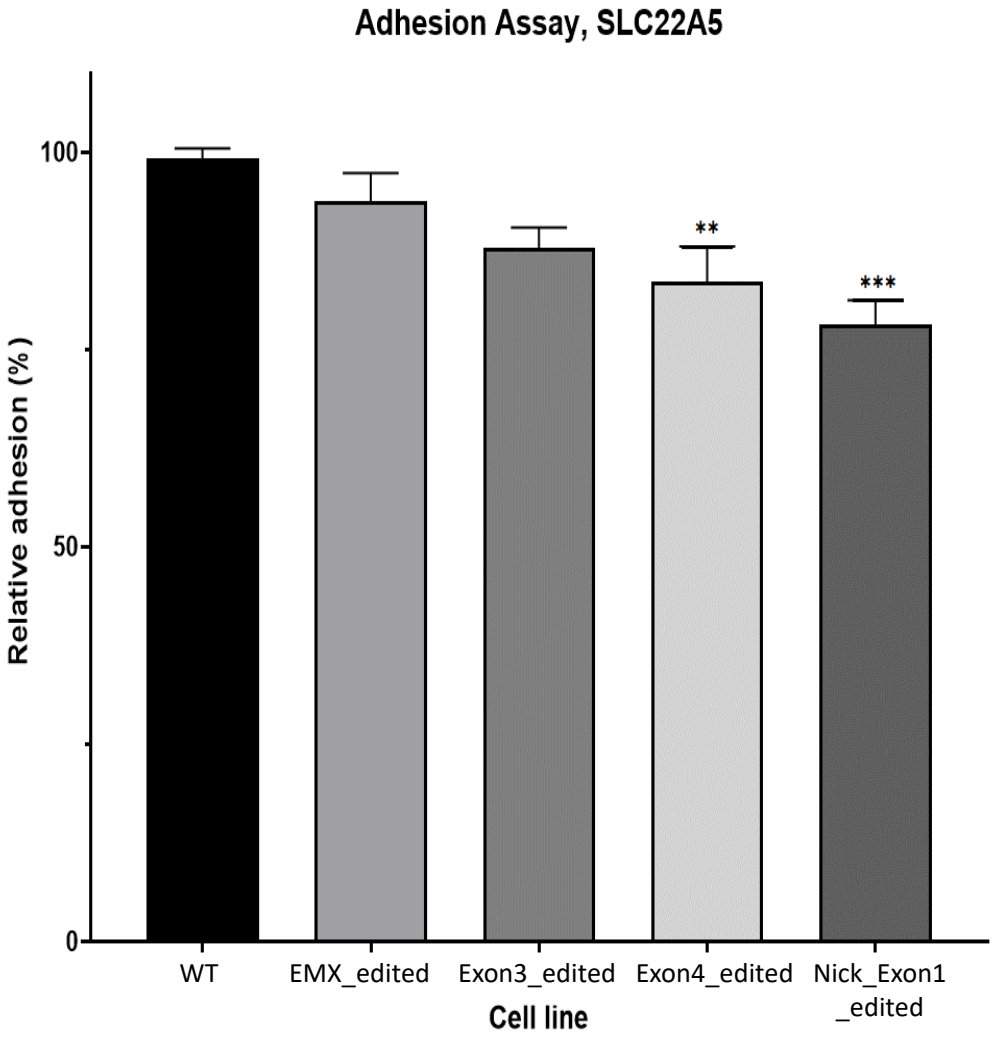


*Figure 4-29. Effect of SLC22A5 CRISPR-mediated edits on cell adhesion in HEK293 wild-type and edited cells. Cell adhesion was analysed using 1% toluidine blue/3% PFA. Data is presented as mean ± SEM. P values were calculated with Student's test, * p <0.05, ** p < 0.01, *** p< 0.001*

Toluidine assay demonstrated a significant decrease in cell adhesion in SLC22A5 edited cells, as shown in Figure 4-30. CRISPR-mediated disruption in Exon4 showed a significant reduction ($p<0.01$) in cell adhesion, compared to

control cells. Furthermore, region targeted with nickase CRISPR showed a significant decrease in cell adhesion ($p<0.001$). However, gene disruption in Exon3 with CRISPR showed no significant difference in cell adhesion.

Wound healing assay

The effect of the gene disruption of the proteins of interest on the migration ability of HEK293 cells was analysed using wound healing assay. Cells were seeded at a density of 3500 cells/well in a 2-well culture insert. Quantification of the migration ability was carried out at various time points using the ImageJ software.

Wild-type HEK293 cells were used as a control for the analysis of migration in all CRISPR cell lines. The results are presented according to the targeted gene. No significant differences were observed between the EMX1 cells and non-transfected cells. The wild-type cells demonstrated a wound closure of 44.89% ± 1.695 at 24 hours and complete wound closure at 48 hours. Similarly, the EMX1 edited cells exhibited a 41.26 % ± 2.081% closure at 24 hours and complete wound closure at 48 hours.

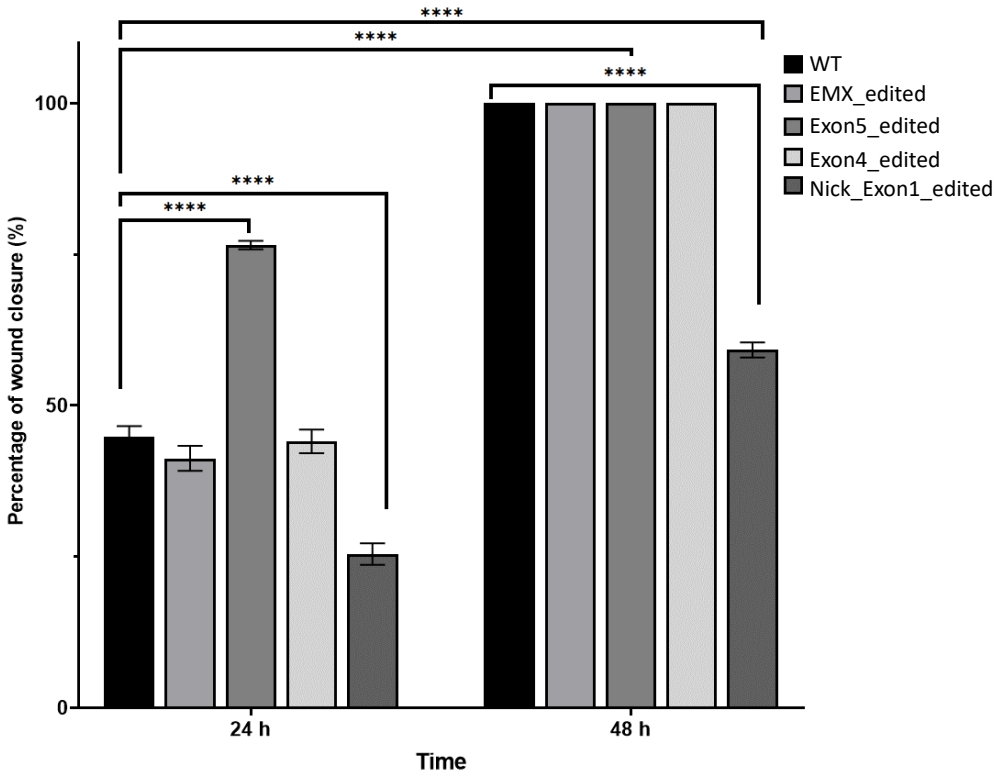


*Figure 4-30. Wound healing assay of wild type and SLC22A1 HEK293 edited cells. Cells were seeded in a 2 well culture-insert (ibidi GmbH) at a density of 3,500 cells per well. Migration was documented for 48 h. For every time point, three pictures were taken using the Celestron HD Microscope Imager. Data is presented as Mean ± SEM * p <0.05, ** p < 0.01, *** p< 0.001*

Results from the CRISPR-mediated disruption of SLC22A1 are presented in Figure 4-30. What stands out in the chart is that targeted Exon5 showed a significant increase ($p< 0.0001$) in wound healing with a closure of 76.56% ± 0.7% at 24h after scratch while there was no significant change when gene was targeted in Exon 4, with a closure similar to wild type cells, as 44.07% ± 1.9% migration at 24h and full wound closure 48 h after. However, use of CRISPR nickase in Exon1 showed a significant difference at the $p = 0.0001$ level, with a wound closure of 25.4% ± 1.8.

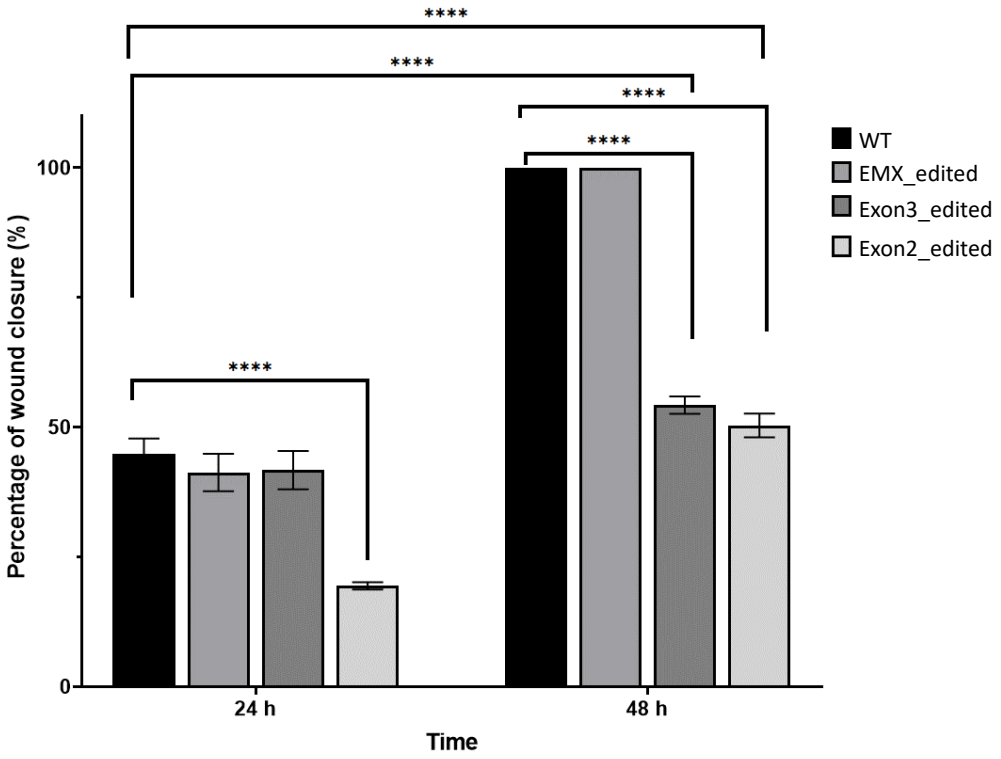


*Figure 4-31. Wound healing assay of wild type and SLC22A4 HEK293 edited cells. Cells were seeded in a 2 well culture-insert (ibidi GmbH) at a density of 3,500 cells per well. Migration was documented for 48 h. For every time point, three pictures were taken using the Celestron HD Microscope Imager. Data is presented as Mean ± SEM **** p< 0.0001*

Transfected cells with targets in SLC22A4 were analysed for wound healing and compared to wild type cells. CRISPR cells showed a decreased migration, compared to wild type. From Figure 4-31, it can be seen that by far the greatest decrease in migration is shown in Exon2 disruption, where wound closure was significantly ($p<0.0001$) reduced to 19.42% ± 0.4123% at 24 h, whereas the wound closure at 48 was 50.34% ± 1.322%, with a mean increase of 30.92% in 24 h. In comparison to wild type cells, where wound closure at 24 was 44.89 % ± 1.695% and wound was completely close at 48 h. Reduction in wound healing when Exon3 was targeted was no significantly different at 24h, however, migration at 48 h showed a significant ($p<0.0001$) slow migration of cells when

compared to the previous time-point, where migration changed from 41.71% ± 2.135% to 54.25 % ± 0.9644 in a 24 h window.

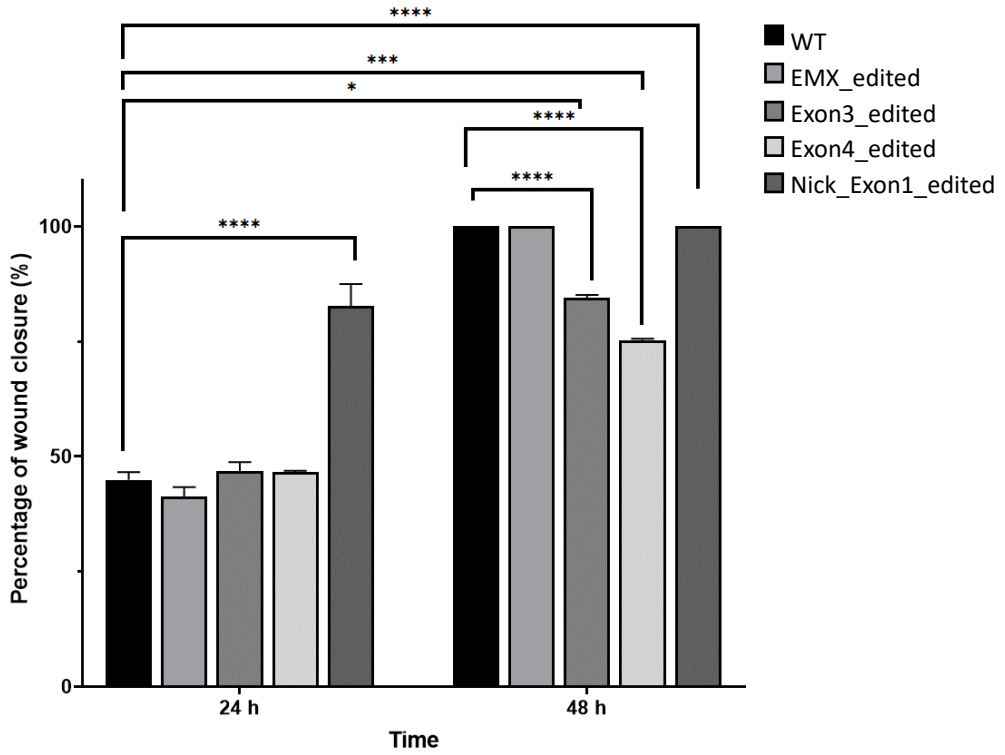


*Figure 4-32. Wound healing assay of wild type and SLC22A5 HEK293 edited cells. Cells were seeded in a 2 well culture-insert (ibidi GmbH) at a density of 3,500 cells per well. Migration was documented for 48 h. For every time point, three pictures were taken using the Celestron HD Microscope Imager. Data is presented as Mean ± SEM * p <0.05, ** p < 0.01, *** p< 0.001*

Data from wound healing assay in HEK293 with CRISPR-mediated mutations in SLC22A5 gene is summarised in Figure 4-32. Exon 4 and Exon 3 showed no significant difference between the migration at 24 h after, with a gap of 46.60 ± 0.2917 and 46.79 ± 1.953, respectively. Surprisingly, migration of cells with CRISPR nickase was increased, showing a significant difference at the p = 0.001 level, with a wound closure of 82.61% ± 4.870% at 24 h. However, at 48 h, Exon4 and Exon3 showed a significant reduction (p<0.001) in wound closure with migration of 75.26% ± 0.3247 and 84.44% ± 0.6474, respectively.

4.3. Discussion

The utilization of the CRISPR-Cas9 system as a powerful tool for gene editing has revolutionized the field of molecular biology. In this study, we aimed to assess the efficiency of the CRISPR-Cas9 system in disrupting SLC22A1, SLC22A4, and SLC22A5 genes in HEK 293 cells and lung cells, A549 and Calu-3 cells. To achieve this, we employed three different strategies, namely the Cas9 protein, nCas9 protein and RNP complexes, as tools for gene editing.

In the current study, the pX459 plasmid was used as a vector to construct six different recombinant plasmids to target the protein transporters SLC22A1, SLC22A4 and SLC22A5. In order to optimize the genome-editing efficiency, the design of sgRNAs was carefully executed, taking into consideration the GC content that is favourable for gene editing. The GC content of the sgRNAs was within the range (40%-60%) that has been reported in previous literature studies (Doench et al., 2014; Liu et al., 2016; T. Wang et al., 2014). This meticulous design process aimed to minimize off-target effects and maximize the efficiency of the CRISPR/Cas9 system in performing gene editing.

After verifying the insertion of sgRNAs, constructs were transfected into cells of the kidney epithelial cell line HEK293 and lung cells A459 and Calu-3. Cell selection was performed using puromycin and transfection of pX459-EMX1 as a control was carried out at simultaneously. T7E1 assay, Sanger sequencing and qPCR were used to identify cell lines with a gene phenotype. nCas9 and RNP complexes were designed as alternative strategies for gene editing.

Calu-3 cells were unable to survive after transfection with the various CRISPR/Cas9 strategies employed. Cells were transfected with Cas9, nCas9 nickase plasmids, and RNP complexes, but no cell survival was observed after selection with puromycin. Additionally, there was no expression of GFP from the nCas9 plasmid and no cell survival after transfection with RNP complexes.

In an attempt to establish a lung epithelial cell line model for CRISPR-Cas9 editing, A549 cells were transfected with nickase Cas9 (nCas9) and green fluorescent protein (GFP) as a selection marker. T7E1 assay showed no alterations in the sequence, and the amplicons were subjected to sequencing. Furthermore, amplicon sequencing confirmed that the targeted genomic region remained unchanged compared to the wild-type (WT) reference sequence. A potential explanation for the lack of observed editing in A549 cells could be their inherent genetic makeup. A549 is a hypotriploid human cell line, characterized by a modal chromosome number of 66 (occurring in approximately 24% of cells). This cell line also exhibits frequent variations in ploidy, with modal numbers of 64 and 67 observed alongside infrequent occurrences of even higher ploidy states (S.-Y. Park et al., 2001; Peng et al., 2010).

The decreased efficiency of Cas9 in targeting heterochromatic (highly condensed) regions highlights the role of DNA accessibility in specificity and efficiency. However, the hetero chromatin state of the target sequence can reduce the diffusion of Cas9, which may result in decreased binding due to its limited accessibility to the PAM sequence (Knight et al., 2015; Verkuijl & Rots, 2019). The sensitivity to CRISPR-Cas9 targeting in amplified genomic regions was also observed for genes that showed a lack of significant mRNA expression

(Aguirre et al., 2016a). mRNA expression profiles of the different transporters were measured in order to assess if expression might impact the efficiency of CRISPR, as the CRISPR technique was unsuccessful in A549 and Calu-3 cells. Results showed that there were no significant differences in the levels of SLC22A1 mRNA between HEK 293 and A549 cells, with a significant decrease in SLC22A1 mRNA levels observed in Calu-3 cells.

The findings from this study indicate that CRISPR transfection in lung cells is challenging, as demonstrated by the lack of efficacy with three different transfection methods. A possible explanation for this is that as the stability of the genome is of utmost importance for the survival of all organisms. Accurate DNA replication and repair processes are crucial to ensure genome stability. These processes involve the use of numerous enzymes that modify DNA and must be regulated carefully. Approximately 90% of DNA double-strand breaks (DSBs) produced by Cas9 in rapidly growing mammalian cells are repaired within a time frame of 1 hour (Goodarzi et al., 2010; Metzger & Iliakis, 1991). CRISPR-Cas interference may trigger genome instability and cell death which may hinder successful genome editing in different cell lines (Cubbon et al., 2018; Kosicki et al., 2018). In addition, the efficiency of the CRISPR/Cas9-mediated genome editing process is influenced by multiple factors, numerous methods have been employed to enhance genome editing efficiency while reducing the occurrence of off-target effects. The efficiency of CRISPR/Cas9-mediated genome editing is influenced by various factors. While cell doubling time is one potential contributor, our results suggest it may not be the sole determinant in the present study.

HEK293 cells, with a doubling time similar to A549 cells (24 hours vs. 22 hours) (Lujan et al., 2019), exhibited demonstrably higher editing efficiency compared to Calu-3 cells (doubling time: 72 hours) (Cervera et al., 2011). This observation suggests that other factors beyond cell division rate may be playing a role in the observed differences. Cancer cell lines, including Calu-3, often exhibit enhanced DNA repair proficiency compared to non-cancerous cell lines like HEK293. This heightened repair activity could potentially counteract CRISPR-Cas9 mediated double-strand breaks, reducing editing efficiency.

Another possible explanation for this might be that CRISPR-Cas9 targeting in cancer cell lines produces two responses: (1) an early anti-proliferative effect from the induction of multiple double-strand DNA breaks, and (2) the essentiality of the target gene from its knock-out and loss of normal protein expression (Aguirre et al., 2016b). This previous study indicates that even a single CRISPR-Cas9-mediated DNA cut leads to a reduction in cell proliferation. In a study conducted by Mu et al., (2019) it was observed that the decreased cell viability and colony forming ability were significantly affected after the transfection of RNP complexes with sgRNA that were in vitro transcribed by T7 polymerase through electroporation. This factor may explain the cell death seen when RNP complexes were used in Calu-3 cells although further investigations are needed to understand the mechanism of resistance and to develop new strategies to increase the efficiency of CRISPR/Cas9-mediated genome editing in Calu-3 cells.

HEK293 are widely used for protein production and easy transfection (Tan et al., 2021; H. Yang et al., 2019), which contrasts with the use of lung epithelial

cells. It has been mentioned that variability in results between different cell types has been attributed to the efficacy of CRISPR edits (Haeussler et al., 2016; Konstantakos et al., 2022). To mitigate this issue, an alternative approach would be to conduct extensive CRISPR activity experiments in A549 and Calu-3 cells. Transporters, as members of the ADME gene group, play crucial roles in maintaining normal physiological processes, homeostasis, and drug response and safety. Future investigations could explore alternative approaches, such as employing lentiviral Cas9 delivery systems for CRISPR knockout (KO) techniques. This method employs lentiviral vectors to introduce CRISPR/Cas9 components, targeting specific genes of interest, into the cells for precise genomic modifications. Where genome-wide knockouts have been performed in Calu-3 and A549 cells (Rebendenne et al., 2021; Wei et al., 2021)

CRISPR-Cas9 editing of SLC22A genes in HEK293 cells revealed distinct effects on cell proliferation, adhesion, and migration, highlighting the potential roles of these transporters in regulating these cellular processes. Editing of SLC22A1, particularly in Exon5, resulted in a significant decrease in cell adhesion compared to wild-type cells. Interestingly, Exon5 editing also showed an increase in wound healing at 24 hours, suggesting a potential disconnect between adhesion and migration in this context. However, editing in Exon4 led to a decrease in cell adhesion without a significant change in wound healing. While the underlying mechanisms need further investigation, these findings suggest SLC22A1 might play a crucial role in cell adhesion and may influence migration depending.

Disruption of SLC22A4 via CRISPR-Cas9 editing led to a significant reduction in both cell proliferation and adhesion. Notably, editing in Exon 2displayed the most pronounced decrease in cell migration, indicating SLC22A4's potential role in these processes. CRISPR-Cas9 editing of SLC22A5 exhibited varying effects on cell migration. While Exon4 and Exon3 disruptions did not significantly affect migration at 24 hours, they showed a decrease at 48 hours. Conversely, editing with Cas9 nickase in Exon1 resulted in a significant increase in migration at 24 hours, followed by a decrease at 48 hours. These complex findings suggest a potential interplay between SLC22A5 function and migration dynamics, warranting further investigation to understand the underlying mechanisms.

CRISPR-Cas9 editing of SLC22A genes in HEK293 cells revealed their potential roles in regulating cell proliferation, adhesion, and migration. These findings provide a valuable starting point for further investigation into the specific mechanisms by which these transporters influence cellular behaviour and their potential contribution to disease progression. There are similarities between the results in this study and those described by Mukherjee, (2012) where reduction of SLC22A1 and SLC22A5 expression by siRNA reduced proliferative activity, therefore affecting wound repair mechanisms. These results further deepen our understanding of the crucial role played by SLC22 transporters in various cellular processes and have implications for the design and development of drugs. Further research is needed to elucidate the specific mechanisms by which SLC22A4 dysfunction disrupts cellular functions.

Unfortunately, attempts to validate protein expression levels through Western blotting were unsuccessful within the scope of this study. This limits the ability to directly correlate the observed functional changes with specific protein structure modifications caused by targeted disruptions in different exons. It is important to acknowledge that CRISPR-Cas9, while a powerful tool for genome editing, can induce unintended consequences beyond the targeted modifications. The study by Cullot et al., (2019) describes the occurrence of megabase-scale chromosomal truncations during CRISPR-Cas9 editing, potentially leading to loss of function for genes flanking the targeted region. While not directly observed in the current study, future investigations may benefit from employing techniques like deep sequencing to assess potential off-target effects and ensure the integrity of the edited loci.

The function of solute carrier (SLC) proteins in facilitating material transport within cells has been well established. Previous studies have focused on the impact of non-synonymous polymorphisms in modulation of pharmacokinetics and therapeutic response. However, recent research has highlighted the potential involvement of specific SLCs in the progression of tumours. (Colas et al., 2016; Liu, 2019; Y. Zhang et al., 2019). Despite the significant role that solute carrier (SLC) proteins play in cellular function, the functions of many SLCs remain largely uninvestigated, particularly regarding their involvement in migration, proliferation, and adhesion. To address this, the present study conducted cell migration, adhesion, and cell proliferation assays on HEK293 edited cell lines to assess any changes.

Previous studies using knockout mice have corroborated the critical role of OCTN1 in the absorption, distribution, and retention of ergothioneine (Kato et al., 2010). In addition, octn1−/− mice displayed a statistically significant, albeit modest, decrease in tolerance to intestinal oxidative stress. Additionally, Shinozaki et al., (2017) found elevated oxidative stress markers in OCTN1-deficient mice with chronic kidney disease (CKD), potentially contributing to the observed exacerbation of kidney fibrosis. These findings collectively highlight the importance of OCTN1 in maintaining optimal tissue ergothioneine levels and mitigating oxidative stress-related pathologies. It is important to acknowledge that gene knockout models, while valuable for elucidating gene function, can sometimes exhibit minimal phenotypic alterations. his might be attributed to compensatory mechanisms involving other transporter gene products in vivo or functional redundancy arising from homologous transporters with overlapping substrate specificities.

Zebrafish lacking OCTN1 displayed no significant morphological or behavioural differences. However, these knockout animals exhibited an elevation in 8-oxo-7,8-dihydroguanine (8-oxo-dG) levels, a widely used marker of oxidative DNA and RNA damage (Pfeiffer et al., 2015). These findings, while derived from a non-mammalian model system, suggest a potential role for OCTN1 in mitigating oxidative stress and raise intriguing questions about the interplay between OCTN1 function and intestinal inflammation. Polymorphisms within the SLC22A4 gene have been linked to an increased susceptibility to chronic inflammatory diseases such as Crohn's disease, ulcerative colitis, and type I diabetes (Cheah & Halliwell, 2012; Gründemann, 2012).

Utilizing a mouse model with a non-functional OCTN2 (SLC22A5) carnitine transporter mutation, Sonne et al., (2012) investigated the impact of carnitine deficiency on the neonatal gut. This depletion coincided with stunted villous growth, indicative of impaired development. Additionally, the OCTN2−/− intestines displayed early signs of inflammation, characterized by infiltration of lymphocytes and macrophages, alongside a breakdown of villous architecture. Furthermore, researchers observed a significant increase in apoptosis within gut samples from OCTN2−/− mice. The carnitine deficiency also triggered a severe immune phenotype, manifesting as atrophy of the thymus, spleen, and lymph nodes, again attributed to heightened apoptosis.

Previous research supports the notion that both SLC22A5 mutations and other SLC22 family members can impact cell proliferation. Where downregulation of members of the SLC22 might contributed in the regulation of cell proliferation, migration, invasion, in liver cancer cells and cholangiocellular carcinoma (CCA) (Fang et al., 2021; Lautem et al., 2013; G. Zhu et al., 2019). Missense variants in SLC22A5 have been identified as contributing to the development of rare monogenic disorders (Koleske et al., 2022). Targeting different exons within SLC22A genes may lead to distinct protein variants, potentially explaining the observed variations in the regulation of cell proliferation, adhesion, and migration and lower functionality.

Previous research has demonstrated that other transporters within the SLC22 family play a role in cell proliferation. A recent study demonstrated that downregulation of members of the SLC22 might contributed in the regulation of cell proliferation, migration, invasion, in liver cancer cells and

cholangiocellular carcinoma (CCA) (Fang et al., 2021; Lautem et al., 2013; G. Zhu et al., 2019). Present results suggest that different exons can lead to different protein structures, thereby affecting the regulation of cell proliferation, adhesion, and migration. Research has primarily centered on the implications of polymorphisms on drug absorption (Kawoosa et al., 2022; Santiago et al., 2006; K. Zhou et al., 2009), however, there is evidence to suggest that there may be additional effects on cell physiology as demonstrated in cancer studies where gene expression has been linked to cancer stage . The findings in this study align with the conclusions reached in prior research, where the suppression of a member of the Solute Carrier Family has been shown to result in the inhibition of cell growth, a decrease in migration and invasion, as well as the arrest of the cell cycle in breast cancer (Yen et al., 2018). Other Solute Carrier Transporters have been identified as playing a role in maintaining the normal rate of cell proliferation (S. Feng et al., 2020; Roblek et al., 2022).

While the molecular mechanisms involved in the impact of transporters belonging to the SLC family on cellular development has been investigated in this study, there are several limitations to consider. The study only focused on one specific cell line, HEK293 cells, and may not accurately represent the effects of these transporters in other cell types or *in vivo*. In addition, the results obtained from this study are limited to the effects on migration, proliferation, and adhesion in HEK293 cells and may not accurately reflect the impact on other cellular processes such as angiogenesis, apoptosis, and metabolism.

Examining the molecular processes involved in the gene expression of transporters within the ADME gene group is essential for a comprehensive

understanding of their influence on cellular development, tumour progression, and the effectiveness and safety of drugs.

4.4. Conclusion

In conclusion, the CRISPR-Cas9 system has proven to be an effective tool for gene editing in HEK 293 cells, but its application in lung epithelial cells such as A549 and Calu-3 proved challenging. The study aimed to assess the efficiency of the CRISPR-Cas9 system in disrupting SLC22A1, SLC22A4, and SLC22A5 genes in HEK 293 cells using three different strategies, Cas9 protein, nCas9 protein, and RNP complexes.

The mRNA expression profiles showed no significant differences in the levels of SLC22A1, SLC22A4, and SLC22A5 between the cell lines. To mitigate the variability in results between different cell types, future investigations might consider conducting extensive CRISPR activity experiments in lung epithelial cells. The results of this study provide valuable insight into the application of the CRISPR-Cas9 system in lung epithelial cells, and highlight the need for further research to understand the mechanism of resistance and develop new strategies to increase the efficiency of CRISPR-Cas9-mediated genome editing. Based on the results from migration, proliferation and adhesion assays in HEK 293 cells, it is possible that SLC22 genes may play a role in tumour progression. The findings from these assays suggest that the contribution of specific SLCs to these processes cannot be overlooked.

CHAPTER 5. ANALYSIS OF TRANSCRIPTIONAL REGULATION OF SLC22A1, SLC22A4 AND SLC22A5 IN LUNG EPITHELIA

CHAPTER 5. ANALYSIS OF TRANSCRIPTIONAL REGULATION OF SLC22A1, SLC22A4 AND SLC22A5 IN LUNG EPITHELIA

5.1. Introduction

Transcriptional regulation refers to alteration in the level of gene expression by changing the rate of transcription. It is a vital process in regulation of gene function in eukaryotes that plays an important role in the accuracy and diversity of genetic information transmission. Transcriptional regulation in eukaryotes includes various processes that are highly related to each other such as DNA methylation, histone modification, chromatin remodelling, and control of transcription via different regulatory factors (Mitsis et al., 2020). Currently, DNA methylation stands as one of the most extensively investigated epigenetic processes. This process entails the addition of a methyl group at the fifth carbon position of cytosine residues, which occurs within the context of a CpG dinucleotide. CpG dinucleotides are not distributed evenly across the human genome, and instead tend to form clusters within compact sequences called CpG islands, which constitute a mere 1% of the entire genome sequence. These CpG islands are enriched in gene promoter regions and serve critical functions in the regulation of gene expression under normal physiological conditions.

The analysis of transcriptional regulation of genes typically involves identifying and characterizing the cis-regulatory elements within or near the gene encoding, as well as the transcription factors that bind to them. Techniques such as ChIP-seq and RNA-seq can be used to identify and quantify the binding of

transcription factors to cis-regulatory elements and the expression of the gene, respectively.

In addition to identifying cis-regulatory elements and transcription factors, the analysis of transcriptional regulation of protein transporters also involves understanding how these elements and factors interact to control gene expression.

Recent studies have highlighted the involvement of SLCs in cancer development. The downregulation of several SLCs has been associated with tumour growth and proliferation, suggesting a role in tumour progression (Rashid et al., 2021). mRNA expression of hOCT1 has been linked to the expression of mRNA encoding for HNF4α and C/EBP in the human liver. Rulcova et al., (2013) found a significant correlation between hepatic mRNA expression of hOCT1 and that of HNF4α and C/EBP. Two important response elements for HNF4α in the 5'-flanking region of the solute carrier 22A1 (SLC22A1) gene, which encodes hOCT1, have been identified by Hyrsova et al., (2016) and Saborowski et al., (2006).

A previous study by (Drenberg et al., 2017) demonstrated that decreased expression of the uptake transporter SLC22A4, is strongly associated with unfavourable event-free and overall survival outcomes in various cohorts of acute myeloid leukemia (AML) patients treated with the cytidine nucleoside analog, cytarabine (Ara-C). A recent analysis found SLC22A4 gene to be regulated by methylation within the promoter region (Buelow et al., 2021).

Methylation of CpG sites in the region spanning -354 to +85 in SLC22A5 has been found to be negatively correlated with the expression of SLC22A5 in

cancer cells (Qu et al., 2013). SLC22A5 has been reported to be regulated via an intronic-estrogen responsive element (ERE) in breast cells (C. Wang et al., 2012). Furthermore, supporting clinical evidence indicates a decrease in the mRNA level of SLC22A5 in liver biopsies of patients treated with the PXR agonist carbamazepine (Oscarson et al., 2006). However, our current understanding of the transcriptional factors involved in gene regulation remains limited.

Understanding the transcriptional regulation of transporter proteins can provide important insights into how these genes are controlled in response to different physiological conditions or in response to drug treatments and can help to identify new targets for drug development. It is vital to precisely characterise regulations, owing notable to the major role played by transporters in pharmacokinetics and some toxic effects of drugs. Due to their critical role in drug absorption, distribution and elimination, alteration in the activities of SLC transporters may result in pharmacokinetic changes and consequently in drug response. The activity of a transporter is subject to regulation by various factors such as genetic polymorphisms, co-medication, environmental toxins, and disease conditions.

In contrast to their well-recognised clinical importance, the regulatory mechanisms accounting for the activities of SLC transporters remain less characterized in general. Although the variety of studies regarding the genetic regulation of SLCs has focused on liver and kidney, not much has been reported on the role of genetic variants in physiological and pathophysiological processes in the lung.

The regulation of transport activity in response to endogenous and exogenous signals may occur at various levels such as transcription, mRNA stability, translation, and posttranslational modification. This diversity of regulatory mechanisms may be advantageous to correspond to various biological signals. In general, transcriptional regulation and posttranslational modification are believed to be responsible for long-term and short-term regulation, respectively.

Depending on the tissue or cell type, additional regulation mechanisms have been described. The present chapter focuses on understanding the transcriptional regulation of transporter proteins, as changes in transport activity are dynamically regulated by increases or decreases in levels of mRNA expression. The tissue-specific expression of transporters is also under transcriptional control, although there is little information about the mechanisms behind lung expression.

Due to the lack of understanding of the role of genetic regulation of SLC22A1, SLC22A4 and SLC22A5, this chapter details the construction of luciferase-based reporter systems for assaying the genetic regulation of these transporters. The dual-luciferase assay has been widely used in cell lines to determine rapidly and accurately the activity of a given promoter. Although this strategy has proved very useful, it does not allow the promoter and gene function to be analysed in the context of the whole organism.

5.2. Results

SLC transporters are present in tissues with varying abundance. The transcriptional regulation is the major determinant of tissue distribution and abundance of these drug transporters.

5.2.1. Analysis of Methylation Array data

Investigation of regulatory mechanisms in lung was expanded to cover lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) methylation. Illumina Human Methylation 450K array data were obtained from the Cancer Genome Atlas Program (TCGA, https://www.cancer.gov/tcga.").

A total of 42 normal tissue samples and 370 primary tumour samples for the LUSC cohort were analysed using a genome-wide approach to identify differentially methylated regions. The LUAD cohort included 32 normal tissue samples and 473 primary tumour samples, expanding the investigation of regulatory mechanisms in lung cancer. CpG sites' methylation quantification was quantified by the beta value (β), calculated as M/(M + U + a), where M and U represent the signal intensities of methylated and unmethylated states, respectively.

In this study, specific genomic regions are defined based on their positions relative to the transcriptional start site (TSS) of genes. "TSS1500" and "TSS200" respectively refer to regions located 200–1500 bases upstream and 0–200 bases upstream of the TSS. Additionally, the 5′ untranslated region (5′UTR) spans from the TSS to the ATG start site, while the "1stExon" represents the gene's first exon. The "Body" region extends from the ATG start site to the stop codon,

and the 3′ untranslated region (3′UTR) lies between the stop codon and the poly-A tail. DNA methylation was correlated with SLC22A1, SLC22A4 and SLC22A5 mRNA expression with paired samples from RNAseq obtained from TCGA-LUSC or LUAD, respectively. Pearson's correlation coefficient was used to quantify the strength and direction of the relationship between DNA methylation and gene expression.

Analysis of DNA methylation at SLC22A regulatory regions and its association with gene expression in lung squamous cell carcinoma

The study investigated the level of DNA methylation at five CpG sites located in the SLC22A1 promoter region. The CpG sites tested were cg05314142, cg13466809, cg22416916, cg24864413, and cg27292431. Figure 1 1 illustrates the DNA methylation levels of three genes, SLC22A1, SLC22A4, and SLC22A5, in the LUSC cohort. The methylation profiles of each gene are displayed as median levels of β values on the y-axis . The differences in DNA methylation levels between the two sample types are visually represented in the figure.

Results showed that CpG site cg22416916 had the highest median beta methylation level of 0.939, with a standard deviation of 0.114. CpG site cg05314142 had a lower median beta methylation level of 0.895, with a standard deviation of 0.101. CpG site cg24864413 had a median beta methylation level of 0.881, with a standard deviation of 0.083. CpG site cg27292431 had a median beta methylation level of 0.870, with a standard deviation of 0.055. Finally, CpG site cg13466809 had the lowest median beta methylation level of 0.73, with a standard deviation of 0.088.

Table 1 1 presents the CpG sites located within the promoter regions of SLC22A1, SLC22A4, and SLC22A5 genes and their correlation to the RNA expression levels. Integrative analysis with RNA expression showed a negative correlation between DNA methylation at CpG site cg13466809 and RNA-seq expression levels (Pearson's correlation coefficient = -0.12). Similarly, CpG site cg05314142 also showed a negative correlation with RNA-seq expression levels (coefficient = -0.09). However, the CpG sites cg22416916 and cg24864413 in the TSS200 region showed weaker, non-significant correlations with RNA-seq expression levels (Pearson's correlation coefficients of -0.02 and 0.002, respectively). Finally, CpG site cg27292431 in the first exon of gene CG27292431 showed a positive, albeit weak, correlation with RNA-seq expression levels (Pearson's correlation coefficient = 0.06).

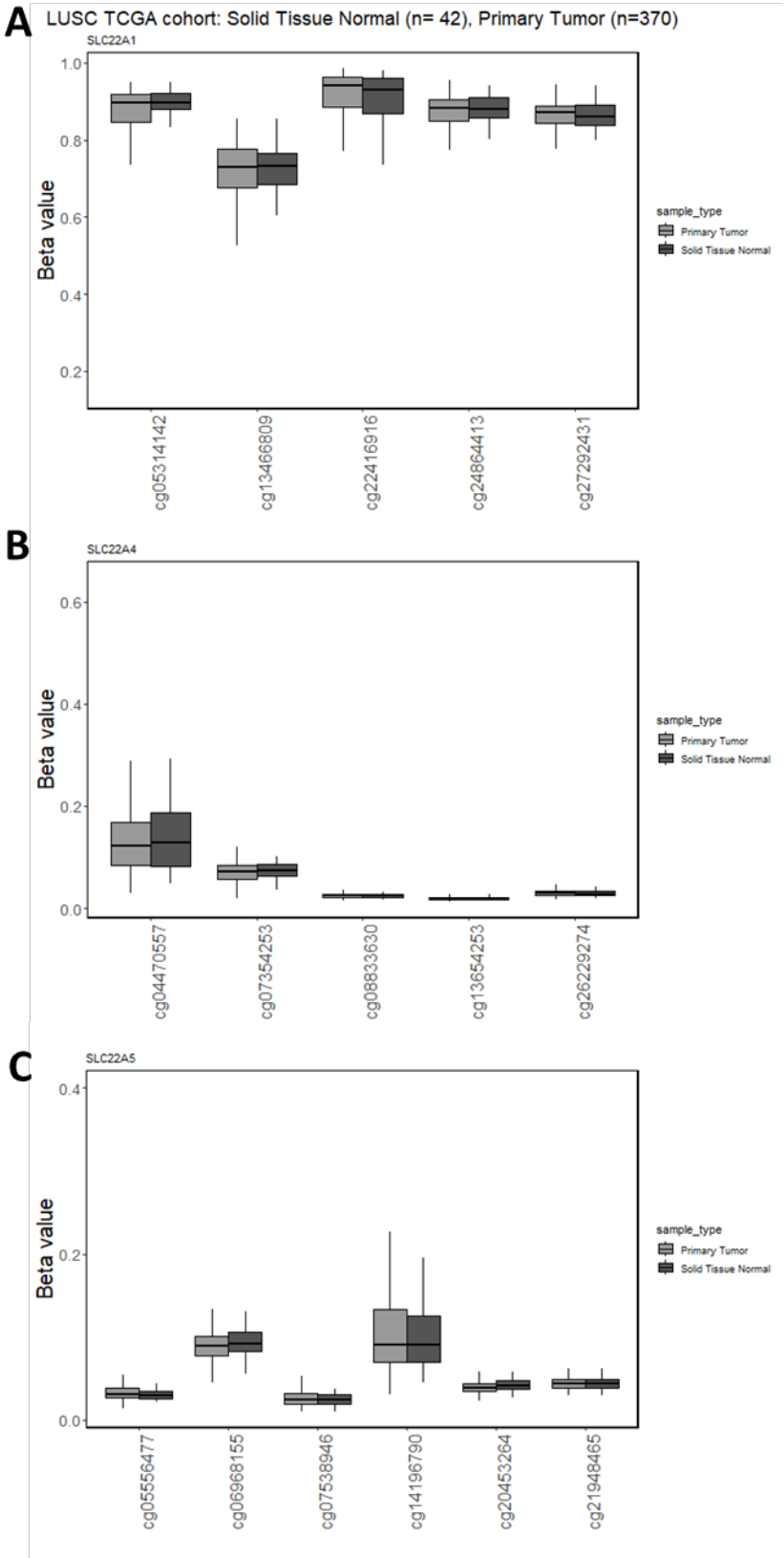

A
LUSC TCGA cohort: Solid Tissue Normal (n= 42), Primary Tumor (n=370)
SLC22A1
Beta value
1.0
0.8
0.6
0.4
0.2
cg05314142
cg13466809
cg22416916
cg24864413
cg27292431
sample_type
Primary Tumor
Solid Tissue Normal
B
SLC22A4
Beta value
0.6
0.4
0.2
0.0
cg04470557
cg07354253
cg08833630
cg13654253
cg26229274
sample_type
Primary Tumor
Solid Tissue Normal
C
SLC22A5
Beta value
0.4
0.2
0.0
cg05556477
cg06968155
cg07538946
cg14196790
cg20453264
cg21948465
sample_type
Primary Tumor
Solid Tissue Normal

Figure 5-1. DNA methylation of SLC22A1, SLC22A4 and SLC22A5 in lung squamous cell carcinoma and solid tissue normal. Methylation profiles of A) SLC22A1, B) SLC22A4, C) SLC22A5. DNA methylation (y-axis) is given as Beta value. Profiles show mean levels.

The DNA methylation levels of five CpG sites in the SLC22A4 gene were analysed and are displayed in Figure 1 1B. The median DNA methylation levels and standard deviations were calculated for each CpG site. CpG site cg04470557 had the highest median DNA methylation level of 0.124, with a standard deviation of 0.087. CpG site cg07354253 had a median DNA methylation level of 0.071, with a standard deviation of 0.021. CpG site cg08833630 had a median DNA methylation level of 0.025, with a standard deviation of 0.005. CpG site cg13654253 had the lowest median DNA methylation level of 0.019, with a standard deviation of 0.003. Finally, CpG site cg26229274 had a median DNA methylation level of 0.029, with a standard deviation of 0.007.

The relationship between DNA methylation and expression levels of the SLC22A4 protein was investigated for five CpG sites located in the promoter region. The results showed a negative correlation between DNA methylation at CpG sites cg13654253 in the TSS150 region, cg08833630 in the TSS200 region, and cg04470557 in the body region of SLC22A4, and mRNA expression levels (Pearson's correlation coefficients of -0.0501, -0.0469, and -0.0453, respectively). However, the CpG site cg07354253 in the TSS1500 region showed a weak, non-significant negative correlation with mRNA expression levels (Pearson's correlation coefficient = -4.00E-04). In contrast, CpG site cg26229274 in the TSS1500 region showed a positive correlation with mRNA expression levels (Pearson's correlation coefficient = 0.093).

Table 5-1. Correlation Analysis of DNA methylation of individual CpG sites and SLC22A1, SLC22A4, SLC22A5 mRNA expression in lung squamous cell carcinoma (n-=473). TSS1500: 200-1500 bases upstream of the transcriptional start site (TSS). TSS200: 0-200 bases upstream of the TSS. 5'UTR: 5' untranslated region located between the TSS and the ATG start site. 1stExon: the first exon of the gene. Body: the region between ATG start site and stop codon.

Gene	Position (hg38)	CpG	Group	Pearson's correlation coefficient	Adj.P.Val
SLC22A1					
	chr6:160,121,912	cg13466809	1stExon;5'UTR	-0.1228	0.0988
	chr6:160,121,679	cg05314142	TSS200	-0.0906	0.2593
	chr6:160,121,738	cg22416916	TSS200	-0.0206	0.8065
	chr6:160,121,700	cg24864413	TSS200	0.0018	0.9938
	chr6:160,122,229	cg27292431	1stExon	0.0623	0.5882
SLC22A4					
	chr5:132,294,217	cg13654253	TSS150	-0.0501	0.6852
	chr5:132,294,288	cg08833630	TSS200	-0.0469	0.6852
	chr5:132,295,302	cg04470557	Body	-0.0453	0.6852
	chr5:132,294,230	cg07354253	TSS1500	-4.00E-04	0.9938
	chr5:132,293,817	cg26229274	TSS1500	0.0927	0.2587
SLC22A5					
	chr5:132,369,458	cg21948465	TSS1500	-0.1553	0.0205
	chr5:132,369,343	cg14196790	TSS1500	-0.1186	0.1069
	chr5:132,369,496	cg07538946	TSS1500	-0.0242	0.8052
	chr5:132,369,420	cg06968155	TSS1500	-0.0232	0.8052
	chr5:132,370,050	cg20453264	1stExon	0.003	0.9938
	chr5:132,369,627	cg05556477	TSS200	0.0269	0.8052

DNA methylation analysis of SLC22A5 revealed distinct median beta values and standard deviations for six CpG sites, Figure 5-1C. The CpG site cg05556477 located in the TSS200 region displayed a median beta value of 0.031 and a standard deviation of 0.016. CpG site cg06968155 situated in the TSS1500 region showed a higher median beta value of 0.089 with a standard deviation of 0.020. In contrast, CpG site cg07538946 located in the TSS1500 region had a lower median beta value of 0.024 with a larger standard deviation of 0.022. CpG site cg14196790, also in the TSS1500 region, had a median beta value of 0.091 with a standard deviation of 0.065. CpG site cg20453264 in the 1st Exon region showed a median beta value of 0.039 with a standard deviation of 0.018. Finally, CpG site cg21948465 situated in the TSS1500 region had a median beta value of 0.043 with a smaller standard deviation of 0.009.

Findings revealed a negative correlation between DNA methylation at two CpG sites (cg21948465 and cg14196790) located in the TSS1500 region of SLC22A5 and mRNA expression levels, with Pearson's correlation coefficients of -0.155 and -0.119, respectively. However, the CpG sites cg07538946 and cg06968155 in the same region showed weaker, non-significant negative correlations with mRNA expression levels (Pearson's correlation coefficients of -0.024 and -0.023, respectively). In addition, CpG site cg05556477 in the TSS200 region showed a weak, non-significant positive correlation with mRNA expression levels (Pearson's correlation coefficient=0.027). CpG site cg20453264 in the first exon of gene SLC22A5 showed no significant correlation with mRNA expression levels.

Analysis of DNA methylation at SLC22A regulatory regions and its association with gene expression in lung squamous cell carcinoma

DNA methylation levels of SLC22A1, SLC22A4, and SLC22A5, in lung adenocarcinoma (LUAD) samples were analysed. The methylation profiles of each gene were depicted as median levels of β values on the y-axis, and the differences in DNA methylation levels between the two sample types were visually represented in Figure 5-2. While Table 5-2 presents the correlation of CpG sites located within the promoter regions of SLC22A1, SLC22A4, and SLC22A5 genes, with the RNA expression levels.

Results of methylation analysis of SLC22A1 in LUAD samples are illustrated in Figure 5-2. The methylation levels of three CpG sites, namely cg05314142, cg13466809, and cg24864413, were determined in terms of median and standard deviation. The median methylation level of cg05314142 was found to be 0.854 with a standard deviation of 0.054, while the median methylation level of cg13466809 was 0.740 with a standard deviation of 0.068. Additionally, the median methylation level of cg24864413 was 0.859690423 with a standard deviation of 0.044. Correlation analysis between DNA methylation and RNA expression of SLC22A1 is described in Table 5-2. The CpG site cg24864413 located in the TSS200 region showed a Pearson's correlation coefficient of 0.018 with RNA expression. CpG site cg05314142, also located in the TSS200 region, had a Pearson's correlation coefficient of 0.026. Finally, the CpG site cg13466809 located in the 1stExon;5'UTR region showed a Pearson's correlation coefficient of 0.048.

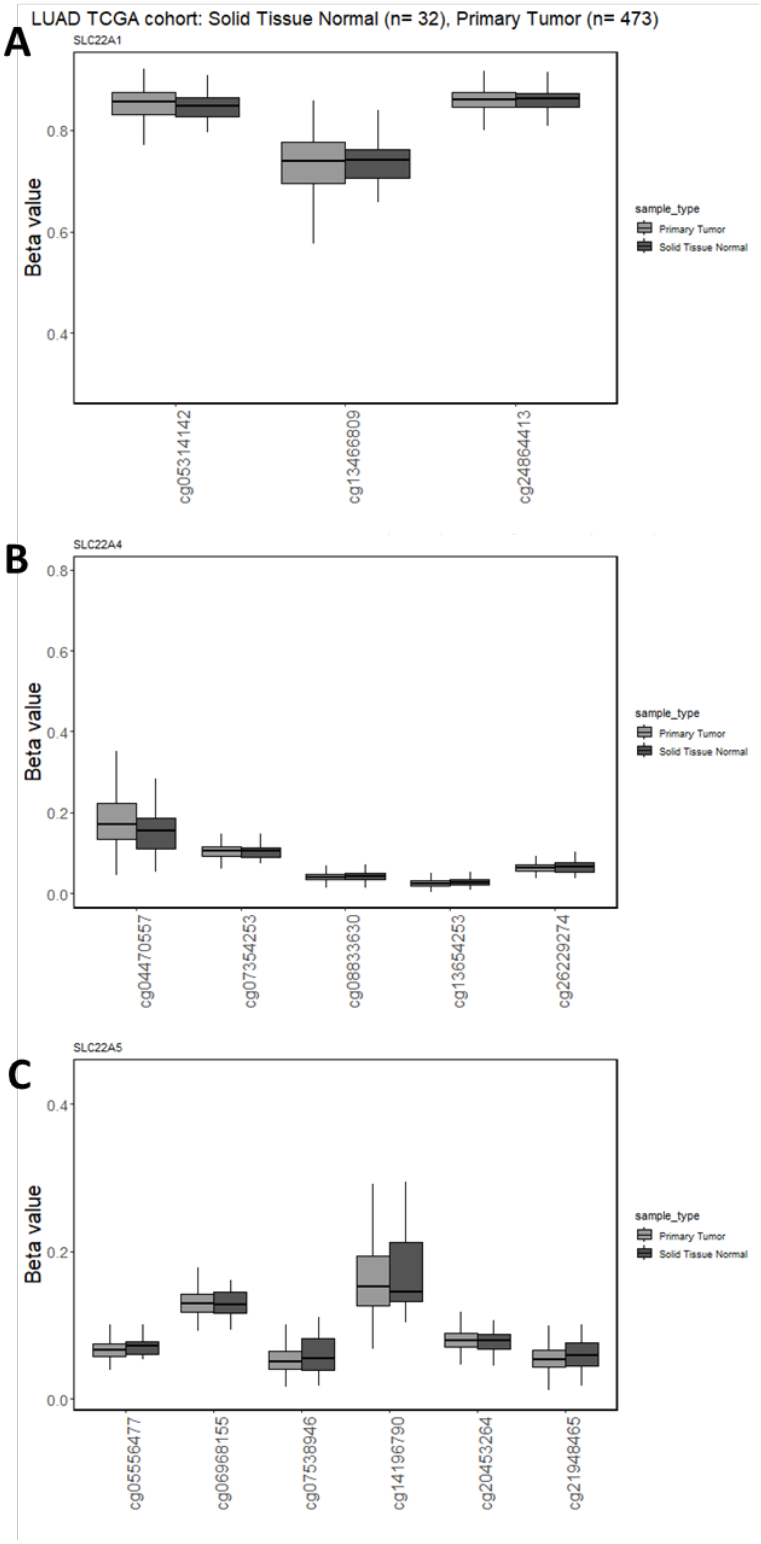


Figure 5-2. DNA methylation of SLC22A1, SLC22A4 and SLC22A5 in lung adenocarcinoma and solid tissue normal. Methylation profiles of A) SLC22A1, B) SLC22A4, C) SLC22A5. DNA methylation (y-axis) is given as Beta value. Profiles show mean levels.

Methylation analysis of SLC22A4 showed 5 different CpG islands located within the promoter region, as seen in Figure 5-2B. The median methylation level of cg04470557 was found to be 0.170 with a standard deviation of 0.081, while the median methylation level of cg07354253 was 0.104 ± 0.024. Additionally, the median methylation level of cg08833630 was 0.040 ± 0.037, and the median methylation level of cg13654253 was 0.0245 ± 0.012. Finally, the median methylation level of cg26229274 was 0.063 ± 0.013. The CpG sites were grouped according to their genomic location, Table 5-2.

The CpG site cg17111895 located in the 1stExon region showed a negative Pearson's correlation coefficient of -0.045 with RNA expression, while cg27372468, located in the 1stExon;5'UTR region, had a negative correlation coefficient of -0.040. CpG site cg08833630 located in the TSS200 region also showed a negative correlation coefficient of -0.0319 with RNA expression, and cg16700673, located in the same TSS200 region, had a correlation coefficient of -0.031.

Additionally, CpG site cg07810106, also located in the TSS200 region, had a correlation coefficient of -0.0262. On the other hand, CpG sites cg26229274 and cg13373085 located in the TSS1500 region showed positive correlation coefficients of 0.010 and 0.011, respectively. CpG site cg12912258 located in the 1stExon;5'UTR region showed a positive correlation coefficient of 0.0257. Finally, the CpG site cg04470557 located in the Body region showed a positive correlation coefficient of 0.096 with RNA expression.

Table 5-2. Correlation Analysis of DNA methylation of individual CpG sites and SLC22A1, SLC22A4, SLC22A5 mRNA expression in lung adenocarcinoma (LUAD). TSS1500: 200-1500 bases upstream of the transcriptional start site (TSS). TSS200: 0-200 bases upstream of the TSS. 5'UTR: 5' untranslated region located between the TSS and the ATG start site. 1stExon: the first exon of the gene. Body: the region between ATG start site and stop codon.

Gene	Location (hg38)	CpG	Group	Pearson's correlation coefficient	adj.P.Val
SLC22A1					
	chr6:160121700	cg24864413	TSS200	0.018	0.8154
	chr6:160121679	cg05314142	TSS200	0.0258	0.7512
	chr6:160121912	cg13466809	1stExon;5'UTR	0.048	0.5871
SLC22A4					
	chr5:132294694	cg17111895	1stExon	-0.0455	0.5871
	chr5:132294519	cg27372468	1stExon;5'UTR	-0.0404	0.6119
	chr5:132294288	cg08833630	TSS200	-0.0319	0.735
	chr5:132294443	cg16700673	TSS200	-0.031	0.735
	chr5:132294429	cg07810106	TSS200	-0.0262	0.7512
	chr5:132293817	cg26229274	TSS1500	0.0099	0.9045
	chr5:132295302	cg04470557	Body	0.0963	0.1885
SLC22A5					
	chr5:132369343	cg14196790	TSS1500	-0.2101	0
	chr5:132369420	cg06968155	TSS1500	-0.1645	0.0028
	chr5:132369496	cg07538946	TSS1500	-0.0648	0.4742
	chr5:132369458	cg21948465	TSS1500	-0.0532	0.5871
	chr5:132370050	cg20453264	1stExon	-0.0086	0.9045
	chr5:132369627	cg05556477	TSS200	0.0919	0.2078

Lastly, methylation analysis of SLC22A5 in LUAD samples from the TCGA returned 6 CpG island located within the promoter region, Figure 5-2C. CpG sites cg14196790 and cg06968155 located in the TSS1500 region showed strong negative Pearson's correlation coefficients of -0.21 and -0.164, respectively, with RNA expression of SLC22A5. CpG site cg07538946 also located in the TSS1500 region showed a negative correlation coefficient of -0.065, and cg21948465 located in the same region had a correlation coefficient of -0.053. CpG site cg20453264 located in the 1stExon region showed a weak negative correlation coefficient of -0.009 with RNA expression. In contrast, CpG site cg05556477 located in the TSS200 region showed a positive correlation coefficient of 0.092 with RNA expression.

5.2.2. Analysis of reporter activity using a luciferase reporter system

Genetic reporter systems are widely used to study eukaryotic gene expression. These assays involve placing a genetic regulatory element upstream of a reporter gene. The Dual-Luciferase assay relies on two different reporter genes, *Renilla* luciferase (Rluc) and *Firefly* luciferase (Fluc), to evaluate the regulation of gene expression. Rluc is fused to a constitutive promoter, whereas the firefly gene is fused to a test promoter. The assay is performed by sub sequentially measuring the luminescence of the *Firefly* and *Renilla* luciferase on the same sample; results are expressed as the ratio of *Fluc* to *Rluc*.

Regulatory elements were analysed in HEK293, Calu-3 and A549 using a luciferase reporter to determine cell-specific regulatory mechanism in SLC22A1, SLC22A4 and SLC22A5.

Partial mapping of the minimal SLC22A1 regulatory elements

To analyse the reporter activity of SLC22A1, SLC22A4, SLC22A5 transporters, the cloned 5' and intron regions, DNA fragments were subcloned into the pGL3-Basic vector in front of a luciferase reporter gene and the constructs were named according to the gene and region amplified. The 5'-end were numbered +1 as the transcription site, according to the UCSC Genome Browser (accessed on 22 February 2023) (Lee et al., 2021).

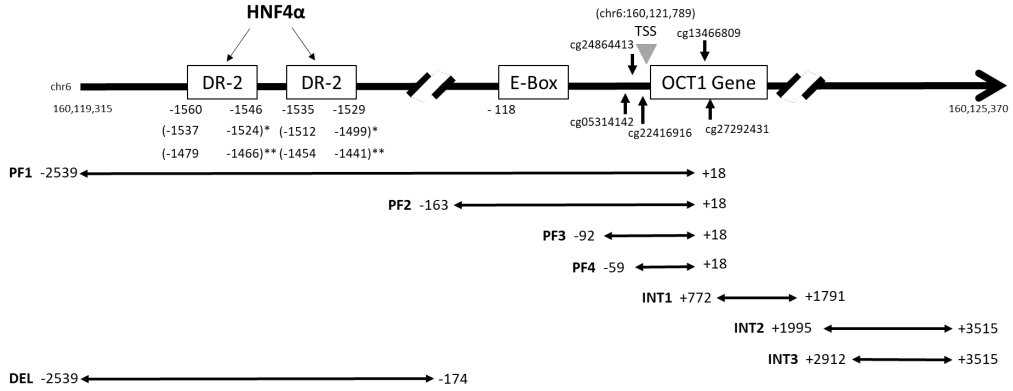


*Figure 5-3. Schematic representation of the SLC22A1 gene control region. Position is relative to the transcription start site (TSS, obtained from UCSC, accessed on 23 February 2023). In parentheses are positions reported by *(Kajiwara et al., 2008) ** (Saborowski et al., 2006).*

To identify the minimal promoter region for transcriptional activity of SLC22A1, seven different regulatory fragments ranging in size were amplified by PCR and amplified into the pGL3 plasmid, map can be found in the Appendix section. Luciferase Reporter system was used to analyse the reporter activity of SLC22A1 plasmid in HEK293, A549, and Calu-3 cells and co-transfected with the reference Rluc plasmid. Regions are illustrated in Figure 5-3. The pGL3-A1_PF1 construct contains the upstream sequence covering the DR-2 binding sites and putative E-Box (5'-CACGTG-3'), the pGL3-A1_PF2 to PF4 constructs

cover different sizes of the elements, such as putative E-box and upstream sequence of SLC22A1 gene.

The position of each promoter fragment (PF) and intronic region (INT) is given in relation to the transcription start site (TSS). Results showed that the pGL3-A1_PF1 fragment (-2539:+18) had a significant activity in A549 cells (3.528 ±1.180, $p \leq 0.001$) whereas HEK293 and Calu-3 showed a lower activity, with 1.236 ± 0.335 and 1.614 ± 0.355, respectively, when compared to background.

pGL3-A1_PF2 fragment (-163:+18) showed a significant activity in HEK293 cells (2.943 ± 0.714, $p \leq 0.001$) but a lower activity in A549 cells (1.096 ± 0.267) and Calu-3 cells (0.349 ± 0.063). Whereas, pGL3-A1_PF3 fragment (-92:+18) showed no significant differences in reporter activity among the three cell lines, with HEK293 cells showing an activity of 0.912 ± 0.199, A549 cells showing an activity of 1.426 ± 0.372, and Calu-3 cells showing an activity of 0.286 ± 0.045. pGL3-A1_PF4 fragment (-59:+18) showed a significant increase in activity in A549 cells (2.579 ± 0.930, $p \leq 0.05$) compared to HEK293 cells (0.42 ±0.126) and a slight increase in activity in Calu-3 cells (1.303 ±0.597).

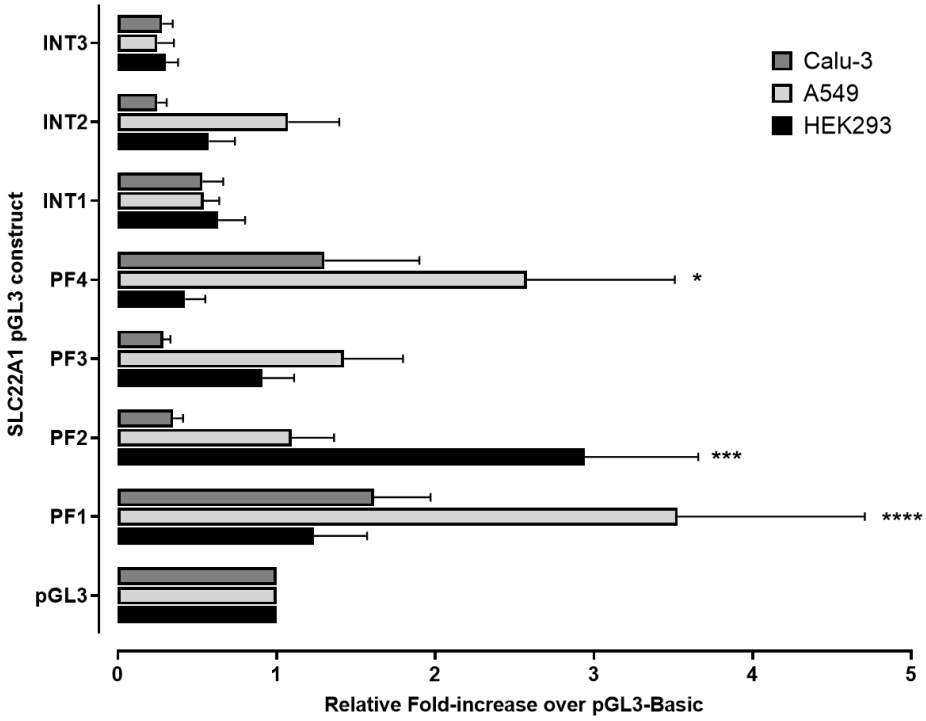


Figure 5-4. Analysis of the organic cation transporter 1, SLC22A1 promoter in A549, HEK293 and Calu-3 cells. A series of promoter constructs were transfected into A549 cells for luciferase assay. Firefly luciferase activity was normalised to Renilla luciferase activity. Data is represented as the relative fold increase compared with pGL3-Basic and are the means ± SEM for replicated.

The intronic regions, INT1 (+772: +1791), INT2 (+1995:+3515), and INT3 (+2912:+3515), showed no significant differences in reporter activity among the three cell lines. HEK293 cells showed an activity of 0.633 ± 0.170, 0.575 ± 0.163, and 0.303 ± 0.075 for pGL3- A1_INT1, pGL3- A1_INT2 and pGL3- A1_INT3, respectively. A549 cells showed a luciferase activity of 0.540 ± 0.099, 1.074 ± 0.321, and 0.250 ± 0.105 for pGL3-A1_INT1, to pGL3- A1_INT3, respectively. Lastly, Calu-3 cells showed an activity of 0.534 ± 0.130, 0.250 ± 0.059, and 0.280 ± 0.066 for pGL3-A1_INT1, pGL3- A1_INT2, and pGL3- A1_INT3, respectively. These results provide insights into the reporter

activity of the SLC22A1 plasmid in different cell lines, which may have implications for the regulation of the gene and its potential therapeutic targets.

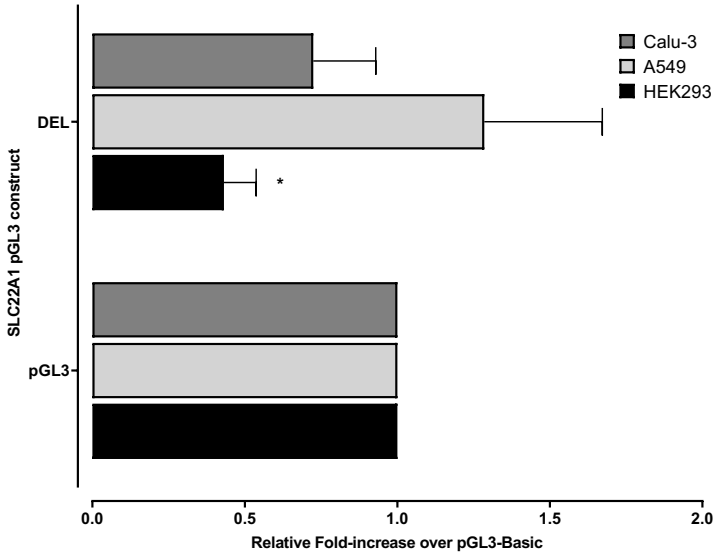


Figure 5-5. Analysis of the organic cation transporter 1, SLC22A1 promoter in A549, HEK293 and Calu-3 cells. Plasmid was prepared by directed mutagenesis to keep the region corresponding to the DR-2 elements. Firefly luciferase activity was normalised to Renilla luciferase activity. Data is represented as the relative fold increase compared with pGL3-Basic and are the means ± SEM for replicated.

To investigate the importance of the DR-2 elements in the SLC22A1 promoter, was analysed using directed mutagenesis to preserve the region corresponding to the DR-2 elements. The results of this analysis are presented as follows: DEL (-2539 to -174) in HEK293 was 0.432 ± 0.105 ($p<0.05$), in A549 was 1.285 ± 0.387, and in Calu-3 was 0.724 ± 0.205. Results are shown in Figure 5-5.

Partial mapping of the minimal SLC22A4 regulatory elements

Three different regulatory fragments of varying sizes were amplified by PCR in order to identify the minimal promoter region of SLC22A4. These fragments

were subsequently cloned into the pGL3 plasmid, and co-transfected with the reference *Rluc* plasmid into HEK293, A549, and Calu-3 cells.

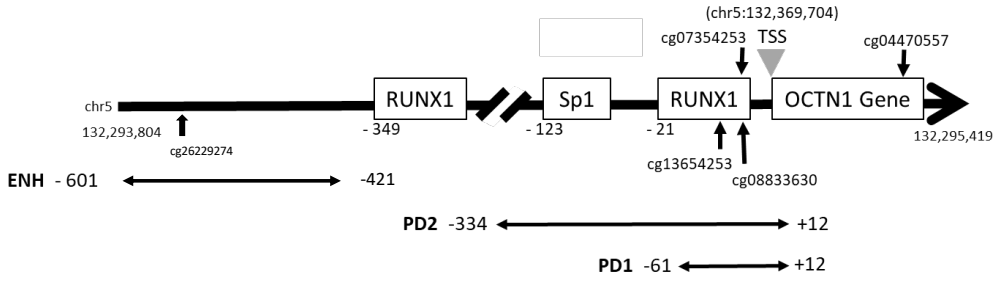


Figure 5-6. Schematic representation of the SLC224 gene control region. Position is relative to the transcription start site (TSS, obtained from UCSC, accessed on 23 February 2023).

According to Figure 5-6, the three constructs, pGL3-A4_PD1, pGL3-A4_PD2, and pGL3-A4_ENH, cover different regions of the SLC22A4 promoter. PD1 construct spans the region from -61 to +12 relative to TSS, which covers the RUNX1 binding site. pGL3-A4_PD2 contains the Sp1 and RUNX1 binding sites and spans the region from -334 to +12 relative to the TSS. ENH construct covers the upstream sequence of the second RUNX1 located at -349 bp from the TSS.

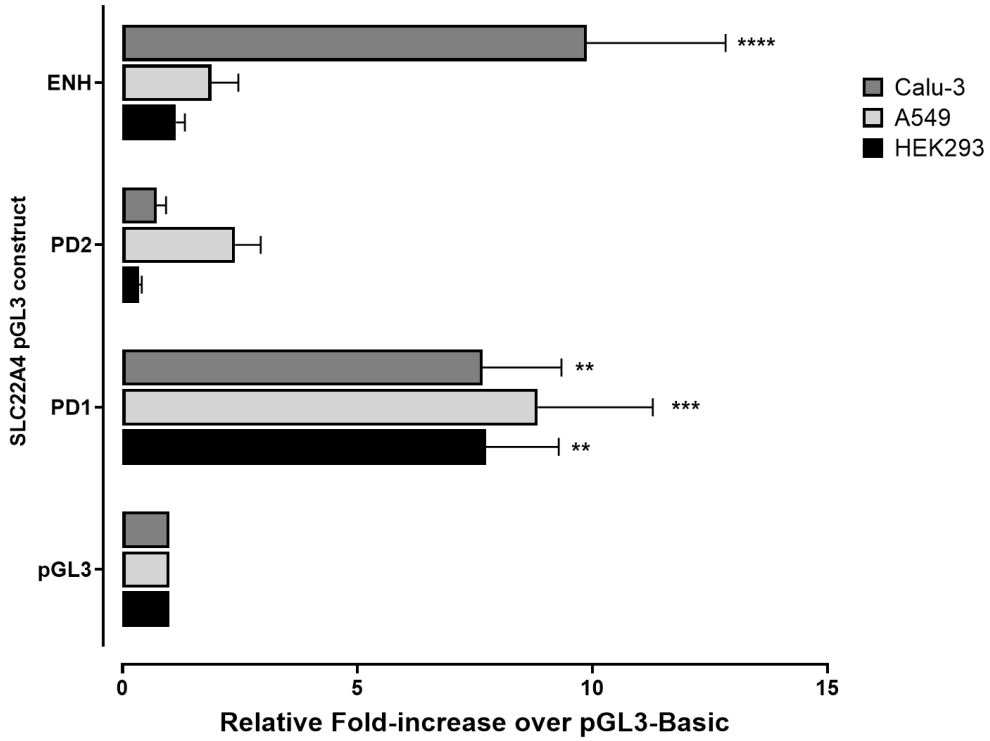


Figure 5-7. Analysis of the SLC22A4 promoter in A549, HEK293 and Calu-3 cells. Firefly luciferase activity was normalised to Renilla luciferase activity. Data is represented as the relative fold increase compared with pGL3-Basic and are the means ± SEM for replicated.

Results revealed that pGL3-A4_PD1 (-61 to +12) exhibited significant reporter activity in HEK293 with a mean value of 7.737± 1.543 ($p<0.01$), in A549 with a mean value of 8.835 ± 2.448 ($p<0.005$), and in Calu-3 with a mean value of 7.664 ± 1.674 ($p<0.01$). pGL3-A4_PD2 (-334 to +12) showed low reporter activity in HEK293 with a mean value of 0.351 ± 0.058, whereas in A549, it exhibited a higher reporter activity with a mean value of 2.392± 0.550, and in Calu-3, the reporter activity was 0.724 ± 0.199. pGL3-A4_ENH (-601 to -421) exhibited moderate reporter activity in HEK293 with a mean value of 1.138 ± 0.188, whereas in A549, it had a higher reporter activity with a mean value of 1.890 ± 0.577, and in Calu-3, the reporter activity was significantly higher with a mean value of 9.876 ± 2.955 ($p<0.0001$). Data is presented in Figure 5-7.

Partial mapping of the minimal SLC22A5 regulatory elements

In order to determine the minimal promoter region of SLC22A5, two regulatory fragments of different lengths were amplified by PCR and cloned into the pGL3 plasmid. Co-transfection of these constructs with the reference Rluc plasmid was performed in HEK293, A549, and Calu-3 cells. According to Figure 5-8, the pGL3-A5_P3 construct covers the upstream sequence of the SLC22A5 gene (-505 to -289), where the pGL3-A5_INT (+8524 to +9562) construct contains the ERE and CRE binding sites located in the first intron.

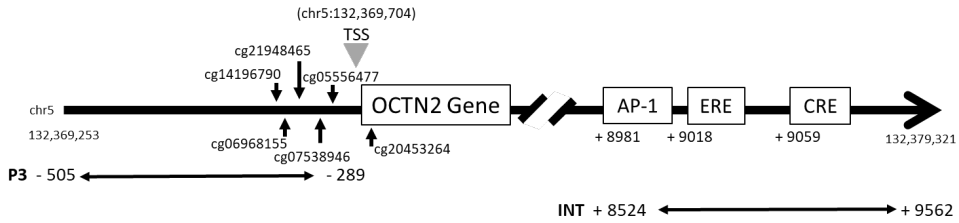


Figure 5-8. Schematic representation of the SLC22A5 gene control region. Position is relative to the transcription start site (TSS, obtained from UCSC, accessed on 23 February 2023).

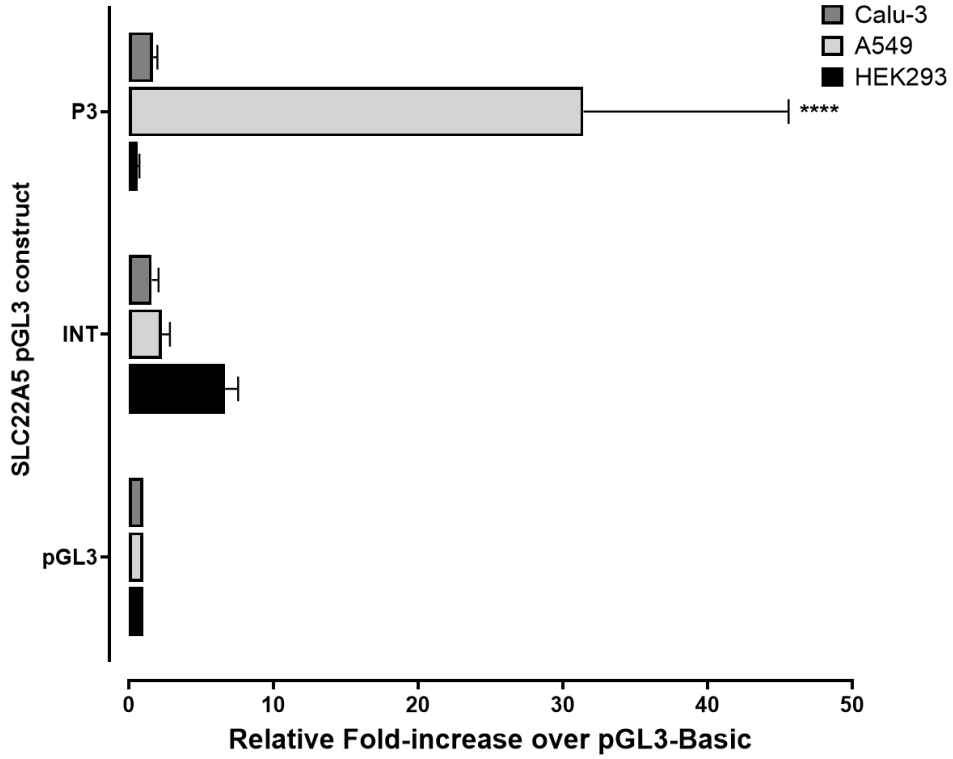


Figure 5-9. Analysis of the organic cation transporter 2 (SLC22A5) promoter in A549, HEK293 and Calu-3 cells. Firefly luciferase activity was normalised to Renilla luciferase activity. Data is represented as the relative fold increase compared with pGL3-Basic and are the means ± SEM for replicated.

The reporter activity of SLC22A5 plasmid was analysed in HEK293, A549, and Calu-3 cells using the Luciferase Reporter system and co-transfected with the reference *Rluc* plasmid. The results are presented in Figure 5, and the relative positions relative to the transcription start site (TSS) and all expression profiles were compared to the empty pGL3 plasmid, indicated in parentheses. The results showed that the activity of the promoter region pGL3-A5_INT (-505 to -289) in HEK293 was 2.28 ± 0.55, in A549 was 1.58 ± 0.46, and in Calu-3 was 6.65 ± 0.90. Furthermore, the promoter region pGL3-A5_P3 (+8524 to +9562) demonstrated a significant difference in HEK293 with a value of 31.42 ± 14.18,

while A549 showed a significantly lower value of 1.67 ± 0.29 ($p<0.0001$), and Calu-3 had a value of 0.6 ± 0.13.

5.3. Discussion

Emerging evidence suggests that DNA methylation plays a role in the regulation of transporter expression, including organic cation transporters (OCTs) Schaeffeler et al. (2011) have reported that DNA methylation of SLC22A1 is associated with the downregulation of SLC22A1 in hepatocellular carcinoma (HCC). there are no further studies assessing DNA methylation in SLC transporters in lung cancer to compare with the present study. mRNA was previously examined in Chapter 4 in Figure 4-23.

Presented findings revealed a negative correlation between DNA methylation at specific CpG sites (cg13466809 and cg05314142) and SLC22A1 RNA-seq expression levels in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) tissues. These observations suggest a potential role for DNA methylation in regulating SLC22A1 expression. However, further studies are required to definitively establish a cause-and-effect relationship.

Previous research demonstrated that hypermethylation of these specific CpG sites in the 5'-UTR and first exon of SLC22A1 is associated with decreased expression in hepatocellular carcinoma (HCC) Al-Abdulla et al., (2019). Interestingly, in our study, the luciferase reporter assay using a construct containing these CpG sites (SLC22A1_PF1) did not show a significant correlation with methylation status, suggesting that other regulatory elements within this region or additional factors might be involved in lung cancer cells.

The promoter region of SLC22A1_PF1 encompasses DR-2 elements and E-Box motifs, which have been implicated in SLC22A1 regulation by other studies.

Luciferase reporter assays using various SLC22A1 promoter fragments revealed a significant increase in activity with a specific region (-2539 +18) in A549 cells. This region encompasses previously identified DR-2 and E-Box elements, which could potentially be involved in SLC22A1 regulation. Interestingly, the observed increase in luciferase activity did not correlate with the methylation status of the CpG sites within this region. These findings suggest the potential involvement of transcription factors, such as HNF-4α, interacting with these elements in A549 cells. The aforementioned region encompasses three CpG sites, namely cg05314142, cg22416916, and cg24864413, along with two DR-2 and E-Box elements upstream of the transcription start site (TSS). Interestingly, the observed increase in luciferase activity did not correlate with the methylation status of the CpG sites within this region. DR-2 elements were identified by Saborowski et al., (2006) to be transactivated by the human nuclear factor 4α (HNF-4α) in the human hepatoma cell line Huh7. Functional importance of DR-2 elements was demonstrated through site-directed mutagenesis.

The Human Protein Atlas data indicates a higher expression of HNF-4α, a transcription factor known to interact with DR-2 elements, in A549 cells compared to Calu-3 cells 3 (20.8 TPM and 10.2 TPM, respectively) (Sjöstedt et al., 2020)(visited on 08 March 2023). This observation aligns with our findings of higher luciferase activity in A549 cells transfected with SLC22A1_PF1.

Additionally, the presence of an E-Box element within the construct could explain the increased expression observed in HEK293 cells, as E-Boxes are known to interact with basal transcription factors and regulate SLC22A1 expression. HNF-4α has previously been implicated in regulation of other SLC transporters by Popowski et al., (2005). The E-Box motif is known to interact with the basal transcription factors, upstream stimulating factors (USFs), and has been shown to regulate SLC22A1 expression in liver cells, as demonstrated in previous research (Bokelmann et al., 2018; Kajiwara et al., 2008).

Methylation of the CpG sites located centrally within the E-Box motif can significantly inhibit the formation of the transcription factor complex and negatively regulate gene expression (Prendergast & Ziff, 1991; Watt & Molloy, 1988). However, methylation status of the E-Box motif in the SLC22A1 promoter region was found to be similar in both liver and kidney tissues, as reported by Aoki et al., (2008). However, as main CpG islands were not included, it might prove the negative regulation in HEK293. Plasmids (pGL3-A1_PF1to pGL3- A1_PF4) harbour three CpG islands that are situated downstream of the E-box and may not exert an effect on regulation. In addition, A549 also showed higher expression with pGL3- A1_PF4, which can be attributed to the lack of CpG islands in the cloned sequence.

A prior study by O'Brien et al., (2013) indicated a binding motif for the hepatocyte nuclear factor 1 (HNF1) located within an evolutionary conserved region (ECR) of intron 1 in the SLC22A1 gene plays a role in the robust expression in the liver. Nonetheless, analysis of the intronic region of SLC22A1 using a luciferase reporter assay revealed no significant variance in expression

across the three cell lines under investigation. The present investigation proposes a synergistic relationship between DNA methylation and transcription factors in governing the expression of the SLC22A1 gene in pulmonary cells. It is possible that additional regulatory elements or cell-type specific factors contribute to SLC22A1 expression in lung cancer.

DNA methylation in SLC22A4 revealed contrasting correlations with disease status depending on the CpG site and lung cancer subtype. While some CpG sites exhibited negative correlations with lung cancer, suggesting a potential tumor-suppressive role, others showed positive correlations. These findings warrant further investigation to understand the complex interplay between DNA methylation and SLC22A4 expression in lung cancer development.

Similar to SLC22A1, luciferase reporter assays for SLC22A4 using constructs encompassing different promoter regions demonstrated varying activity across cell lines. Methylation analysis of SLC22A4 showed that in lung squamous cell carcinoma, the CpGs located in TSS150, TSS200, and the body exhibited negative correlations with the disease. Among these CpGs, cg26229274 located in TSS1500 showed a positive correlation with lung squamous cell carcinoma. On the other hand, in lung adenocarcinoma, the CpGs located in 1stExon, 5'UTR, TSS200, and TSS1500 showed negative correlations with the disease. Interestingly, cg04470557 located in the body of SLC22A4 showed a positive correlation with lung adenocarcinoma.

Reporter activity of SLC22A4 was investigated using three luciferase constructs, pGL3-A4_PD1, pGL3-A4_PD2, and pGL3-A4_ENH, which cover different regions of the promoter. pGL3-A4_PD1exhibited significant reporter

activity in all three cell lines, while pGL3-A4_PD2 showed low reporter activity in HEK293 but higher reporter activity in A549. pGL3-A4_ENH showed moderate reporter activity in HEK293 and A549 but significantly higher reporter activity in Calu-3.

This suggests that different regions of the SLC22A4 promoter may have cell-type specific regulatory functions. Previous research has identified RUNX1 as a transcription factor involved in SLC22A4 promoter activity. Maeda et al., (2007) found that transcription factor, RUNX1, is involved in regulating the activity of the SLC22A4 promoter through a luciferase-reporter gene assay and gel shift assay. Expression of SLC22A4 has been shown to be regulated by various factors, including RUNX1 and, inflammatory cytokines, and nuclear factor-kappaB (NF-κB), and inflammatory cytokines such as interleukin-1beta and tumour necrosis factor-alpha (Tokuhiro et al., 2003). Both pGL3-A4_ENH and pGL3-A4_PD1 contain a RUNX1 binding site. pGL3- A4_PD2 also contained the RUNX1 binding site, however, expression stayed basal compared to empty vector. RUNX1 transcription factor can act as both a transcriptional activator and repressor by interacting with various cofactors.

Findings suggest that DNA methylation plays a critical role in the regulation of SLC22A5 expression in both LUAD and LUSC. A negative correlation was found between DNA methylation levels of certain CpG sites in the TSS1500 region and gene expression, which suggests that hypermethylation of these CpG sites may contribute to the downregulation of SLC22A5 expression. The TSS1500 region is located upstream of the transcription start site and is known to be involved in the regulation of gene expression. Our results are consistent

with Chapter 3 that reported the downregulation of SLC22A5 in lung cancer, and present results suggest that DNA methylation may be one of the mechanisms responsible for this dysregulation.

Interestingly, DNA methylation levels of the CpG sites in the first exon (cg20453264) and the TSS200 region (cg05556477) did not show any significant correlation with SLC22A5 expression in both LUAD and LUSC samples. Suggesting that these CpG sites may not play a critical role in the regulation of SLC22A5 expression in lung cancer.

Results from promoter analysis demonstrated a significant ($p<0.001$) 30-fold increase in SLC22A5 expression in A549 cells with plasmid pGL3-A5_P3 (-505 to - 289) while no significant difference was observed in the intronic region compared to the empty vector. Although the higher mRNA expression of SLC22A5 in A549 and Calu-3 cells partially supports the observed increase in

expression with plasmid pGL3-A5_P3, further investigations are required to fully comprehend the underlying mechanisms involved. Wang et al., (2012) have previously studied the core promoter (-527 to -39) of SLC22A5 in breast cells which showed a relative activity of 17 compared to empty luciferase plasmid. Authors demonstrated that oestrogen can induce SLC22A5 expression in an oestrogen receptor (ER)-dependent manner. Although the proximal promoter of SLC22A5 was not found to be responsive to oestrogen. Knockdown was linked to suppression of cell proliferation, whereas, a previous analysis of oestrogen impact in NSLC showed that ERs might promote cell proliferation, migration, invasion and apoptosis in lung cancer (Gao et al., 2019); Novakova et al., 2022). The current study did not investigate the induction of SLC22A5 by

oestrogen receptor (ER) in lung cells. Therefore, further investigation is needed to fully elucidate the potential role of ER in the induction of SLC22A5 expression in lung cells.

In a recent investigation, analysis of activation of the SLC22A5 in the murine skeletal muscle cells (C2C12 myoblasts) demonstrated an interaction between the peroxisome proliferator-activated receptor-gamma coactivator (PGC-1α) and the myocyte enhancer factor 2 (MEF2) binding sites in the promoter region (Novakova et al., 2022). Present findings also suggest the involvement of PGC1α in SLC22A5 activation, as PGC1α levels in A549 cells have been shown to be much higher than other lung cancer cell lines (Oh et al., 2021).

Overall, the study suggests that DNA methylation may play a critical role in the regulation of SLC22A1 expression in lung cancer, with certain CpG sites exhibiting strong associations with RNAseq expression levels. SLC22A1 gene has exhibited hypermethylation in both LUAD and LUSC, as evidenced by β values exceeding 0.5. However, further elucidation is necessary to fully understand the underlying mechanisms of this interplay.

5.4. Conclusion

This study investigated the relationship between DNA methylation and the expression of SLC22A1, SLC22A4, and SLC22A5 in lung cancer. Current investigation identified potential associations between DNA methylation at specific CpG sites and mRNA expression levels for all three transporters. Additionally, luciferase reporter assays provided insights into potential regulatory elements within the SLC22A1 and SLC22A5 promoter regions.

However, while these findings offer promising leads, further research is necessary to definitively establish the regulatory mechanisms involved. Our study primarily focused on DNA methylation as a potential regulator, but the interplay with other factors like transcription factors requires more in-depth investigation. Additionally, the observed correlations with DNA methylation need functional validation to confirm their causal role in SLC transporter expression.

Furthermore, the study identified potential roles for HNF-4α and RUNX1 in SLC22A1 and SLC22A4 regulation, respectively. Further studies focusing on these transcription factors could provide valuable insights. In conclusion, this study lays the groundwork for future investigations into the complex regulatory mechanisms governing SLC transporter expression in lung cancer.

CHAPTER 6. CONCLUDING REMARKS

CHAPTER 6. CONCLUDING REMARKS AND FUTURE WORK

The lung epithelia represent a unique environment in terms of the composition of the extracellular matrix, the presence of specialized cells such as ciliated cells and mucus-secreting cells, and the exposure to inhaled substances. SLC22A transporters belong to the solute carrier (SLC) family of membrane transport proteins and play a crucial role in the transport of endogenous and exogenous compounds across cellular membranes. An association exists between polymorphisms in the genes encoding SLC22A4 and SLC22A5 transporters and an extended time to progression in patients with unresectable gastrointestinal stromal tumours (GIST) treated with imatinib therapy (Angelini et al., 2013).

In the lung epithelia, SLC22A transporters are involved in the uptake and elimination of a variety of substrates, including drugs, toxins, and metabolites (Selo et al., 2020). Therefore, the expression and function of SLC22A transporters may differ from those observed in other tissues, and may be regulated by specific factors such as environmental pollutants or microbial products. Understanding the mechanisms that regulate the expression and function of these transporters in the lung epithelia may offer new insights into their role in lung diseases and provide novel therapeutic targets for the treatment of these conditions. In addition, alterations in SLC22A expression and function have been implicated in several lung diseases, including chronic obstructive pulmonary disease (COPD), asthma, and cancer. For example, studies have shown that the expression of SLC22A transporters is dysregulated in COPD, which may contribute to the altered metabolism and signalling observed in this

disease. Similarly, alterations in SLC22A expression have been associated with the pathogenesis of asthma, and targeting these transporters may offer novel therapeutic opportunities for the treatment of this disease.

Furthermore, recent studies have highlighted the importance of SLC22A transporters in the development and progression of lung cancer (Brecht et al., 2020). Dysregulation of these transporters can lead to altered cellular uptake of organic cations, resulting in changes in metabolic pathways and downstream signaling events in cancer cells. Therefore, understanding the role of SLC22A transporters in lung cancer may offer new avenues for targeted therapy.

The aim of the present study was to explore the physiological functions of OCT/OCNs transporters in the human bronchial epithelium. Specifically, the investigation was aimed at elucidating the involvement of these transporters in cellular proliferation and survival, as well as characterizing the mechanisms governing the transcription of SLC22A1, SLC22A4, and SLC22A5 in lung epithelial cells.

6.1. Synopsis of findings/ Summary of Observations

The ensuing pages will undertake a critical review of the diverse observations recorded during the course of this investigation and assess their significance from a comprehensive viewpoint.

6.1.1. Downregulation of SLC22A1, SLC22A4 and SLC22A5 in lung malignancies and severe respiratory diseases.

Prior research highlighted a connection between elevated SLC gene expression and cancer, as a response to the augmented nutrient requirements and waste elimination demands of cancer cells (Ganapathy et al., 2009). In more recent studies, specific genes from the SLC22 family have been suggested to act as potential tumour suppressors, such as SLC22A1 in cholangiocellular carcinoma (Lautem et al., 2013) and SLC22A7 in hepatocellular carcinoma (Yasui et al., 2014). Chronic Obstructive Pulmonary Disease (COPD) and asthma are both chronic inflammatory lung diseases that affect millions of people worldwide. Lung cancer is one of the leading causes of cancer-related deaths worldwide. The two main histological subtypes of lung cancer are adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC). In this study, mRNA expression levels of SLC22A1 , SLC22A4 and SLC22A5 transporters were analysed in COPD, LUSC, LUAD and asthma from publicly available datasets to investigate their potential relationship in disease progression.

The results of described in Chapter 3 showed that SLC22A5 mRNA expression was significantly increased in COPD patients. Findings suggest that alterations in the expression of SLC22As transporters may play a role in the pathogenesis of COPD, as previous studies have highlighted the contribution in inflammatory diseases (Koepsell, 2020). Specifically, the upregulation of SLC22A5 may indicate an increased demand for carnitine transport in COPD patients, as carnitine is essential for mitochondrial metabolism and energy production in cells. On the other hand, no significant differences were observed in

OCT/OCTN mRNA expression levels between healthy and asthmatic tissue samples. These results suggest that the expression of OCT/OCTN transporters may not be a major factor in the development or progression of asthma.

However, it is important to note that this study only analysed mRNA expression levels, and further investigations are needed to fully understand the role of OCT/OCTN transporters in asthma pathogenesis and Single-cell RNA sequencing (scRNA-seq) might shed a light into the different expression profiles found in the lung epithelia.

Findings suggest that SLC22A4 may play a role in the development and progression of lung cancer. The downregulation of SLC22A4 could potentially result in reduced uptake of organic cations, which may contribute to altered metabolism and signalling in cancer cells.

In addition, results showed a statistically significant downregulation of SLC22A4 and SLC22A5 in LUSC samples, further supporting the role of these genes in the progression of lung cancer. The observed downregulation of SLC22A4 and SLC22A5 could potentially result in altered cellular uptake of organic cations, leading to changes in metabolic pathways and downstream signalling events in LUSC cells. Interestingly, although SLC22A1 showed a broader expression profile in LUSC samples, the difference was not statistically significant. This suggests that SLC22A1 may not be as important in the development and progression of LUSC as SLC22A4 and SLC22A5.

Overall, findings described in Chapter 3 contribute to a better understanding of the molecular mechanisms underlying the development and progression of COPD. More importantly, SLC22 family genes, particularly SLC22A4 and

SLC22A5, may be involved in the pathogenesis of adenocarcinoma and LUSC. Future studies could investigate the functional implications of these transporter alterations and their potential as therapeutic targets for these respiratory diseases.

6.1.2. Potential effects on cell proliferation, adhesion and migration in SLC22A1 , SLC22A4 and SLC22A5 genes

The thesis aimed to investigate the role of OCT/OCTNs transporters in cellular physiology using CRISPR/Cas9, nCas9, and RNP complexes. These gene-editing tools were applied to HEK293 cells and two lung cell lines, A549 and Calu-3, with the objective of inactivating the target transporters. However, the efficiency of inactivating the transporters was found to be relatively low in the lung cells. However, the gene editing efficiency was found to be relatively low in the lung cells.

Gene editing of OCT/OCTNs in HEK293 resulted in slower proliferation rates, decreased migration, and reduced adhesion. These findings align with previous studies by Mukherjee, (2012) where reduction of SLC22A1 and SLC22A5 expression by siRNA reduced proliferative activity, therefore affecting wound repair mechanisms. According to the literature, SLC22A5 expression has been positively associated with tumour size in breast cancer (dinarvand et al., 2023). A previous study conducted by Wang et al., (2012) revealed that the growth of breast cancer cell lines was significantly impeded upon the knockdown of SLC22A5, indicating the fundamental role of SLC22A5 in cell proliferation. Similarly, Fink et al., (2019) demonstrated that the loss of LN18 cell viability was observed upon siRNA-mediated silencing of SLC22A5.

Nakamura et al., (2007) reported that inhibiting SLC22A4 in K562 cells resulted in a reduction in growth rate and significantly impacted erythroid differentiation. The precise mechanism remains unclear, but it's thought that the decrease in growth could be caused by cell cycle inhibition or induction of apoptosis.

These findings suggest that SLC22A4 and SLC22A5 may be involved in transporting physiological compounds essential for cell proliferation. Inhibiting SLC22A5 could disrupt carnitine transport, leading to abnormal lipid metabolism, decreased cellular energy, and reduced cell proliferation.

Given the vital role of physiological compounds in cell proliferation, it is hypothesized that the SLC22A family of transporters, particularly OCT/OCTNs, plays a crucial role in facilitating the transport of these compounds. This is particularly relevant for lung epithelia, where proper transporter function is essential for maintaining cellular homeostasis.

This study has laid the groundwork for further investigation into the functional roles of OCT/OCTN transporters in lung cancer. To overcome the limitations encountered with CRISPR/Cas9 ribonucleoprotein (RNP) complexes, particularly the relatively low editing efficiency in lung cancer cell lines, future studies could explore alternative gene editing approaches. Lentiviral delivery of CRISPR machinery offers a potentially more efficient and stable method for inactivating target genes. A recent study by Biering et al., (2022) successfully employed lentiviral Cas9-Blast and dCAS–VP64_Blast vectors for functional studies in Calu-3 lung cancer cells.

In addition to enhanced editing efficiency, lentiviral CRISPR would also enable the generation of stable cell lines with inactivated OCT/OCTN transporters. This

would allow for more comprehensive analysis of the long-term effects of transporter inactivation on cell physiology, migration, adhesion, and response to chemotherapeutic agents.

6.1.3. Transcriptional regulation of SLC22A1, SLC22A4 and SLC22A5

Data from Chapter 5 showed analysis of expression regulation of SLC22A1, SLC22A4, and SLC22A5 in the lungs is currently scarce. Comprehending the impact of regulatory mechanisms on the function of pulmonary transporters is of utmost importance in the development of therapeutics for respiratory illnesses. Salani et al., (2012) found that metformin can increase the expression of SLC22A1 in a dose-dependent manner in lung cell lines. A previous study by Mukherjee et al., (2017) showed upregulation of SLC22A1, SLC22A4 and SLC22A5 after LPS exposure However, studies have failed to explain the underlying mechanisms. The findings of this study provide valuable insights into the relationship between DNA methylation at specific CpG sites and the expression levels of OCT/OCTNs in lung cancer.

In LUSC, SLC22A1 exhibited hypermethylation, while SLC22A4 and SLC22A5 showed hypomethylation. Additionally, the methylation status of OCT/OCTNs demonstrated an inverse relationship with their RNA expression profiles in LUSC. The present finding supports previous investigations where expression of OCT/OCTNs transporters have been found to be negatively correlated to their methylation status in different cancer cells (Pochini et al., 2022; Qu et al., 2013; Schaeffeler et al., 2011).

Chapter 5 provided evidence of a potential interplay between two DR-2 elements located in the SLC22A1 gene, and CpG islands situated within the first exon (cg13466809 and cg27292431), which suggests cooperative activation of the promoter by HNF-4α in the A549 cell line. This observation may be partially explained by the higher expression of HNF-4α in A549 cells compared to other cell lines, as reported by *The Human Protein Atlas* (visited on 08 March 2023). This study supports evidence from previous observations (Aoki et al., 2008) where methylation of the proximal promoter region of SLC22A1 is unlikely to play a role in the tissue-specific expression of this gene. However, the plasmid containing the E-Box element exhibited increased expression in HEK293 cells. In contrast, intronic regions displayed no significant differences in reporter activity among the three cell lines.

Reporter activity of SLC22A4 suggests that the different regions have varying degrees of reporter activity in different cell lines. These results suggest that the expression of SLC22A1 might be regulated by RUNX1. This transcription factor can act as both a transcriptional activator and repressor by interacting with various cofactors, which could explain the difference in reporter activity in the three cell lines.

In the luciferase reporter assay for SLC22A5, A549 showed an increased reporter activity, compared to the other cell lines. mRNA expression in Chapter 4 indicated this cell line with a higher expression. According to the literature, promoter region of SLC22A5 contains the myocyte enhancer factor 2 (MEF2) binding sites. Activation of the SLC22A5 gene promoter in mouse skeletal muscle cells was previously demonstrated by an interaction between the

peroxisome proliferator-activated receptor-gamma coactivator (PGC-1α) and MEF2 binding sites in the promoter (Novakova et al., 2022). Higher expression of SLC22A5 can be explained as A549 has shown to have a higher expression of PGC1α than other lung cancer cell lines (Oh et al., 2021).

6.2. Future Work

Due to the clinical significance of SLC transporters in drug absorption and future development of inhaled drugs, future studies could aim to use lentiviral systems as a means of delivering CRISPR-Cas9 to these cells to further explore the regulation of SLC transporters in lung epithelial cells. According to recent research, CRISPR lentiviral systems have emerged as a prominent contender for in vivo genome editing delivery and have demonstrated significant efficacy in lung adenocarcinoma cells (B. Li et al., 2020; R. Park et al., 2022). This approach may provide a more efficient method of genetic manipulation and lead to the development of cell lines with specific gene knockouts or knock-ins.

While RNA-seq and microarrays for methylation analysis of expression and methylation profiles of specific proteins represents a significant step forward in understanding the role of methylation in gene regulation, there are several limitations to the study. One limitation is the inability to capture methylation data for all CpG sites in the promoter region due to the constraints of the project timeline. This could potentially lead to incomplete information and incomplete understanding of the role of methylation in the regulation of these proteins. Furthermore, a deeper understanding of the underlying regulatory mechanisms governing these genes in lung epithelia cells could be obtained with additional techniques such Electrophoretic mobility shift assay (EMSA) to identify

transcription factors in lung epithelial cells, co-transfection with transcription factors and luciferase plasmids prepared in the present study to confirm activation. Single-cell RNA sequencing (scRNA-seq), for example, can uncover regulatory relationships between transporters within different lung cell types; in contrast to bulk RNAseq, where gene expressions are average across large population of cells.

REFERENCES

Aguirre, A. J., Meyers, R. M., Weir, B. A., Vazquez, F., Zhang, C.-Z., Ben-David, U., Cook, A., Ha, G., Harrington, W. F., Doshi, M. B., Kost-Alimova, M., Gill, S., Xu, H., Ali, L. D., Jiang, G., Pantel, S., Lee, Y., Goodale, A., Cherniack, A. D., … Hahn, W. C. (2016a). Genomic copy number dictates a gene-independent cell response to CRISPR-Cas9 targeting. *Cancer Discovery*, *6*(8), 914–929. https://doi.org/10.1158/2159-8290.CD-16-0154

Aguirre, A. J., Meyers, R. M., Weir, B. A., Vazquez, F., Zhang, C.-Z., Ben-David, U., Cook, A., Ha, G., Harrington, W. F., Doshi, M. B., Kost-Alimova, M., Gill, S., Xu, H., Ali, L. D., Jiang, G., Pantel, S., Lee, Y., Goodale, A., Cherniack, A. D., … Hahn, W. C. (2016b). Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer Discovery*, *6*(8), 914–929. https://doi.org/10.1158/2159-8290.CD-16-0154

Al-Abdulla, R., Lozano, E., Macias, R. I. R., Monte, M. J., Briz, O., O'Rourke, C. J., Serrano, M. A., Banales, J. M., Avila, M. A., Martinez-Chantar, M. L., Geier, A., Andersen, J. B., & Marin, J. J. G. (2019). Epigenetic events involved in organic cation transporter 1-dependent impaired response of hepatocellular carcinoma to sorafenib. *British Journal of Pharmacology*, *176*(6), 787–800. https://doi.org/10.1111/bph.14563

Al-Jayyoussi, G., Price, D. F., Kreitmeyr, K., Keogh, J. P., Smith, M. W., Gumbleton, M., & Morris, C. J. (2015). Absorption of ipratropium and l

-carnitine into the pulmonary circulation of the ex-vivo rat lung is driven by passive processes rather than active uptake by OCT/OCTN transporters. *International Journal of Pharmaceutics*, *496*(2), 834–841. https://doi.org/10.1016/j.ijpharm.2015.10.036

Amo, V. L. D., Juste, S. S., & Gantz, V. M. (2022). A nickase Cas9 gene-drive system promotes super-Mendelian inheritance in Drosophila. *Cell Reports*, *39*(8), 110843. https://doi.org/10.1016/j.celrep.2022.110843

Andreev, E., Brosseau, N., Carmona, E., Mes-Masson, A.-M., & Ramotar, D. (2016). The human organic cation transporter OCT1 mediates high affinity uptake of the anticancer drug daunorubicin. *Scientific Reports*, *6*, 20508. https://doi.org/10.1038/srep20508

Angelini, S., Pantaleo, M. A., Ravegnini, G., Zenesini, C., Cavrini, G., Nannini, M., Fumagalli, E., Palassini, E., Saponara, M., Di Battista, M., Casali, P. G., Hrelia, P., Cantelli-Forti, G., & Biasco, G. (2013). Polymorphisms in OCTN1 and OCTN2 transporters genes are associated with prolonged time to progression in unresectable gastrointestinal stromal tumours treated with imatinib therapy. *Pharmacological Research*, *68*(1), 1–6. https://doi.org/10.1016/j.phrs.2012.10.015

Aoki, M., Terada, T., Kajiwara, M., Ogasawara, K., Ikai, I., Ogawa, O., Katsura, T., & Inui, K. (2008). Kidney-specific expression of human organic cation transporter 2 (OCT2/SLC22A2) is regulated by DNA methylation. *American Journal of Physiology-Renal Physiology*, *295*(1), F165–F170. https://doi.org/10.1152/ajprenal.90257.2008

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., & Soboleva, A. (2013). NCBI GEO: Archive for functional genomics data sets—update. *Nucleic Acids Research*, *41*(D1), D991–D995. https://doi.org/10.1093/nar/gks1193

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., & Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research*, *125*(1–2), 279–284. https://doi.org/10.1016/s0166-4328(01)00297-2

Berg, T., Hegelund Myrbäck, T., Olsson, M., Seidegård, J., Werkström, V., Zhou, X.-H., Grunewald, J., Gustavsson, L., & Nord, M. (2014). Gene expression analysis of membrane transporters and drug-metabolizing enzymes in the lung of healthy and COPD subjects. *Pharmacology Research & Perspectives*, *2*(4), e00054. https://doi.org/10.1002/prp2.54

Bicer, E. M., Forbes, B., Somers, G., Blomberg, A., Behndig, A. F., & Mudway, I. (2012). Characterizing The Composition Of Human Respiratory Tract Lining Fluids In Health And Disease. In *C58. ENVIRONMENTAL EXPOSURES: MECHANISMS* (1–315, pp. A4661–A4661). American Thoracic Society. https://doi.org/10.1164/ajrccm-conference.2012.185.1_MeetingAbstracts.A4661

Biering, S. B., Sarnik, S. A., Wang, E., Zengel, J. R., Leist, S. R., Schäfer, A., Sathyan, V., Hawkins, P., Okuda, K., Tau, C., Jangid, A. R., Duffy, C. V., Wei, J., Gilmore, R. C., Alfajaro, M. M., Strine, M. S., Nguyenla, X.,

Van Dis, E., Catamura, C., … Hsu, P. D. (2022). Genome-wide bidirectional CRISPR screens identify mucins as host factors modulating SARS-CoV-2 infection. *Nature Genetics*, *54*(8), 1078–1089. https://doi.org/10.1038/s41588-022-01131-x

Binz, R. L., Tian, E., Sadhukhan, R., Zhou, D., Hauer-Jensen, M., & Pathak, R. (2019). Identification of novel breakpoints for locus- and region-specific translocations in 293 cells by molecular cytogenetics before and after irradiation. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-47002-0

Bleasby, K., Castle, J. C., Roberts, C. J., Cheng, C., Bailey, W. J., Sina, J. F., Kulkarni, A. V., Hafey, M. J., Evers, R., Johnson, J. M., Ulrich, R. G., & Slatter, J. G. (2006). Expression profiles of 50 xenobiotic transporter genes in humans and pre-clinical species: A resource for investigations into drug disposition. *Xenobiotica*, *36*(10–11), 963–988. https://doi.org/10.1080/00498250600861751

Bokelmann, K., Brockmöller, J., & Tzvetkov, M. V. (2018). Impact of Promoter Polymorphisms on the Transcriptional Regulation of the Organic Cation Transporter OCT1 (SLC22A1). *Journal of Personalized Medicine*, *8*(4), 42. https://doi.org/10.3390/jpm8040042

Bolstad, B. M., Irizarry, R. A., Åstrand, M., & Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, *19*(2), 185–193. https://doi.org/10.1093/bioinformatics/19.2.185

Bosquillon, C. (2010). Drug transporters in the lung—Do they play a role in the biopharmaceutics of inhaled drugs? *Journal of Pharmaceutical Sciences*, *99*(5), 2240–2255. https://doi.org/10.1002/jps.21995

Bosquillon, C., Madlova, M., Patel, N., Clear, N., & Forbes, B. (2017). A Comparison of Drug Transport in Pulmonary Absorption Models: Isolated Perfused rat Lungs, Respiratory Epithelial Cell Lines and Primary Cell Culture. *Pharmaceutical Research*, *34*(12), 2532–2540. https://doi.org/10.1007/s11095-017-2251-y

Bourgon, R., Gentleman, R., & Huber, W. (2010). Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(21), 9546–9551. https://doi.org/10.1073/pnas.0914005107

Brecht, K., Schäfer, A. M., & Meyer zu Schwabedissen, H. E. (2020). Uptake Transporters of the SLC21, SLC22A, and SLC15A Families in Anticancer Therapy—Modulators of Cellular Entry or Pharmacokinetics? *Cancers*, *12*(8), Article 8. https://doi.org/10.3390/cancers12082263

Buelow, D. R., Anderson, J. T., Pounds, S. B., Shi, L., Lamba, J. K., Hu, S., Gibson, A. A., Goodwin, E. A., Sparreboom, A., & Baker, S. D. (2021). DNA Methylation-Based Epigenetic Repression of SLC22A4 Promotes Resistance to Cytarabine in Acute Myeloid Leukemia. *Clinical and Translational Science*, *14*(1), 137–142. https://doi.org/10.1111/cts.12861

Bur, M., & Lehr, C.-M. (2008). Pulmonary cell culture models to study the safety and efficacy of innovative aerosol medicines. *Expert Opinion on Drug Delivery*, *5*(6), 641–652. https://doi.org/10.1517/17425247.5.6.641

Bylund, L., Kytölä, S., Lui, W.-O., Larsson, C., & Weber, G. (2004). Analysis of the cytogenetic stability of the human embryonal kidney cell line 293 by cytogenetic and STR profiling approaches. *Cytogenetic and Genome Research*, *106*(1), 28–32. https://doi.org/10.1159/000078556

Camelo, A., Dunmore, R., Sleeman, M. A., & Clarke, D. L. (2014). The epithelium in idiopathic pulmonary fibrosis: Breaking the barrier. *Frontiers in Pharmacology*, *4*. https://doi.org/10.3389/fphar.2013.00173

Campeau, E., & Gobeil, S. (2011). RNA interference in mammals: Behind the screen. *Briefings in Functional Genomics*, *10*(4), 215–226. https://doi.org/10.1093/bfgp/elr018

Cervera, L., Gutiérrez, S., Gòdia, F., & Segura, M. M. (2011). Optimization of HEK 293 cell growth by addition of non-animal derived components using design of experiments. *BMC Proceedings*, *5*(Suppl 8), P126. https://doi.org/10.1186/1753-6561-5-S8-P126

César-Razquin, A., Snijder, B., Frappier-Brinton, T., Isserlin, R., Gyimesi, G., Bai, X., Reithmeier, R. A., Hepworth, D., Hediger, M. A., Edwards, A. M., & Superti-Furga, G. (2015). A Call for Systematic Research on Solute Carriers. *Cell*, *162*(3), 478–487. https://doi.org/10.1016/j.cell.2015.07.022

Chaudhry, R., & Bordoni, B. (2019). Anatomy, Thorax, Lungs. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK470197/

Cheah, I. K., & Halliwell, B. (2012). Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, *1822*(5), 784–793. https://doi.org/10.1016/j.bbadis.2011.09.017

Cheah, I. K., & Halliwell, B. (2021). Ergothioneine, recent developments. *Redox Biology*, *42*, 101868. https://doi.org/10.1016/j.redox.2021.101868

Chen, K. G., & Sikic, B. I. (2012). Molecular Pathways: Regulation and Therapeutic Implications of Multidrug Resistance. *Clinical Cancer Research*, *18*(7), 1863–1869. https://doi.org/10.1158/1078-0432.CCR-11-1590

Chen, L., Hong, C., Chen, E. C., Yee, S. W., Xu, L., Almof, E. U., Wen, C., Fujii, K., Johns, S. J., Stryke, D., Ferrin, T. E., Simko, J., Chen, X., Costello, J. F., & Giacomini, K. M. (2013). Genetic and epigenetic regulation of the organic cation transporter 3, SLC22A3. *The Pharmacogenomics Journal*, *13*(2), 110–120. https://doi.org/10.1038/tpj.2011.60

Chen, Z., Shi, T., Zhang, L., Zhu, P., Deng, M., Huang, C., Hu, T., Jiang, L., & Li, J. (2016). Mammalian drug efflux transporters of the ATP binding cassette (ABC) family in multidrug resistance: A review of the past

decade. *Cancer Letters*, *370*(1), 153–164. https://doi.org/10.1016/j.canlet.2015.10.010

Chiang, T.-W. W., le Sage, C., Larrieu, D., Demir, M., & Jackson, S. P. (2016). CRISPR-Cas9D10A nickase-based genotypic and phenotypic screening to enhance genome editing. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep24356

Chronic obstructive pulmonary disease (COPD). (n.d.). Retrieved 15 January 2023, from https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)

Ciarimboli, G., & Schlatter, E. (2005). Regulation of organic cation transport. *Pflügers Archiv*, *449*(5), 423–441. https://doi.org/10.1007/s00424-004-1355-5

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., Sabedot, T. S., Malta, T. M., Pagnotta, S. M., Castiglioni, I., Ceccarelli, M., Bontempi, G., & Noushmehr, H. (2016). TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Research*, *44*(8), e71. https://doi.org/10.1093/nar/gkv1507

Colas, C., Ung, P. M.-U., & Schlessinger, A. (2016). SLC Transporters: Structure, Function, and Drug Discovery. *MedChemComm*, *7*(6), 1069–1081. https://doi.org/10.1039/C6MD00005C

Courcot, E., Leclerc, J., Lafitte, J.-J., Mensier, E., Jaillard, S., Gosset, P., Shirali, P., Pottier, N., Broly, F., & Lo-Guidice, J.-M. (2012). Xenobiotic metabolism and disposition in human lung cell models: Comparison with

in vivo expression profiles. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, *40*(10), 1953–1965. https://doi.org/10.1124/dmd.112.046896

Cryan, S., Sivadas, N., & Garciacontreras, L. (2007). In vivo animal models for drug delivery across the lung mucosal barrier☆. *Advanced Drug Delivery Reviews*, *59*(11), 1133–1151. https://doi.org/10.1016/j.addr.2007.08.023

Cubbon, A., Ivancic-Bace, I., & Bolt, E. L. (2018). CRISPR-Cas immunity, DNA repair and genome stability. *Bioscience Reports*, *38*(5), BSR20180457. https://doi.org/10.1042/BSR20180457

Cui, Y., Xu, J., Cheng, M., Liao, X., & Peng, S. (2018). Review of CRISPR/Cas9 sgRNA Design Tools. *Interdisciplinary Sciences: Computational Life Sciences*, *10*(2), 455–465. https://doi.org/10.1007/s12539-018-0298-z

Cullot, G., Boutin, J., Toutain, J., Prat, F., Pennamen, P., Rooryck, C., Teichmann, M., Rousseau, E., Lamrissi-Garcia, I., Guyonnet-Duperat, V., Bibeyran, A., Lalanne, M., Prouzet-Mauléon, V., Turcq, B., Ged, C., Blouin, J.-M., Richard, E., Dabernat, S., Moreau-Gaudry, F., & Bedel, A. (2019). CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nature Communications*, *10*(1), 1136. https://doi.org/10.1038/s41467-019-09006-2

D'Argenio, G., Petillo, O., Margarucci, S., Torpedine, A., Calarco, A., Koverech, A., Boccia, A., Paolella, G., & Peluso, G. (2010). Colon

OCTN2 Gene Expression Is Up-regulated by Peroxisome Proliferator-activated Receptor γ in Humans and Mice and Contributes to Local and Systemic Carnitine Homeostasis. *The Journal of Biological Chemistry*, *285*(35), 27078–27087. https://doi.org/10.1074/jbc.M110.109678

DeGorter, M. K., Xia, C. Q., Yang, J. J., & Kim, R. B. (2012). Drug Transporters in Drug Efficacy and Toxicity. *Annual Review of Pharmacology and Toxicology*, *52*(1), 249–273. https://doi.org/10.1146/annurev-pharmtox-010611-134529

DeWitt, M. A., Corn, J. E., & Carroll, D. (2017). Genome editing via delivery of Cas9 ribonucleoprotein. *Methods*, *121–122*, 9–15. https://doi.org/10.1016/j.ymeth.2017.04.003

dinarvand, N., Karimi, F., Azizi, R., Rastaghi, S., Sheikhi, A., & Pourfarzam, M. (2023). Evaluation of the gene encoding carnitine transporter (OCTN2/SLC22A5) expression in human breast cancer and its association with clinicopathological characteristics. *Molecular Biology Reports*, *50*(3), 2061–2066. https://doi.org/10.1007/s11033-022-08152-z

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, *34*(2), 184–191. https://doi.org/10.1038/nbt.3437

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation. *Nature Biotechnology*, *32*(12), Article 12. https://doi.org/10.1038/nbt.3026

Drenberg, C. D., Gibson, A. A., Pounds, S. B., Shi, L., Rhinehart, D. P., Li, L., Hu, S., Du, G., Nies, A. T., Schwab, M., Pabla, N., Blum, W., Gruber, T. A., Baker, S. D., & Sparreboom, A. (2017). OCTN1 Is a High-Affinity Carrier of Nucleoside Analogues. *Cancer Research*, *77*(8), 2102–2111. https://doi.org/10.1158/0008-5472.CAN-16-2548

Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., Guo, L., & Zhang, Y. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. *Cell Research*, *24*(8), 1009–1012. https://doi.org/10.1038/cr.2014.87

Edemir, B. (2020). Identification of Prognostic Organic Cation and Anion Transporters in Different Cancer Entities by In Silico Analysis. *International Journal of Molecular Sciences*, *21*(12), 4491. https://doi.org/10.3390/ijms21124491

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Ehrhardt, C., Kneuer, C., Bies, C., Lehr, C.-M., Kim, K.-J., & Bakowsky, U. (2005). Salbutamol is actively absorbed across human bronchial

epithelial cell layers. *Pulmonary Pharmacology & Therapeutics*, *18*(3), 165–170. https://doi.org/10.1016/j.pupt.2004.11.007

Ehrhardt, C., Kneuer, C., Fiegel, J., Hanes, J., Schaefer, U. F., Kim, K.-J., & Lehr, C.-M. (2002). Influence of apical fluid volume on the development of functional intercellular junctions in the human epithelial cell line 16HBE14o-: Implications for the use of this cell line as an in vitro model for bronchial drug absorption studies. *Cell and Tissue Research*, *308*(3), 391–400. https://doi.org/10.1007/s00441-002-0548-5

Endter, S., Francombe, D., Gumbleton, M., & Ehrhardt, C. (2009). RT-PCR analysis of ABC, SLC and SLCO drug transporters in human lung epithelial cell models. *Journal of Pharmacy and Pharmacology*, *61*(5), 583–591. https://doi.org/10.1211/jpp/61.05.0006

Fang, X., Liu, Y., Xiao, W., Zhao, N., Zhu, C., Yu, D., & Zhao, Y. (2021). Prognostic SLC family genes promote cell proliferation, migration, and invasion in hepatocellular carcinoma. *Acta Biochimica et Biophysica Sinica*, *53*(8), 1065–1075. https://doi.org/10.1093/abbs/gmab076

Fellmann, C., Gowen, B. G., Lin, P.-C., Doudna, J. A., & Corn, J. E. (2017). Cornerstones of CRISPR–Cas in drug discovery and therapy. *Nature Reviews Drug Discovery*, *16*(2), 89–100. https://doi.org/10.1038/nrd.2016.238

Feng, B., Dresser, M. J., Shu, Y., Johns, S. J., & Giacomini, K. M. (2001). Arginine 454 and Lysine 370 Are Essential for the Anion Specificity of

the Organic Anion Transporter, rOAT3 †. *Biochemistry*, *40*(18), 5511–5520. https://doi.org/10.1021/bi002841o

Feng, S., Zacharioudaki, E., Millen, K., & Bray, S. J. (2020). The SLC36 transporter Pathetic is required for neural stem cell proliferation and for brain growth under nutrition restriction. *Neural Development*, *15*(1), 10. https://doi.org/10.1186/s13064-020-00148-4

Fernandes, C. A., & Vanbever, R. (2009). Preclinical models for pulmonary drug delivery. *Expert Opinion on Drug Delivery*, *6*(11), 1231–1245. https://doi.org/10.1517/17425240903241788

Fink, M. A., Paland, H., Herzog, S., Grube, M., Vogelgesang, S., Weitmann, K., Bialke, A., Hoffmann, W., Rauch, B. H., Schroeder, H. W. S., & Bien-Möller, S. (2019). L-Carnitine-Mediated Tumor Cell Protection and Poor Patient Survival Associated with OCTN2 Overexpression in Glioblastoma Multiforme. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *25*(9), 2874–2886. https://doi.org/10.1158/1078-0432.CCR-18-2380

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, *391*(6669), 806–811. https://doi.org/10.1038/35888

Fiszer-Kierzkowska, A., Vydra, N., Wysocka-Wycisk, A., Kronekova, Z., Jarząb, M., Lisowska, K. M., & Krawczyk, Z. (2011). Liposome-based DNA carriers may induce cellular stress response and change gene

expression pattern in transfected cells. *BMC Molecular Biology*, *12*(1), 27. https://doi.org/10.1186/1471-2199-12-27

Florea, B. I., Cassara, M. L., Junginger, H. E., & Borchard, G. (2003). Drug transport and metabolism characteristics of the human airway epithelial cell line Calu-3. *Journal of Controlled Release*, *87*(1–3), 131–138. https://doi.org/10.1016/S0168-3659(02)00356-5

Forbes, B., Shah, A., Martin, G. P., & Lansley, A. B. (2003). The human bronchial epithelial cell line 16HBE14o- as a model system of the airways for studying drug transport. *International Journal of Pharmaceutics*, *257*(1–2), 161–167. https://doi.org/10.1016/s0378-5173(03)00129-7

Foster, K. A., Oster, C. G., Mayer, M. M., Avery, M. L., & Audus, K. L. (1998). Characterization of the A549 Cell Line as a Type II Pulmonary Epithelial Cell Model for Drug Metabolism. *Experimental Cell Research*, *243*(2), 359–366. https://doi.org/10.1006/excr.1998.4172

Franks, T. J., Colby, T. V., Travis, W. D., Tuder, R. M., Reynolds, H. Y., Brody, A. R., Cardoso, W. V., Crystal, R. G., Drake, C. J., Engelhardt, J., Frid, M., Herzog, E., Mason, R., Phan, S. H., Randell, S. H., Rose, M. C., Stevens, T., Serge, J., Sunday, M. E., … Williams, M. C. (2008). Resident Cellular Components of the Human Lung. *Proceedings of the American Thoracic Society*, *5*(7), 763–766. https://doi.org/10.1513/pats.200803-025HR

Freitas-Lima, L. C., Budu, A., Arruda, A. C., Perilhão, M. S., Barrera-Chimal, J., Araujo, R. C., & Estrela, G. R. (2020). PPAR-α Deletion Attenuates Cisplatin Nephrotoxicity by Modulating Renal Organic Transporters MATE-1 and OCT-2. *International Journal of Molecular Sciences*, *21*(19), Article 19. https://doi.org/10.3390/ijms21197416

Friedman, J. R., Richbart, S. D., Merritt, J. C., Brown, K. C., Nolan, N. A., Akers, A. T., Lau, J. K., Robateau, Z. R., Miles, S. L., & Dasgupta, P. (2019). Acetylcholine signaling system in progression of lung cancers. *Pharmacology & Therapeutics*, *194*, 222–254. https://doi.org/10.1016/j.pharmthera.2018.10.002

Ganapathy, V., Thangaraju, M., & Prasad, P. D. (2009). Nutrient transporters in cancer: Relevance to Warburg hypothesis and beyond. *Pharmacology & Therapeutics*, *121*(1), 29–40. https://doi.org/10.1016/j.pharmthera.2008.09.005

Gao, X., Cai, Y., Wang, Z., He, W., Cao, S., Xu, R., & Chen, H. (2019). Estrogen receptors promote NSCLC progression by modulating the membrane receptor signaling network: A systems biology perspective. *Journal of Translational Medicine*, *17*(1), 308. https://doi.org/10.1186/s12967-019-2056-3

GINA Report: Global Strategy for Asthma Management and Prevention. (n.d.). Global Initiative for Asthma - GINA. Retrieved 15 January 2023, from https://ginasthma.org/reports/2019-gina-report-global-strategy-for-asthma-management-and-prevention/

Goel, M. K., Khanna, P., & Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research*, *1*(4), 274–278. https://doi.org/10.4103/0974-7788.76794

Goodarzi, A. A., Jeggo, P., & Lobrich, M. (2010). The influence of heterochromatin on DNA double strand break repair: Getting the strong, silent type to relax. *DNA Repair*, *9*(12), 1273–1282. https://doi.org/10.1016/j.dnarep.2010.09.013

Gorbunov, D., Gorboulev, V., Shatskaya, N., Mueller, T., Bamberg, E., Friedrich, T., & Koepsell, H. (2008). High-affinity cation binding to organic cation transporter 1 induces movement of helix 11 and blocks transport after mutations in a modeled interaction domain between two helices. *Molecular Pharmacology*, *73*(1), 50–61. https://doi.org/10.1124/mol.107.040170

Grandjean-Forestier, F., Stenger, C., Robert, J., Verdier, M., & Ratinaud, M.-H. (2009). The P-Glycoprotein 170: Just a Multidrug Resistance Protein or a Protean Molecule? In A. Boumendjel, J. Boutonnat, & J. Robert (Eds.), *ABC Transporters and Multidrug Resistance* (pp. 15–46). John Wiley & Sons, Inc. https://doi.org/10.1002/9780470495131.ch1

Gromicho, M., Dinis, J., Magalhães, M., Fernandes, A. R., Tavares, P., Laires, A., Rueff, J., & Rodrigues, A. S. (2011). Development of imatinib and dasatinib resistance: Dynamics of expression of drug transporters ABCB1, ABCC1, ABCG2, MVP, and SLC22A1. *Leukemia & Lymphoma*, *52*(10), 1980–1990. https://doi.org/10.3109/10428194.2011.584005

Groneberg, D. A., Witt, C., Wagner, U., Chung, K. F., & Fischer, A. (2003). Fundamentals of pulmonary drug delivery. *Respiratory Medicine*, *97*(4), 382–387. https://doi.org/10.1053/rmed.2002.1457

Gründemann, D. (2012). The ergothioneine transporter controls and indicates ergothioneine activity—A review. *Preventive Medicine*, *54*, S71–S74. https://doi.org/10.1016/j.ypmed.2011.12.001

Gründemann, D., Gorboulev, V., Gambaryan, S., Veyhl, M., & Koepsell, H. (1994). Drug excretion mediated by a new prototype of polyspecific transporter. *Nature*, *372*(6506), 549–552. https://doi.org/10.1038/372549a0

Gründemann, D., Harlfinger, S., Golz, S., Geerts, A., Lazar, A., Berkels, R., Jung, N., Rubbert, A., & Schömig, E. (2005). Discovery of the ergothioneine transporter. *Proceedings of the National Academy of Sciences*, *102*(14), 5256–5261. https://doi.org/10.1073/pnas.0408624102

Gu, Y., Xu, Z., Zhou, J., Wen, X., Jin, Y., Yuan, Q., Xia, P., Feng, Y., Yang, L., Lin, J., & Qian, J. (2022). SLC22A3 methylation-mediated gene silencing predicts adverse prognosis in acute myeloid leukemia. *Clinical Epigenetics*, *14*, 162. https://doi.org/10.1186/s13148-022-01373-w

Gumbleton, M., Al-Jayyoussi, G., Crandon-Lewis, A., Francombe, D., Kreitmeyr, K., Morris, C. J., & Smith, M. W. (2011). Spatial expression and functionality of drug transporters in the intact lung: Objectives for

further research. *Advanced Drug Delivery Reviews*, *63*(1), 110–118. https://doi.org/10.1016/j.addr.2010.09.008

Guo, W., Sun, S., Guo, L., Song, P., Xue, X., Zhang, H., Zhang, G., Li, R., Gao, Y., Qiu, B., Tan, F., Xue, Q., Gao, S., & He, J. (2020). Elevated SLC2A1 Expression Correlates with Poor Prognosis in Patients with Surgically Resected Lung Adenocarcinoma: A Study Based on Immunohistochemical Analysis and Bioinformatics. *DNA and Cell Biology*, *39*(4), 631–644. https://doi.org/10.1089/dna.2019.5291

Gupta, R. M., & Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *The Journal of Clinical Investigation*, *124*(10), 4154–4161. https://doi.org/10.1172/JCI72992

Gupta, S., Wulf, G., Henjakovic, M., Koepsell, H., Burckhardt, G., & Hagos, Y. (2012). Human organic cation transporter 1 is expressed in lymphoma cells and increases susceptibility to irinotecan and paclitaxel. *The Journal of Pharmacology and Experimental Therapeutics*, *341*(1), 16–23. https://doi.org/10.1124/jpet.111.190561

Guschin, D. Y., Waite, A. J., Katibah, G. E., Miller, J. C., Holmes, M. C., & Rebar, E. J. (2010). A Rapid and General Assay for Monitoring Endogenous Gene Modification. In J. P. Mackay & D. J. Segal (Eds.), *Engineered Zinc Finger Proteins: Methods and Protocols* (pp. 247–256). Humana Press. https://doi.org/10.1007/978-1-60761-753-2_15

Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly,

J.-S., & Concordet, J.-P. (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology*, *17*(1), 148. https://doi.org/10.1186/s13059-016-1012-2

Haghi, M., Ong, H. X., Traini, D., & Young, P. (2014). Across the pulmonary epithelial barrier: Integration of physicochemical properties and human cell models to study pulmonary drug formulations. *Pharmacology & Therapeutics*, *144*(3), 235–252. https://doi.org/10.1016/j.pharmthera.2014.05.003

Han, H. (2018). RNA Interference to Knock Down Gene Expression. In J. K. DiStefano (Ed.), *Disease Gene Identification: Methods and Protocols* (pp. 293–302). Springer. https://doi.org/10.1007/978-1-4939-7471-9_16

Hatano, T., Saiki, S., Okuzumi, A., Mohney, R. P., & Hattori, N. (2016). Identification of novel biomarkers for Parkinson's disease by metabolomic technologies. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(3), 295–301. https://doi.org/10.1136/jnnp-2014-309676

Heise, M., Lautem, A., Knapstein, J., Schattenberg, J. M., Hoppe-Lotichius, M., Foltys, D., Weiler, N., Zimmermann, A., Schad, A., Gründemann, D., Otto, G., Galle, P. R., Schuchmann, M., & Zimmermann, T. (2012). Downregulation of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) in human hepatocellular carcinoma and their prognostic significance. *BMC Cancer*, *12*, 109. https://doi.org/10.1186/1471-2407-12-109

Helander, H. F., & Fändriks, L. (2014). Surface area of the digestive tract—Revisited. *Scandinavian Journal of Gastroenterology*, *49*(6), 681–689. https://doi.org/10.3109/00365521.2014.898326

Hendrickx, R., Johansson, J. G., Lohmann, C., Jenvert, R.-M., Blomgren, A., Börjesson, L., & Gustavsson, L. (2013). Identification of novel substrates and structure-activity relationship of cellular uptake mediated by human organic cation transporters 1 and 2. *Journal of Medicinal Chemistry*, *56*(18), 7232–7242. https://doi.org/10.1021/jm400966v

Hermanns, M. I., Unger, R. E., Kehe, K., Peters, K., & Kirkpatrick, C. J. (2004). Lung epithelial cell lines in coculture with human pulmonary microvascular endothelial cells: Development of an alveolo-capillary barrier in vitro. *Laboratory Investigation*, *84*(6), 736. https://doi.org/10.1038/labinvest.3700081

Hickman, J. (2016). Transepithelial/endothelial Electrical Resistance (TEER) theory and applications for microfluidic body-on-a-chip devices. *Journal of Rare Diseases Research & Treatment*, *1*(3), 46–52. https://doi.org/10.29245/2572-9411/2016/3.1026

Hiemstra, P. S., Grootaers, G., van der Does, A. M., Krul, C. A. M., & Kooter, I. M. (2018). Human lung epithelial cell cultures for analysis of inhaled toxicants: Lessons learned and future directions. *Toxicology in Vitro*, *47*, 137–146. https://doi.org/10.1016/j.tiv.2017.11.005

Hlobilová, M. (2017). *Effect of L-carnitine uptake in skin cells* [Other]. https://doi.org/10.26226/morressier.595a9c55d462b80296c9fa93

Hong, M. (2017). Biochemical studies on the structure–function relationship of major drug transporters in the ATP-binding cassette family and solute carrier family. *Advanced Drug Delivery Reviews*, *116*, 3–20. https://doi.org/10.1016/j.addr.2016.06.003

Horvath, G., Schmid, N., Fragoso, M. A., Schmid, A., Conner, G. E., Salathe, M., & Wanner, A. (2007). Epithelial organic cation transporters ensure pH-dependent drug absorption in the airway. *American Journal of Respiratory Cell and Molecular Biology*, *36*(1), 53–60. https://doi.org/10.1165/rcmb.2006-0230OC

Hou, S., Wu, J., Li, X., & Shu, H. (2015). Practical, regulatory and clinical considerations for development of inhalation drug products. *Asian Journal of Pharmaceutical Sciences*, *10*(6), 490–500. https://doi.org/10.1016/j.ajps.2015.08.008

Hu, C., Tao, L., Cao, X., & Chen, L. (2019). The solute carrier transporters and the brain: Physiological and pharmacological implications. *Asian Journal of Pharmaceutical Sciences*. https://doi.org/10.1016/j.ajps.2019.09.002

Hu, D. G., Mackenzie, P. I., Nair, P. C., McKinnon, R. A., & Meech, R. (2020). The Expression Profiles of ADME Genes in Human Cancers and Their Associations with Clinical Outcomes. *Cancers*, *12*(11), Article 11. https://doi.org/10.3390/cancers12113369

Hu, K., Li, K., Lv, J., Feng, J., Chen, J., Wu, H., Cheng, F., Jiang, W., Wang, J., Pei, H., Chiao, P. J., Cai, Z., Chen, Y., Liu, M., & Pang, X. (2020).

Suppression of the SLC7A11/glutathione axis causes synthetic lethality in KRAS-mutant lung adenocarcinoma. *The Journal of Clinical Investigation*, *130*(4), 1752–1766. https://doi.org/10.1172/JCI124049

Hutter, V., Chau, D. Y. S., Hilgendorf, C., Brown, A., Cooper, A., Zann, V., Pritchard, D. I., & Bosquillon, C. (2014). Digoxin net secretory transport in bronchial epithelial cell layers is not exclusively mediated by P-glycoprotein/MDR1. *European Journal of Pharmaceutics and Biopharmaceutics*, *86*(1), 74–82. https://doi.org/10.1016/j.ejpb.2013.06.010

Hyrsova, L., Smutny, T., Trejtnar, F., & Pavek, P. (2016). Expression of organic cation transporter 1 (OCT1): Unique patterns of indirect regulation by nuclear receptors and hepatospecific gene regulation. *Drug Metabolism Reviews*, *48*(2), 139–158. https://doi.org/10.1080/03602532.2016.1188936

Ibrahim, M., & Garcia-Contreras, L. (2013). Mechanisms of absorption and elimination of drugs administered by inhalation. *Therapeutic Delivery*, *4*(8), 1027–1045. https://doi.org/10.4155/tde.13.67

Ingoglia, F., Visigalli, R., Rotoli, B. M., Barilli, A., Riccardi, B., Puccini, P., & Dall'Asta, V. (2015a). Functional characterization of the organic cation transporters (OCTs) in human airway pulmonary epithelial cells. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1848*(7), 1563–1572. https://doi.org/10.1016/j.bbamem.2015.04.001

Ingoglia, F., Visigalli, R., Rotoli, B. M., Barilli, A., Riccardi, B., Puccini, P., & Dall'Asta, V. (2015b). Functional characterization of the organic cation transporters (OCTs) in human airway pulmonary epithelial cells. *Biochimica Et Biophysica Acta*, *1848*(7), 1563–1572. https://doi.org/10.1016/j.bbamem.2015.04.001

International Transporter Consortium, Giacomini, K. M., Huang, S.-M., Tweedie, D. J., Benet, L. Z., Brouwer, K. L. R., Chu, X., Dahlin, A., Evers, R., Fischer, V., Hillgren, K. M., Hoffmaster, K. A., Ishikawa, T., Keppler, D., Kim, R. B., Lee, C. A., Niemi, M., Polli, J. W., Sugiyama, Y., … Zhang, L. (2010). Membrane transporters in drug development. *Nature Reviews. Drug Discovery*, *9*(3), 215–236. https://doi.org/10.1038/nrd3028

Jonker, J. W., Wagenaar, E., Mol, C. A., Buitelaar, M., Koepsell, H., Smit, J. W., & Schinkel, A. H. (2001). Reduced hepatic uptake and intestinal excretion of organic cations in mice with a targeted disruption of the organic cation transporter 1 (Oct1 [Slc22a1]) gene. *Molecular and Cellular Biology*, *21*(16), 5471–5477. https://doi.org/10.1128/MCB.21.16.5471-5477.2001

Jonker, J. W., Wagenaar, E., van Eijl, S., & Schinkel, A. H. (2003). Deficiency in the Organic Cation Transporters 1 and 2 (Oct1/Oct2 [Slc22a1/Slc22a2]) in Mice Abolishes Renal Secretion of Organic Cations. *Molecular and Cellular Biology*, *23*(21), 7902–7908. https://doi.org/10.1128/MCB.23.21.7902-7908.2003

Jung, E. S., Park, H. J., Kong, K. A., Choi, J. H., & Cheon, J. H. (2017). Association study between OCTN1 functional haplotypes and Crohn's disease in a Korean population. *The Korean Journal of Physiology & Pharmacology : Official Journal of the Korean Physiological Society and the Korean Society of Pharmacology*, *21*(1), 11–17. https://doi.org/10.4196/kjpp.2017.21.1.11

Juraszek, B., Czarnecka-Herok, J., & Nałęcz, K. A. (2021). Glioma cells survival depends both on fatty acid oxidation and on functional carnitine transport by SLC22A5. *Journal of Neurochemistry*, *156*(5), 642–657. https://doi.org/10.1111/jnc.15124

Kajiwara, M., Terada, T., Asaka, J., Aoki, M., Katsura, T., Ikai, I., & Inui, K. (2008). Regulation of basal core promoter activity of human organic cation transporter 1 (OCT1/SLC22A1). *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *295*(6), G1211–G1216. https://doi.org/10.1152/ajpgi.90360.2008

Kato, Y., Kubo, Y., Iwata, D., Kato, S., Sudo, T., Sugiura, T., Kagaya, T., Wakayama, T., Hirayama, A., Sugimoto, M., Sugihara, K., Kaneko, S., Soga, T., Asano, M., Tomita, M., Matsui, T., Wada, M., & Tsuji, A. (2010). Gene Knockout and Metabolome Analysis of Carnitine/Organic Cation Transporter OCTN1. *Pharmaceutical Research*, *27*(5), 832–840. https://doi.org/10.1007/s11095-010-0076-z

Kawoosa, F., Shah, Z. A., Masoodi, S. R., Amin, A., Rasool, R., Fazili, K. M., Dar, A. H., Lone, A., & ul Bashir, S. (2022). Role of human organic cation transporter-1 (OCT-1/SLC22A1) in modulating the response to

metformin in patients with type 2 diabetes. *BMC Endocrine Disorders*, *22*(1), 140. https://doi.org/10.1186/s12902-022-01033-3

Keller, T., Egenberger, B., Gorboulev, V., Bernhard, F., Uzelac, Z., Gorbunov, D., Wirth, C., Koppatz, S., Dötsch, V., Hunte, C., Sitte, H. H., & Koepsell, H. (2011). The Large Extracellular Loop of Organic Cation Transporter 1 Influences Substrate Affinity and Is Pivotal for Oligomerization. *The Journal of Biological Chemistry*, *286*(43), 37874–37886. https://doi.org/10.1074/jbc.M111.289330

Khanppnavar, B., Maier, J., Herborg, F., Gradisch, R., Lazzarin, E., Luethi, D., Yang, J.-W., Qi, C., Holy, M., Jäntsch, K., Kudlacek, O., Schicker, K., Werge, T., Gether, U., Stockner, T., Korkhov, V. M., & Sitte, H. H. (2022). Structural basis of organic cation transporter-3 inhibition. *Nature Communications*, *13*(1), Article 1. https://doi.org/10.1038/s41467-022-34284-8

Kim, C. H., Kim, H., Lee, Y., Lee, H., Lim, K.-S., Park, S. J., Huh, J.-W., Kim, Y.-H., Lee, D.-S., Kim, K. M., Hur, J. K., & Lee, S. H. (2022). *Highly precise genome editing using enhanced CRISPR-Cas12a nickase module* (p. 2022.08.27.505535). bioRxiv. https://doi.org/10.1101/2022.08.27.505535

Knight, S. C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L. B., Bosanac, L., Zhang, E. T., El Beheiry, M., Masson, J.-B., Dahan, M., Liu, Z., Doudna, J. A., & Tjian, R. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. *Science (New York, N.Y.)*, *350*(6262), 823–826. https://doi.org/10.1126/science.aac6572

Knudsen, L., & Ochs, M. (2018). The micromechanics of lung alveoli: Structure and function of surfactant and tissue components. *Histochemistry and Cell Biology*, *150*(6), 661–676. https://doi.org/10.1007/s00418-018-1747-9

Koepsell, H. (2013). The SLC22 family with transporters of organic cations, anions and zwitterions. *Molecular Aspects of Medicine*, *34*(2–3), 413–435. https://doi.org/10.1016/j.mam.2012.10.010

Koepsell, H. (2020). Organic Cation Transporters in Health and Disease. *Pharmacological Reviews*, *72*(1), 253–319. https://doi.org/10.1124/pr.118.015578

Koepsell, H., & Endou, H. (2004). The SLC22 drug transporter family. *Pflügers Archiv*, *447*(5), 666–676. https://doi.org/10.1007/s00424-003-1089-9

Koepsell, H., Lips, K., & Volk, C. (2007). Polyspecific Organic Cation Transporters: Structure, Function, Physiological Roles, and Biopharmaceutical Implications. *Pharmaceutical Research*, *24*(7), 1227–1251. https://doi.org/10.1007/s11095-007-9254-z

Koleske, M. L., McInnes, G., Brown, J. E. H., Thomas, N., Hutchinson, K., Chin, M. Y., Koehl, A., Arkin, M. R., Schlessinger, A., Gallagher, R. C., Song, Y. S., Altman, R. B., & Giacomini, K. M. (2022). Functional genomics of OCTN2 variants informs protein-specific variant effect predictor for Carnitine Transporter Deficiency. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(46), e2210247119. https://doi.org/10.1073/pnas.2210247119

Konstantakos, V., Nentidis, A., Krithara, A., & Paliouras, G. (2022). CRISPR–Cas9 gRNA efficiency prediction: An overview of predictive tools and the role of deep learning. *Nucleic Acids Research*, *50*(7), 3616–3637. https://doi.org/10.1093/nar/gkac192

Korn, T., Kühlkamp, T., Track, C., Schatz, I., Baumgarten, K., Gorboulev, V., & Koepsell, H. (2001). The Plasma Membrane-associated Protein RS1 Decreases Transcription of the Transporter SGLT1 in Confluent LLC-PK1 Cells*. *Journal of Biological Chemistry*, *276*(48), 45330–45340. https://doi.org/10.1074/jbc.M105975200

Kosicki, M., Tomberg, K., & Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nature Biotechnology*, *36*(8), 765–771. https://doi.org/10.1038/nbt.4192

Kreft, M. E., Jerman, U. D., Lasič, E., Hevir-Kene, N., Rižner, T. L., Peternel, L., & Kristan, K. (2015). The characterization of the human cell line Calu-3 under different culture conditions and its use as an optimized in vitro model to investigate bronchial epithelial function. *European Journal of Pharmaceutical Sciences*, *69*, 1–9. https://doi.org/10.1016/j.ejps.2014.12.017

Kroiss, M., Leyerer, M., Gorboulev, V., Kühlkamp, T., Kipp, H., & Koepsell, H. (2006). Transporter regulator RS1 (RSC1A1) coats the trans-Golgi network and migrates into the nucleus. *American Journal of Physiology. Renal Physiology*, *291*(6), F1201-1212. https://doi.org/10.1152/ajprenal.00067.2006

Kurosaki, T., & Maquat, L. E. (2016). Nonsense-mediated mRNA decay in humans at a glance. *Journal of Cell Science*, *129*(3), 461–467. https://doi.org/10.1242/jcs.181008

Labun, K., Krause, M., Torres Cleuren, Y., & Valen, E. (2021). CRISPR Genome Editing Made Easy Through the CHOPCHOP Website. *Current Protocols*, *1*(4), e46. https://doi.org/10.1002/cpz1.46

Lai, Y., Xue, J., Liu, C.-W., Gao, B., Chi, L., Tu, P., Lu, K., & Ru, H. (2019). Serum Metabolomics Identifies Altered Bioenergetics, Signaling Cascades in Parallel with Exposome Markers in Crohn's Disease. *Molecules*, *24*(3), Article 3. https://doi.org/10.3390/molecules24030449

Lautem, A., Heise, M., Gräsel, A., Hoppe-Lotichius, M., Weiler, N., Foltys, D., Knapstein, J., Schattenberg, J. M., Schad, A., Zimmermann, A., Otto, G., Lang, H., Galle, P. R., Schuchmann, M., & Zimmermann, T. (2013). Downregulation of organic cation transporter 1 (SLC22A1) is associated with tumor progression and reduced patient survival in human cholangiocellular carcinoma. *International Journal of Oncology*, *42*(4), 1297–1304. https://doi.org/10.3892/ijo.2013.1840

Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., & Ritchie, M. E. (2018). RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Research*, *5*, ISCB Comm J-1408. https://doi.org/10.12688/f1000research.9005.3

Lee, B. T., Barber, G. P., Benet-Pagès, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J. N., Hinrichs, A. S., Lee, C. M.,

Muthuraman, P., Nassar, L. R., Nguy, B., Pereira, T., Perez, G., Raney, B. J., Rosenbloom, K. R., Schmelter, D., Speir, M. L., … Kent, W. J. (2021). The UCSC Genome Browser database: 2022 update. *Nucleic Acids Research*, *50*(D1), D1115–D1122. https://doi.org/10.1093/nar/gkab959

Leung, E., Hong, J., Fraser, A. G., Merriman, T. R., Vishnu, P., & Krissansen, G. W. (2006). Polymorphisms in the organic cation transporter genes SLC22A4 and SLC22A5 and Crohn's disease in a New Zealand Caucasian cohort. *Immunology and Cell Biology*, *84*(2), 233–236. https://doi.org/10.1111/j.1440-1711.2006.01423.x

Li, B., Clohisey, S. M., Chia, B. S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L. D., Hoover, P., Parkinson, N. J., Nachshon, A., Smith, N., Regan, T., Farr, D., Gutmann, M. U., Bukhari, S. I., Law, A., Sangesland, M., Gat-Viks, I., Digard, P., … Hacohen, N. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nature Communications*, *11*, 164. https://doi.org/10.1038/s41467-019-13965-x

Li, D., Qi, C., Zhou, J., Wen, Z., Zhu, X., Xia, H., & Song, J. (2020). LPS-induced inflammation delays the transportation of ASP+ due to down-regulation of OCTN1/2 in alveolar epithelial cells. *Journal of Drug Targeting*, *28*(4), 437–447. https://doi.org/10.1080/1061186X.2019.1678169

LI, Q., YANG, H., PENG, X., GUO, D., DONG, Z., POLLI, J. E., & SHU, Y. (2013). Ischemia/Reperfusion-Inducible Protein Modulates the Function

of Organic Cation Transporter 1 and Multidrug and Toxin Extrusion1. *Molecular Pharmaceutics*, *10*(7), 2578–2587. https://doi.org/10.1021/mp400013t

Lin, L., Yee, S. W., Kim, R. B., & Giacomini, K. M. (2015). SLC Transporters as Therapeutic Targets: Emerging Opportunities. *Nature Reviews. Drug Discovery*, *14*(8), 543–560. https://doi.org/10.1038/nrd4626

Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., & Callewaert, N. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications*, *5*, 4767. https://doi.org/10.1038/ncomms5767

Lindeboom, R. G. H., Supek, F., & Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nature Genetics*, *48*(10), 1112–1118. https://doi.org/10.1038/ng.3664

Lips, K. S., Volk, C., Schmitt, B. M., Pfeil, U., Arndt, P., Miska, D., Ermert, L., Kummer, W., & Koepsell, H. (2005). Polyspecific cation transporters mediate luminal release of acetylcholine from bronchial epithelium. *American Journal of Respiratory Cell and Molecular Biology*, *33*(1), 79–88. https://doi.org/10.1165/rcmb.2004-0363OC

Liu, X. (2019). SLC Family Transporters. *Advances in Experimental Medicine and Biology*, *1141*, 101–202. https://doi.org/10.1007/978-981-13-7647-4_3

Liu, X., Homma, A., Sayadi, J., Yang, S., Ohashi, J., & Takumi, T. (2016). Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Scientific Reports*, *6*(1), Article 1. https://doi.org/10.1038/srep19675

Longo, N., Frigeni, M., & Pasquali, M. (2016). Carnitine transport and fatty acid oxidation. *Biochimica Et Biophysica Acta*, *1863*(10), 2422–2435. https://doi.org/10.1016/j.bbamcr.2016.01.023

Lozano, E., Briz, O., Macias, R., Serrano, M., Marin, J., & Herraez, E. (2018). Genetic Heterogeneity of SLC22 Family of Transporters in Drug Disposition. *Journal of Personalized Medicine*, *8*(2), 14. https://doi.org/10.3390/jpm8020014

Lozano, E., Herraez, E., Briz, O., Robledo, V. S., Hernandez-Iglesias, J., Gonzalez-Hernandez, A., & Marin, J. J. G. (2013). Role of the Plasma Membrane Transporter of Organic Cations OCT1 and Its Genetic Variants in Modern Liver Pharmacology. *BioMed Research International*, *2013*, 1–13. https://doi.org/10.1155/2013/692071

Lujan, H., Criscitiello, M. F., Hering, A. S., & Sayes, C. M. (2019). Refining In Vitro Toxicity Models: Comparing Baseline Characteristics of Lung Cell Types. *Toxicological Sciences*, *168*(2), 302–314. https://doi.org/10.1093/toxsci/kfz001

Lund, A. H., Duch, M., & Skou Pedersen, F. (1996). Increased Cloning Efficiency by Temperature-Cycle Ligation. *Nucleic Acids Research*, *24*(4), 800–801. https://doi.org/10.1093/nar/24.4.800

Lung Cancer Guide. (n.d.). American Cancer Society. Retrieved 15 January 2023, from https://www.cancer.org/cancer/lung-cancer.html

Macdonald, C., Shao, D., Oli, A., & Agu, R. U. (2013). Characterization of Calu-3 cell monolayers as a model of bronchial epithelial transport: Organic cation interaction studies. *Journal of Drug Targeting*, *21*(1), 97–106. https://doi.org/10.3109/1061186X.2012.731068

Maeda, T., Hirayama, M., Kobayashi, D., Miyazawa, K., & Tamai, I. (2007). Mechanism of the Regulation of Organic Cation/Carnitine Transporter 1 (SLC22A4) by Rheumatoid Arthritis-Associated Transcriptional Factor RUNX1 and Inflammatory Cytokines. *Drug Metabolism and Disposition*, *35*(3), 394–401. https://doi.org/10.1124/dmd.106.012112

Martovetsky, G., Tee, J. B., & Nigam, S. K. (2013). Hepatocyte Nuclear Factors 4α and 1α Regulate Kidney Developmental Expression of Drug-Metabolizing Enzymes and Drug Transporters. *Molecular Pharmacology*, *84*(6), 808–823. https://doi.org/10.1124/mol.113.088229

Mathias, N. R., Timoszyk, J., Stetsko, P. I., Megill, J. R., Smith, R. L., & Wall, D. A. (2002). Permeability Characteristics of Calu-3 Human Bronchial Epithelial Cells: In Vitro - In Vivo Correlation to Predict Lung Absorption in Rats. *Journal of Drug Targeting*, *10*(1), 31–40. https://doi.org/10.1080/10611860290007504

McDowell, E. M., Barrett, L. A., Glavin, F., Harris, C. C., & Trump, B. F. (1978). The Respiratory Epithelium. I. Human Bronchus. *JNCI: Journal*

of the National Cancer Institute, *61*(2), 539–549. https://doi.org/10.1093/jnci/61.2.539

Mehendale, H. M., Angevine, L. S., & Ohmiya, Y. (1981). The isolated perfused lung—A critical evaluation. *Toxicology*, *21*(1), 1–36. https://doi.org/10.1016/0300-483X(81)90013-5

Metzger, L., & Iliakis, G. (1991). Kinetics of DNA double-strand break repair throughout the cell cycle as assayed by pulsed field gel electrophoresis in CHO cells. *International Journal of Radiation Biology*, *59*(6), 1325–1339. https://doi.org/10.1080/09553009114551201

Meyer, M. J., Seitz, T., Brockmöller, J., & Tzvetkov, M. V. (2017). Effects of genetic polymorphisms on the OCT1 and OCT2-mediated uptake of ranitidine. *PLoS ONE*, *12*(12). https://doi.org/10.1371/journal.pone.0189521

Min, K. A., Rosania, G. R., & Shin, M. C. (2016). Human Airway Primary Epithelial Cells Show Distinct Architectures on Membrane Supports under Different Culture Conditions. *Cell Biochemistry and Biophysics*, *74*(2), 191–203. https://doi.org/10.1007/s12013-016-0719-8

Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G. P., & Eliopoulos, E. (2020). Transcription factors and evolution: An integral part of gene expression (Review). *World Academy of Sciences Journal*, *2*(1), 3–8. https://doi.org/10.3892/wasj.2020.32

Mo, J., Shi, S., Zhang, Q., Gong, T., Sun, X., & Zhang, Z. (2011). Synthesis, Transport and Mechanism of a Type I Prodrug: L-Carnitine Ester of

Prednisolone. *Molecular Pharmaceutics*, *8*(5), 1629–1640. https://doi.org/10.1021/mp100412z

Moffatt, M. F., Gut, I. G., Demenais, F., Strachan, D. P., Bouzigon, E., Heath, S., von Mutius, E., Farrall, M., Lathrop, M., & Cookson, W. O. C. M. (2010). A Large-Scale, Consortium-Based Genomewide Association Study of Asthma. *New England Journal of Medicine*, *363*(13), 1211–1221. https://doi.org/10.1056/NEJMoa0906312

Moliva, J. I., Rajaram, M. V. S., Sidiki, S., Sasindran, S. J., Guirado, E., Pan, X. J., Wang, S.-H., Ross, P., Lafuse, W. P., Schlesinger, L. S., Turner, J., & Torrelles, J. B. (2014). Molecular composition of the alveolar lining fluid in the aging lung. *Age*, *36*(3). https://doi.org/10.1007/s11357-014-9633-4

Morey, J. S., Ryan, J. C., & Van Dolah, F. M. (2006). Microarray validation: Factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological Procedures Online*, *8*(1), Article 1. https://doi.org/10.1251/bpo126

Morrow, J. D., Zhou, X., Lao, T., Jiang, Z., DeMeo, D. L., Cho, M. H., Qiu, W., Cloonan, S., Pinto-Plata, V., Celli, B., Marchetti, N., Criner, G. J., Bueno, R., Washko, G. R., Glass, K., Quackenbush, J., Choi, A. M. K., Silverman, E. K., & Hersh, C. P. (2017). Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. *Scientific Reports*, *7*, 44232. https://doi.org/10.1038/srep44232

Motohashi, H., & Inui, K. (2013). Organic Cation Transporter OCTs (SLC22) and MATEs (SLC47) in the Human Kidney. *The AAPS Journal*, *15*(2), 581–588. https://doi.org/10.1208/s12248-013-9465-7

Mu, W., Tang, N., Cheng, C., Sun, W., Wei, X., & Wang, H. (2019). In vitro transcribed sgRNA causes cell death by inducing interferon release. *Protein & Cell*, *10*(6), 461–465. https://doi.org/10.1007/s13238-018-0605-9

Mukherjee, M. (2012). *An in vitro analysis of the expression and function of bronchial Organic cation Transporters*. University of Nottingham.

Mukherjee, M., Cingolani, E., Pritchard, D. I., & Bosquillon, C. (2017). Enhanced expression of Organic Cation Transporters in bronchial epithelial cell layers following insults associated with asthma – Impact on salbutamol transport. *European Journal of Pharmaceutical Sciences*, *106*, 62–70. https://doi.org/10.1016/j.ejps.2017.05.052

Mukherjee, M., Latif, M. L., Pritchard, D. I., & Bosquillon, C. (2013). In-cell Western™ detection of organic cation transporters in bronchial epithelial cell layers cultured at an air-liquid interface on Transwell(®) inserts. *Journal of Pharmacological and Toxicological Methods*, *68*(2), 184–189. https://doi.org/10.1016/j.vascn.2013.05.007

Mukherjee, M., Pritchard, D. I., & Bosquillon, C. (2012). Evaluation of air-interfaced Calu-3 cell layers for investigation of inhaled drug interactions with organic cation transporters in vitro. *International*

Journal of Pharmaceutics, *426*(1), 7–14. https://doi.org/10.1016/j.ijpharm.2011.12.036

Müller, J., Lips, K. S., Metzner, L., Neubert, R. H. H., Koepsell, H., & Brandsch, M. (2005). Drug specificity and intestinal membrane localization of human organic cation transporters (OCT). *Biochemical Pharmacology*, *70*(12), 1851–1860. https://doi.org/10.1016/j.bcp.2005.09.011

Nakamura, T., Nakanishi, T., Haruta, T., Shirasaka, Y., Keogh, J. P., & Tamai, I. (2010). Transport of Ipratropium, an Anti-Chronic Obstructive Pulmonary Disease Drug, Is Mediated by Organic Cation/Carnitine Transporters in Human Bronchial Epithelial Cells: Implications for Carrier-Mediated Pulmonary Absorption. *Molecular Pharmaceutics*, *7*(1), 187–195. https://doi.org/10.1021/mp900206j

Nakamura, T., Sugiura, S., Kobayashi, D., Yoshida, K., Yabuuchi, H., Aizawa, S., Maeda, T., & Tamai, I. (2007). Decreased Proliferation and Erythroid Differentiation of K562 Cells by siRNA-induced Depression of OCTN1 (SLC22A4) Transporter Gene. *Pharmaceutical Research*, *24*(9), 1628–1635. https://doi.org/10.1007/s11095-007-9290-8

Nakanishi, T., Hasegawa, Y., Haruta, T., Wakayama, T., & Tamai, I. (2013). In Vivo Evidence of Organic Cation Transporter-Mediated Tracheal Accumulation of the Anticholinergic Agent Ipratropium in Mice. *Journal of Pharmaceutical Sciences*, *102*(9), 3373–3381. https://doi.org/10.1002/jps.23603

Narai, A., Arai, S., & Shimizu, M. (1997). Rapid decrease in transepithelial electrical resistance of human intestinal Caco-2 cell monolayers by cytotoxic membrane perturbents. *Toxicology in Vitro*, *11*(4), 347–354. https://doi.org/10.1016/S0887-2333(97)00026-X

Neuhoff, S., Ungell, A.-L., Zamora, I., & Artursson, P. (2003). pH-Dependent Bidirectional Transport of Weakly Basic Drugs Across Caco-2 Monolayers: Implications for Drug–Drug Interactions. *Pharmaceutical Research*, *20*(8), 1141–1148. https://doi.org/10.1023/A:1025032511040

Nicholls, G., Youdim, K., & Royal Society of Chemistry (Great Britain) (Eds.). (2016). *Drug transporters*. Royal Society Of Chemistry.

Niemeier, R. W. (1984). The isolated perfused lung. *Environmental Health Perspectives*, *56*, 35–41.

Nies, A. T., Koepsell, H., Winter, S., Burk, O., Klein, K., Kerb, R., Zanger, U. M., Keppler, D., Schwab, M., & Schaeffeler, E. (2009). Expression of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) is affected by genetic factors and cholestasis in human liver. *Hepatology*, *50*(4), 1227–1240. https://doi.org/10.1002/hep.23103

Nigam, S. K. (2018). The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annual Review of Pharmacology and Toxicology*, *58*, 663–687. https://doi.org/10.1146/annurev-pharmtox-010617-052713

Nishimura, M., & Naito, S. (2005). Tissue-specific mRNA Expression Profiles of Human ATP-binding Cassette and Solute Carrier Transporter

Superfamilies. *Drug Metabolism and Pharmacokinetics*, *20*(6), 452–477. https://doi.org/10.2133/dmpk.20.452

Novakova, K., Török, M., Panajatovic, M., Bouitbir, J., Duong, F. H. T., Handschin, C., & Krähenbühl, S. (2022). PGC-1α and MEF2 Regulate the Transcription of the Carnitine Transporter OCTN2 Gene in C2C12 Cells and in Mouse Skeletal Muscle. *International Journal of Molecular Sciences*, *23*(20), 12304. https://doi.org/10.3390/ijms232012304

Ober, C., & Nicolae, D. L. (2011). Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nature Genetics*, *43*(9), Article 9. https://doi.org/10.1038/ng.888

O'Brien, V. P., Bokelmann, K., Ramírez, J., Jobst, K., Ratain, M. J., Brockmöller, J., & Tzvetkov, M. V. (2013). Hepatocyte nuclear factor 1 regulates the expression of the organic cation transporter 1 via binding to an evolutionary conserved region in intron 1 of the OCT1 gene. *The Journal of Pharmacology and Experimental Therapeutics*, *347*(1), 181–192. https://doi.org/10.1124/jpet.113.206359

Oh, T.-I., Lee, M., Lee, Y.-M., Kim, G.-H., Lee, D., You, J. S., Kim, S. H., Choi, M., Jang, H., Park, Y.-M., Shin, H.-W., Shin, D. H., & Lim, J.-H. (2021). PGC1α Loss Promotes Lung Cancer Metastasis through Epithelial-Mesenchymal Transition. *Cancers*, *13*(8), 1772. https://doi.org/10.3390/cancers13081772

Öhlinger, K., Kolesnik, T., Meindl, C., Gallé, B., Absenger-Novak, M., Kolb-Lenz, D., & Fröhlich, E. (2019). Air-liquid interface culture changes

surface properties of A549 cells. *Toxicology in Vitro: An International Journal Published in Association with BIBRA*, *60*, 369–382. https://doi.org/10.1016/j.tiv.2019.06.014

O'Keefe, E. P. (2020). siRNAs and shRNAs: Tools for Protein Knockdown by Gene Silencing. *Materials and Methods*. /method/siRNAs-and-shRNAs-Tools-for-Protein-Knockdown-by-Gene-Silencing.html

Olsson, B., Bondesson, E., Borgström, L., Edsbäcker, S., Eirefelt, S., Ekelund, K., Gustavsson, L., & Hegelund-Myrbäck, T. (2011). Pulmonary Drug Metabolism, Clearance, and Absorption. In H. D. C. Smyth & A. J. Hickey (Eds.), *Controlled Pulmonary Drug Delivery* (pp. 21–50). Springer New York. https://doi.org/10.1007/978-1-4419-9745-6_2

Oostendorp, R. L., Buckle, T., Beijnen, J. H., van Tellingen, O., & Schellens, J. H. M. (2009). The effect of P-gp (Mdr1a/1b), BCRP (Bcrp1) and P-gp/BCRP inhibitors on the in vivo absorption, distribution, metabolism and excretion of imatinib. *Investigational New Drugs*, *27*(1), 31–40. https://doi.org/10.1007/s10637-008-9138-z

Oscarson, M., Zanger, U. M., Rifki, O. F., Klein, K., Eichelbaum, M., & Meyer, U. A. (2006). Transcriptional profiling of genes induced in the livers of patients treated with carbamazepine. *Clinical Pharmacology & Therapeutics*, *80*(5), 440–456. https://doi.org/10.1016/j.clpt.2006.08.013

Otter, M., Csader, S., Keiser, M., & Oswald, S. (2022). Expression and Functional Contribution of Different Organic Cation Transporters to the

Cellular Uptake of Doxorubicin into Human Breast Cancer and Cardiac Tissue. *International Journal of Molecular Sciences*, *23*(1), Article 1. https://doi.org/10.3390/ijms23010255

Panduga, V., Stocks, M. J., & Bosquillon, C. (2017). Ipratropium is 'luminally recycled' by an inter-play between apical uptake and efflux transporters in Calu-3 bronchial epithelial cell layers. *International Journal of Pharmaceutics*, *532*(1), 328–336. https://doi.org/10.1016/j.ijpharm.2017.08.112

Park, R., Park, Y.-I., Park, Y., Lee, S., So, J., & Park, J. (2022). CTRP1 Knockout Attenuates Tumor Progression in A549 and HCT116 Cancer Cells. *Cancers*, *14*(18), Article 18. https://doi.org/10.3390/cancers14184495

Park, S.-Y., Choi, H.-C., Chun, Y.-H., Kim, H., & Park, S.-H. (2001). Characterization of chromosomal aberrations in lung cancer cell lines by cross-species color banding. *Cancer Genetics and Cytogenetics*, *124*(1), 62–70. https://doi.org/10.1016/S0165-4608(00)00329-0

Patil, J. S., & Sarasija, S. (2012). Pulmonary drug delivery strategies: A concise, systematic review. *Lung India : Official Organ of Indian Chest Society*, *29*(1), 44–49. https://doi.org/10.4103/0970-2113.92361

Patton, J. S. (1996). Mechanisms of macromolecule absorption by the lungs. *Advanced Drug Delivery Reviews*, *19*(1), 3–36. https://doi.org/10.1016/0169-409X(95)00113-L

Pawlik, A., Paradowska-Gorycka, A., Safranow, K., Dziedziejko, V., Dutkiewicz, G., Słucznowska-Głabowska, S., Juzyszyn, Z., & Drozdzik, M. (2019). SLC22A5 polymorphism associated with risk of extra-articular manifestations in rheumatoid arthritis patients. *Reumatologia*, *57*(1), 3–7. https://doi.org/10.5114/reum.2019.83233

Pedersen, B. P., Kumar, H., Waight, A. B., Risenmay, A. J., Roe-Zurz, Z., Chau, B. H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A., Johri, A. K., & Stroud, R. M. (2013). Crystal structure of a eukaryotic phosphate transporter. *Nature*, *496*(7446), Article 7446. https://doi.org/10.1038/nature12042

Peltekova, V. D., Wintle, R. F., Rubin, L. A., Amos, C. I., Huang, Q., Gu, X., Newman, B., Oene, M. V., Cescon, D., Greenberg, G., Griffiths, A. M., St George-Hyslop, P. H., & Siminovitch, K. A. (2004). Functional variants of OCTN cation transporter genes are associated with Crohn disease. *Nature Genetics*, *36*(5), Article 5. https://doi.org/10.1038/ng1339

Peng, K.-J., Wang, J.-H., Su, W.-T., Wang, X.-C., Yang, F.-T., & Nie, W.-H. (2010). Characterization of two human lung adenocarcinoma cell lines by reciprocal chromosome painting. *Dong Wu Xue Yan Jiu = Zoological Research*, *31*(2), 113–121. https://doi.org/10.3724/SP.J.1141.2010.02113

Pfeiffer, C., Bach, M., Bauer, T., Campos da Ponte, J., Schömig, E., & Gründemann, D. (2015). Knockout of the ergothioneine transporter ETT in zebrafish results in increased 8-oxoguanine levels. *Free Radical*

Biology and Medicine, *83*, 178–185. https://doi.org/10.1016/j.freeradbiomed.2015.02.026

Pilcer, G., & Amighi, K. (2010). Formulation strategy and use of excipients in pulmonary drug delivery. *International Journal of Pharmaceutics*, *392*(1), 1–19. https://doi.org/10.1016/j.ijpharm.2010.03.017

Pinkerton, K. E., Gehr, P., Castañeda, A., & Crapo, J. D. (2015). Chapter 9—Architecture and Cellular Composition of the Air–Blood Tissue Barrier. In R. A. Parent (Ed.), *Comparative Biology of the Normal Lung (Second Edition)* (pp. 105–117). Academic Press. https://doi.org/10.1016/B978-0-12-404577-4.00009-6

Plopper, C. G. (1996). Structure and Function of the Lung. In T. C. Jones, D. L. Dungworth, & U. Mohr (Eds.), *Respiratory System* (pp. 135–150). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-61042-4_13

Plopper, C. G., & Hyde, D. M. (2015). Chapter 7—Epithelial Cells of the Bronchiole. In R. A. Parent (Ed.), *Comparative Biology of the Normal Lung (Second Edition)* (pp. 83–92). Academic Press. https://doi.org/10.1016/B978-0-12-404577-4.00007-2

Pochini, L., Galluccio, M., Scalise, M., Console, L., Pappacoda, G., & Indiveri, C. (2022). OCTN1: A Widely Studied but Still Enigmatic Organic Cation Transporter Linked to Human Pathology and Drug Interactions. *International Journal of Molecular Sciences*, *23*(2), 914. https://doi.org/10.3390/ijms23020914

Pochini, L., Scalise, M., Galluccio, M., Amelio, L., & Indiveri, C. (2011). Reconstitution in liposomes of the functionally active human OCTN1 (SLC22A4) transporter overexpressed in Escherichia coli. *Biochemical Journal*, *439*(2), 227–233. https://doi.org/10.1042/BJ20110544

Polydisperse Microparticle Transport and Deposition to the Terminal Bronchioles in a Heterogeneous Vasculature Tree | Scientific Reports. (n.d.). Retrieved 5 June 2020, from https://www.nature.com/articles/s41598-018-34804-x

Popowski, K., Eloranta, J. J., Saborowski, M., Fried, M., Meier, P. J., & Kullak-Ublick, G. A. (2005). The human organic anion transporter 2 gene is transactivated by hepatocyte nuclear factor-4 alpha and suppressed by bile acids. *Molecular Pharmacology*, *67*(5), 1629–1638. https://doi.org/10.1124/mol.104.010223

Portelli, M. A., Rakkar, K., Hu, S., Guo, Y., Adcock, I. M., & Sayers, I. (2021). Translational Analysis of Moderate to Severe Asthma GWAS Signals Into Candidate Causal Genes and Their Functional, Tissue-Dependent and Disease-Related Associations. *Frontiers in Allergy*, *2*, 738741. https://doi.org/10.3389/falgy.2021.738741

Prendergast, G. C., & Ziff, E. B. (1991). Methylation-Sensitive Sequence-Specific DNA Binding by the c-Myc Basic Region. *Science*, *251*(4990), 186–189. https://doi.org/10.1126/science.1987636

Qi, C., Zhou, J., Wang, Z., Fang, X., Li, D., Jin, Y., & Song, J. (2020). Cigarette smoke extract combined with lipopolysaccharide reduces OCTN1/2

expression in human alveolar epithelial cells in vitro and rat lung in vivo under inflammatory conditions. *International Immunopharmacology*, *87*, 106812. https://doi.org/10.1016/j.intimp.2020.106812

Qian, X., Wang, X., Luo, J., Liu, Y., Pang, J., Zhang, H., Xu, Z., Xie, J., Jiang, X., & Ling, W. (2019). Hypouricemic and nephroprotective roles of anthocyanins in hyperuricemic mice. *Food & Function*, *10*(2), 867–878. https://doi.org/10.1039/C8FO02124D

Qosa, H., Mohamed, L. A., Alqahtani, S., Abuasal, B. S., Hill, R. A., & Kaddoumi, A. (2016). Transporters as Drug Targets in Neurological Diseases. *Clinical Pharmacology and Therapeutics*, *100*(5), 441–453. https://doi.org/10.1002/cpt.435

Qu, Q., Qu, J., Zhan, M., Wu, L.-X., Zhang, Y.-W., Lou, X.-Y., Fu, L.-J., & Zhou, H.-H. (2013). Different Involvement of Promoter Methylation in the Expression of Organic Cation/Carnitine Transporter 2 (OCTN2) in Cancer Cell Lines. *PLOS ONE*, *8*(10), e76474. https://doi.org/10.1371/journal.pone.0076474

Rackley, C. R., & Stripp, B. R. (2012). Building and maintaining the epithelium of the lung. *The Journal of Clinical Investigation*, *122*(8), 2724–2730. https://doi.org/10.1172/JCI60519

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), Article 11. https://doi.org/10.1038/nprot.2013.143

Rashid, K., Ahmad, A., Liang, L., Liu, M., Cui, Y., & Liu, T. (2021). Solute carriers as potential oncodrivers or suppressors: Their key functions in malignant tumor formation. *Drug Discovery Today*, *26*(7), 1689–1701. https://doi.org/10.1016/j.drudis.2021.03.004

Rayner, R. E., Makena, P., Prasad, G. L., & Cormet-Boyaka, E. (2019). Optimization of Normal Human Bronchial Epithelial (NHBE) Cell 3D Cultures for in vitro Lung Model Studies. *Scientific Reports*, *9*. https://doi.org/10.1038/s41598-018-36735-z

Rebendenne, A., Roy, P., Bonaventure, B., Chaves, V. A. L., Desmarets, L., Rouillé, Y., Tauziet, M., Arnaud-Arnould, M., Giovannini, D., Lee, Y., DeWeirdt, P., Hegde, M., Garcia de, G. F., McKellar, J., Wencker, M., Dubuisson, J., Belouzard, S., Moncorgé, O., Doench, J. G., & Goujon, C. (2021). Bidirectional genome-wide CRISPR screens reveal host factors regulating SARS-CoV-2, MERS-CoV and seasonal HCoVs. *Research Square*, rs.3.rs-555275. https://doi.org/10.21203/rs.3.rs-555275/v1

Reynolds, S. D., Pinkerton, K. E., & Mariassy, A. T. (2015). Chapter 6—Epithelial Cells of Trachea and Bronchi. In R. A. Parent (Ed.), *Comparative Biology of the Normal Lung (Second Edition)* (pp. 61–81). Academic Press. https://doi.org/10.1016/B978-0-12-404577-4.00006-0

Roblek, M., Bicher, J., van Gogh, M., György, A., Seeböck, R., Szulc, B., Damme, M., Olczak, M., Borsig, L., & Siekhaus, D. E. (2022). The Solute Carrier MFSD1 Decreases the Activation Status of β1 Integrin

and Thus Tumor Metastasis. *Frontiers in Oncology*, *12*. https://www.frontiersin.org/articles/10.3389/fonc.2022.777634

Rokicki, W., Rokicki, M., Wojtacha, J., & Dżeljijli, A. (2016). The role and importance of club cells (Clara cells) in the pathogenesis of some respiratory diseases. *Kardiochirurgia i Torakochirurgia Polska = Polish Journal of Cardio-Thoracic Surgery*, *13*(1), 26–30. https://doi.org/10.5114/kitp.2016.58961

Rotoli, B. M., Visigalli, R., Barilli, A., Ferrari, F., Bianchi, M. G., Lascia, M. D., Riccardi, B., Puccini, P., & Dall'Asta, V. (2020). Functional analysis of OCTN2 and ATB0,+ in normal human airway epithelial cells. *PLOS ONE*, *15*(2), e0228568. https://doi.org/10.1371/journal.pone.0228568

Rulcova, A., Krausova, L., Smutny, T., Vrzal, R., Dvorak, Z., Jover, R., & Pavek, P. (2013). Glucocorticoid receptor regulates organic cation transporter 1 (OCT1, SLC22A1) expression via HNF4α upregulation in primary human hepatocytes. *Pharmacological Reports: PR*, *65*(5), 1322–1335. https://doi.org/10.1016/s1734-1140(13)71491-9

Saborowski, M., Kullak-Ublick, G. A., & Eloranta, J. J. (2006). The Human Organic Cation Transporter-1 Gene Is Transactivated by Hepatocyte Nuclear Factor-4α. *Journal of Pharmacology and Experimental Therapeutics*, *317*(2), 778–785. https://doi.org/10.1124/jpet.105.099929

Sakagami, M. (2006). In vivo, in vitro and ex vivo models to assess pulmonary absorption and disposition of inhaled therapeutics for systemic delivery.

Advanced Drug Delivery Reviews, *58*(9), 1030–1060. https://doi.org/10.1016/j.addr.2006.07.012

Sakamoto, A., Matsumaru, T., Yamamura, N., Suzuki, S., Uchida, Y., Tachikawa, M., & Terasaki, T. (2015). Drug Transporter Protein Quantification of Immortalized Human Lung Cell Lines Derived from Tracheobronchial Epithelial Cells (Calu-3 and BEAS2-B), Bronchiolar–Alveolar Cells (NCI-H292 and NCI-H441), and Alveolar Type II-like Cells (A549) by Liquid Chromatography–Tandem Mass Spectrometry. *Journal of Pharmaceutical Sciences*, *104*(9), 3029–3038. https://doi.org/10.1002/jps.24381

Sakamoto, A., Matsumaru, T., Yamamura, N., Uchida, Y., Tachikawa, M., Ohtsuki, S., & Terasaki, T. (2013). Quantitative expression of human drug transporter proteins in lung tissues: Analysis of regional, gender, and interindividual differences by liquid chromatography–tandem mass spectrometry. *Journal of Pharmaceutical Sciences*, *102*(9), 3395–3406. https://doi.org/10.1002/jps.23606

Salani, B., Maffioli, S., Hamoudane, M., Parodi, A., Ravera, S., Passalacqua, M., Alama, A., Nhiri, M., Cordera, R., & Maggi, D. (2012). Caveolin-1 is essential for metformin inhibitory effect on IGF1 action in non-small-cell lung cancer cells. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *26*(2), 788–798. https://doi.org/10.1096/fj.11-192088

Salomon, J. J., Endter, S., Tachon, G., Falson, F., Buckley, S. T., & Ehrhardt, C. (2012). Transport of the fluorescent organic cation 4-(4-

(dimethylamino)styryl)-N-methylpyridinium iodide (ASP+) in human respiratory epithelial cells. *European Journal of Pharmaceutics and Biopharmaceutics: Official Journal of Arbeitsgemeinschaft Fur Pharmazeutische Verfahrenstechnik e.V*, *81*(2), 351–359. https://doi.org/10.1016/j.ejpb.2012.03.001

Salomon, J. J., Hagos, Y., Petzke, S., Kühne, A., Gausterer, J. C., Hosoya, K., & Ehrhardt, C. (2015). Beta-2 Adrenergic Agonists Are Substrates and Inhibitors of Human Organic Cation Transporter 1. *Molecular Pharmaceutics*, *12*(8), 2633–2641. https://doi.org/10.1021/mp500854e

Sander, J. D., & Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology*, *32*(4), 347–355. https://doi.org/10.1038/nbt.2842

Santiago, J. L., Martínez, A., de la Calle, H., Fernández-Arquero, M., Figueredo, M. Á., de la Concha, E. G., & Urcelay, E. (2006). Evidence for the association of the SLC22A4 and SLC22A5 genes with Type 1 Diabetes: A case control study. *BMC Medical Genetics*, *7*(1), 54. https://doi.org/10.1186/1471-2350-7-54

Sato, T., Masuda, S., Yonezawa, A., Tanihara, Y., Katsura, T., & Inui, K. (2008). Transcellular transport of organic cations in double-transfected MDCK cells expressing human organic cation transporters hOCT1/hMATE1 and hOCT2/hMATE1. *Biochemical Pharmacology*, *76*(7), 894–903. https://doi.org/10.1016/j.bcp.2008.07.005

Schaeffeler, E., Hellerbrand, C., Nies, A. T., Winter, S., Kruck, S., Hofmann, U., van der Kuip, H., Zanger, U. M., Koepsell, H., & Schwab, M. (2011). DNA methylation is associated with downregulation of the organic cation transporter OCT1 (SLC22A1) in human hepatocellular carcinoma. *Genome Medicine*, *3*(12), 82. https://doi.org/10.1186/gm298

Schubert, M. S., Thommandru, B., Woodley, J., Turk, R., Yan, S., Kurgan, G., McNeill, M. S., & Rettig, G. R. (2021). *Improved methods and optimized design for CRISPR Cas9 and Cas12a homology-directed repair* (p. 2021.04.07.438685). bioRxiv. https://doi.org/10.1101/2021.04.07.438685

Schulz, C., Farkas, L., Wolf, K., Krätzel, K., Eissner, G., & Pfeifer, M. (2002). Differences in LPS-Induced Activation of Bronchial Epithelial Cells (BEAS-2B) and Type II-Like Pneumocytes (A-549). *Scandinavian Journal of Immunology*, *56*(3), 294–302. https://doi.org/10.1046/j.1365-3083.2002.01137.x

Selo, M. A., Sake, J. A., Ehrhardt, C., & Salomon, J. J. (2020). Organic Cation Transporters in the Lung—Current and Emerging (Patho)Physiological and Pharmacological Concepts. *International Journal of Molecular Sciences*, *21*(23), 9168. https://doi.org/10.3390/ijms21239168

Sharpe, J. J., & Cooper, T. A. (2017). Unexpected consequences: Exon skipping caused by CRISPR-generated mutations. *Genome Biology*, *18*(1), 109. https://doi.org/10.1186/s13059-017-1240-0

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., & Skarnes, W. C. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*, *11*(4), Article 4. https://doi.org/10.1038/nmeth.2857

Shi, W., Oshlack, A., & Smyth, G. K. (2010). Optimizing the noise versus bias trade-off for Illumina whole genome expression BeadChips. *Nucleic Acids Research*, *38*(22), e204. https://doi.org/10.1093/nar/gkq871

Shinozaki, Y., Furuichi, K., Toyama, T., Kitajima, S., Hara, A., Iwata, Y., Sakai, N., Shimizu, M., Kaneko, S., Isozumi, N., Nagamori, S., Kanai, Y., Sugiura, T., Kato, Y., & Wada, T. (2017). Impairment of the carnitine/organic cation transporter 1–ergothioneine axis is mediated by intestinal transporter dysfunction in chronic kidney disease. *Kidney International*, *92*(6), 1356–1369. https://doi.org/10.1016/j.kint.2017.04.032

Shrine, N., Portelli, M. A., John, C., Artigas, M. S., Bennett, N., Hall, R., Lewis, J., Henry, A. P., Billington, C. K., Ahmad, A., Packer, R. J., Shaw, D., Pogson, Z. E. K., Fogarty, A., McKeever, T. M., Singapuri, A., Heaney, L. G., Mansur, A. H., Chaudhuri, R., … Sayers, I. (2019). Moderate-to-severe asthma in individuals of European ancestry: A genome-wide association study. *The Lancet Respiratory Medicine*, *7*(1), 20–34. https://doi.org/10.1016/S2213-2600(18)30389-8

Shu, Y., Brown, C., Castro, R., Shi, R., Lin, E., Owen, R., Sheardown, S., Yue, L., Burchard, E., Brett, C., & Giacomini, K. (2008). Effect of Genetic

Variation in the Organic Cation Transporter 1, OCT1, on Metformin Pharmacokinetics. *Clinical Pharmacology & Therapeutics*, *83*(2), 273–280. https://doi.org/10.1038/sj.clpt.6100275

Simoff, I., Karlgren, M., Backlund, M., Lindström, A.-C., Gaugaz, F. Z., Matsson, P., & Artursson, P. (2016). Complete Knockout of Endogenous Mdr1 (Abcb1) in MDCK Cells by CRISPR-Cas9. *Journal of Pharmaceutical Sciences*, *105*(2), 1017–1021. https://doi.org/10.1016/S0022-3549(15)00171-9

Sjöstedt, E., Zhong, W., Fagerberg, L., Karlsson, M., Mitsios, N., Adori, C., Oksvold, P., Edfors, F., Limiszewska, A., Hikmet, F., Huang, J., Du, Y., Lin, L., Dong, Z., Yang, L., Liu, X., Jiang, H., Xu, X., Wang, J., … Mulder, J. (2020). An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science*, *367*(6482), eaay5947. https://doi.org/10.1126/science.aay5947

SLC22A5 solute carrier family 22 member 5 [Homo sapiens (human)]—Gene—NCBI. (n.d.). Retrieved 5 October 2019, from https://www.ncbi.nlm.nih.gov/gene/6584

Sledz, C. A., & Williams, B. R. G. (2005). RNA interference in biology and disease. *Blood*, *106*(3), 787–794. https://doi.org/10.1182/blood-2004-12-4643

Smith, D. A., Di, L., & Kerns, E. H. (2010). The effect of plasma protein binding on in vivo efficacy: Misconceptions in drug discovery. *Nature Reviews. Drug Discovery*, *9*(12), 929–939. https://doi.org/10.1038/nrd3287

Smith, E., Ottosson, F., Hellstrand, S., Ericson, U., Orho-Melander, M., Fernandez, C., & Melander, O. (2020). Ergothioneine is associated with reduced mortality and decreased risk of cardiovascular disease. *Heart*, *106*(9), 691–697. https://doi.org/10.1136/heartjnl-2019-315485

Smyth, G. K. (2005). limma: Linear Models for Microarray Data. In R. Gentleman, V. J. Carey, W. Huber, R. A. Irizarry, & S. Dudoit (Eds.), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (pp. 397–420). Springer. https://doi.org/10.1007/0-387-29362-0_23

Sonne, S., Shekhawat, P. S., Matern, D., Ganapathy, V., & Ignatowicz, L. (2012). Carnitine Deficiency in OCTN2−/− Newborn Mice Leads to a Severe Gut and Immune Phenotype with Widespread Atrophy, Apoptosis and a Pro-Inflammatory Response. *PLOS ONE*, *7*(10), e47729. https://doi.org/10.1371/journal.pone.0047729

Sporty, J. L., Horálková, L., & Ehrhardt, C. (2008). *In vitro* cell culture models for the assessment of pulmonary drug disposition. *Expert Opinion on Drug Metabolism & Toxicology*, *4*(4), 333–345. https://doi.org/10.1517/17425255.4.4.333

Sreenivasan Tantuan, S., & Viljoen, C. D. (2018). Imatinib Affects the Expression of SLC22A1 in a Non-Linear Concentration-Dependent Manner Within 24 Hours. *Medical Science Monitor Basic Research*, *24*, 59–62. https://doi.org/10.12659/msmbr.909124

Stavropoulou, V., Xie, J., Henriksson, M., Tomkinson, B., Imreh, S., & Masucci, M. G. (2005). Mitotic infidelity and centrosome duplication errors in cells overexpressing tripeptidyl-peptidase II. *Cancer Research*, *65*(4), 1361–1368. https://doi.org/10.1158/0008-5472.CAN-04-2085

Steimer, A., Haltner, E., & Lehr, C.-M. (2005). Cell Culture Models of the Respiratory Tract Relevant to Pulmonary Drug Delivery. *Journal of Aerosol Medicine*, *18*(2), 137–182. https://doi.org/10.1089/jam.2005.18.137

Stepanenko, A. A., & Dmitrenko, V. V. (2015). HEK293 in cell biology and cancer research: Phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene*, *569*(2), 182–190. https://doi.org/10.1016/j.gene.2015.05.065

Stewart, C. E., Torr, E. E., Mohd Jamili, N. H., Bosquillon, C., & Sayers, I. (2012). *Evaluation of Differentiated Human Bronchial Epithelial Cell Culture Systems for Asthma Research* [Research Article]. Journal of Allergy; Hindawi. https://doi.org/10.1155/2012/943982

Sun, D., Chen, Q., Gai, Z., Zhang, F., Yang, X., Hu, W., Chen, C., Yang, G., Hörmann, S., Kullak-Ublick, Gerd. A., & Visentin, M. (2021). The Role of the Carnitine/Organic Cation Transporter Novel 2 in the Clinical Outcome of Patients With Locally Advanced Esophageal Carcinoma Treated With Oxaliplatin. *Frontiers in Pharmacology*, *12*, 684545. https://doi.org/10.3389/fphar.2021.684545

Sutherland, R., Meeson, A., & Lowes, S. (2021). *SLCO and SLC22 Solute Carriers: Establishing Differential Expression Patterns in Human Breast Cancer Cell Lines and Tumour Sample*. https://doi.org/10.21203/rs.3.rs-349903/v1

Tamai, I. (2013). Pharmacological and pathophysiological roles of carnitine/organic cation transporters (OCTNs: SLC22A4, SLC22A5 and Slc22a21). *Biopharmaceutics & Drug Disposition*, *34*(1), 29–44. https://doi.org/10.1002/bdd.1816

Tamai, I., Yabuuchi, H., Nezu, J., Sai, Y., Oku, A., Shimane, M., & Tsuji, A. (1997). Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1. *FEBS Letters*, *419*(1), 107–111. https://doi.org/10.1016/s0014-5793(97)01441-5

Tan, E., Chin, C. S. H., Lim, Z. F. S., & Ng, S. K. (2021). HEK293 Cell Line as a Platform to Produce Recombinant Proteins and Viral Vectors. *Frontiers in Bioengineering and Biotechnology*, *9*. https://www.frontiersin.org/articles/10.3389/fbioe.2021.796991

The Human Protein Atlas. (n.d.). Retrieved 13 January 2023, from https://www.proteinatlas.org/

Thomas, J., Wang, L., Clark, R. E., & Pirmohamed, M. (2004). Active transport of imatinib into and out of cells: Implications for drug resistance. *Blood*, *104*(12), 3739–3745. https://doi.org/10.1182/blood-2003-12-4276

Thomas, P., & Smart, T. G. (2005). HEK293 cell line: A vehicle for the expression of recombinant proteins. *Journal of Pharmacological and*

Toxicological Methods, *51*(3), 187–200. https://doi.org/10.1016/j.vascn.2004.08.014

Tissue expression of SLC22A1—Summary—The Human Protein Atlas. (n.d.). Retrieved 14 October 2019, from https://www.proteinatlas.org/ENSG00000175003-SLC22A1/tissue

Tokuhiro, S., Yamada, R., Chang, X., Suzuki, A., Kochi, Y., Sawada, T., Suzuki, M., Nagasaki, M., Ohtsuki, M., Ono, M., Furukawa, H., Nagashima, M., Yoshino, S., Mabuchi, A., Sekine, A., Saito, S., Takahashi, A., Tsunoda, T., Nakamura, Y., & Yamamoto, K. (2003). An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nature Genetics*, *35*(4), Article 4. https://doi.org/10.1038/ng1267

Tronde, A., Nordén, B. o, Marchner, H., Wendel, A.-K., Lennernäs, H., & Bengtsson, U. H. (2003). Pulmonary Absorption Rate and Bioavailability of Drugs in Vivo in Rats: Structure–Absorption Relationships and Physicochemical Profiling of Inhaled Drugs. *Journal of Pharmaceutical Sciences*, *92*(6), 1216–1233. https://doi.org/10.1002/jps.10386

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D. H., Kim, J., Hyun Hwang, T., & Lum, L. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nature Communications*, *10*(1), 4056. https://doi.org/10.1038/s41467-019-12028-5

Tzvetkov, M. V., Matthaei, J., Pojar, S., Faltraco, F., Vogler, S., Prukop, T., Seitz, T., & Brockmöller, J. (2018). Increased Systemic Exposure and Stronger Cardiovascular and Metabolic Adverse Reactions to Fenoterol in Individuals with Heritable OCT1 Deficiency. *Clinical Pharmacology & Therapeutics*, *103*(5), 868–878. https://doi.org/10.1002/cpt.812

Urban, T., Brown, C., Castro, R., Shah, N., Mercer, R., Huang, Y., Brett, C., Burchard, E., & Giacomini, K. (2008). Effects of Genetic Variation in the Novel Organic Cation Transporter, OCTN1, on the Renal Clearance of Gabapentin. *Clinical Pharmacology & Therapeutics*, *83*(3), 416–421. https://doi.org/10.1038/sj.clpt.6100271

Verkuijl, S. A., & Rots, M. G. (2019). The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Current Opinion in Biotechnology*, *55*, 68–73. https://doi.org/10.1016/j.copbio.2018.07.005

Visentin, M., Gai, Z., Torozi, A., Hiller, C., & Kullak-Ublick, G. A. (2017). Colistin is Substrate of the Carnitine/Organic Cation Transporter 2 (OCTN2, SLC22A5). *Drug Metabolism and Disposition*, *45*(12), 1240–1244. https://doi.org/10.1124/dmd.117.077248

Volk, C., Gorboulev, V., Kotzsch, A., Müller, T. D., & Koepsell, H. (2009). Five amino acids in the innermost cavity of the substrate binding cleft of organic cation transporter 1 interact with extracellular and intracellular corticosterone. *Molecular Pharmacology*, *76*(2), 275–289. https://doi.org/10.1124/mol.109.054783

Wang, C., Uray, I. P., Mazumdar, A., Mayer, J. A., & Brown, P. H. (2012). SLC22A5/OCTN2 expression in breast cancer is induced by estrogen via a novel intronic estrogen-response element (ERE). *Breast Cancer Research and Treatment*, *134*(1), 101–115. https://doi.org/10.1007/s10549-011-1925-0

Wang, D., Zhang, C., Wang, B., Li, B., Wang, Q., Liu, D., Wang, H., Zhou, Y., Shi, L., Lan, F., & Wang, Y. (2019). Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12281-8

Wang, Q., Jiang, Z.-P., Zeng, J., Zhu, Y., Cai, H.-L., Xiang, D.-X., He, Q., Shi, X.-L., Zhong, A.-N., Zhao, X.-L., & Xu, P. (2020). Effects of Trough Concentration and Solute Carrier Polymorphisms on Imatinib Efficacy in Chinese Patients with Chronic Myeloid Leukemia. *Journal of Pharmacy & Pharmaceutical Sciences: A Publication of the Canadian Society for Pharmaceutical Sciences, Societe Canadienne Des Sciences Pharmaceutiques*, *23*(1), 1–9. https://doi.org/10.18433/jpps30559

Wang, T., Li, J., Chen, F., Zhao, Y., He, X., Wan, D., & Gu, J. (2007). Choline transporters in human lung adenocarcinoma: Expression and functional implications. *Acta Biochimica Et Biophysica Sinica*, *39*(9), 668–674. https://doi.org/10.1111/j.1745-7270.2007.00323.x

Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science (New York, N.Y.)*, *343*(6166), 80–84. https://doi.org/10.1126/science.1246981

Watt, F., & Molloy, P. L. (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development*, *2*(9), 1136–1143. https://doi.org/10.1101/gad.2.9.1136

Wei, J., Alfajaro, M. M., DeWeirdt, P. C., Hanna, R. E., Lu-Culligan, W. J., Cai, W. L., Strine, M. S., Zhang, S.-M., Graziano, V. R., Schmitz, C. O., Chen, J. S., Mankowski, M. C., Filler, R. B., Ravindra, N. G., Gasque, V., Miguel, F. J. de, Patil, A., Chen, H., Oguntuyo, K. Y., … Wilen, C. B. (2021). Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. *Cell*, *184*(1), 76-91.e13. https://doi.org/10.1016/j.cell.2020.10.028

Weibel, E. R. (1963). Geometric and Dimensional Airway Models of Conductive, Transitory and Respiratory Zones of the Human Lung. In E. R. Weibel (Ed.), *Morphometry of the Human Lung* (pp. 136–142). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-87553-3_11

Weinhold, L., Wahl, S., Pechlivanis, S., Hoffmann, P., & Schmid, M. (2016). A statistical model for the analysis of beta values in DNA methylation studies. *BMC Bioinformatics*, *17*(1), 480. https://doi.org/10.1186/s12859-016-1347-4

Weiss, J., Theile, D., Ketabi-Kiyanvash, N., Lindenmaier, H., & Haefeli, W. E. (2007). Inhibition of MRP1/ABCC1, MRP2/ABCC2, and MRP3/ABCC3 by Nucleoside, Nucleotide, and Non-Nucleoside

Reverse Transcriptase Inhibitors. *Drug Metabolism and Disposition*, *35*(3), 340–344. https://doi.org/10.1124/dmd.106.012765

Whisenant, T. C., & Nigam, S. K. (2022). Organic Anion Transporters (OAT) and Other SLC22 Transporters in Progression of Renal Cell Carcinoma. *Cancers*, *14*(19), 4772. https://doi.org/10.3390/cancers14194772

Wittekindt, O. H. (2017). Tight junctions in pulmonary epithelia during lung inflammation. *Pflugers Archiv*, *469*(1), 135–147. https://doi.org/10.1007/s00424-016-1917-3

Wu, J., Wang, Y., Liu, G., Jia, Y., Yang, J., Shi, J., Dong, J., Wei, J., & Liu, X. (2017). Characterization of air-liquid interface culture of A549 alveolar epithelial cells. *Brazilian Journal of Medical and Biological Research*, *51*(2). https://doi.org/10.1590/1414-431X20176950

Wu, X., Prasad, P. D., Leibach, F. H., & Ganapathy, V. (1998). cDNA Sequence, Transport Function, and Genomic Organization of Human OCTN2, a New Member of the Organic Cation Transporter Family. *Biochemical and Biophysical Research Communications*, *246*(3), 589–595. https://doi.org/10.1006/bbrc.1998.8669

Xu, H., Kita, Y., Bang, U., Gee, P., & Hotta, A. (2021). Optimized electroporation of CRISPR-Cas9/gRNA ribonucleoprotein complex for selection-free homologous recombination in human pluripotent stem cells. *STAR Protocols*, *2*(4), 100965. https://doi.org/10.1016/j.xpro.2021.100965

Yabuuchi, H., Tamai, I., Nezu, J.-I., Sakamoto, K., Oku, A., Shimane, M., Sai, Y., & Tsuji, A. (1999). *Novel Membrane Transporter OCTN1 Mediates Multispecific, Bidirectional, and pH-Dependent Transport of Organic Cations*. *289*, 6.

Yamada, R., Tokuhiro, S., Chang, X., & Yamamoto, K. (2004). SLC22A4 and RUNX1: Identification of RA susceptible genes. *Journal of Molecular Medicine*, *82*(9), 558–564. https://doi.org/10.1007/s00109-004-0547-y

Yang, H., Wang, J., Zhao, M., Zhu, J., Zhang, M., Wang, Z., Gao, Y., Zhu, W., & Lu, H. (2019). Feasible development of stable HEK293 clones by CRISPR/Cas9-mediated site-specific integration for biopharmaceuticals production. *Biotechnology Letters*, *41*(8–9), 941–950. https://doi.org/10.1007/s10529-019-02702-5

Yang, Y., Qiu, J.-G., Li, Y., Di, J.-M., Zhang, W.-J., Jiang, Q.-W., Zheng, D.-W., Chen, Y., Wei, M.-N., Huang, J.-R., Wang, K., & Shi, Z. (2016). Targeting ABCB1-mediated tumor multidrug resistance by CRISPR/Cas9-based genome editing. *American Journal of Translational Research*, *8*(9), 3986–3994.

Yasui, Y., Kudo, A., Kurosaki, M., Matsuda, S., Muraoka, M., Tamaki, N., Suzuki, S., Hosokawa, T., Ueda, K., Matsunaga, K., Nakanishi, H., Tsuchiya, K., Itakura, J., Takahashi, Y., Tanaka, S., Asahina, Y., Enomoto, N., Arii, S., & Izumi, N. (2014). Reduced Organic Anion Transporter Expression Is a Risk Factor for Hepatocellular Carcinoma in Chronic Hepatitis C Patients: A Propensity Score Matching Study. *Oncology*, *86*(1), 53–62. https://doi.org/10.1159/000356643

Yen, M.-C., Chou, S.-K., Kan, J.-Y., Kuo, P.-L., Hou, M.-F., & Hsu, Y.-L. (2018). Solute Carrier Family 27 Member 4 (SLC27A4) Enhances Cell Growth, Migration, and Invasion in Breast Cancer Cells. *International Journal of Molecular Sciences*, *19*(11), 3434. https://doi.org/10.3390/ijms19113434

You, G., & Morris, M. E. (2014). Overview of Drug Transporter Families. In G. You & M. E. Morris (Eds.), *Drug Transporters* (pp. 1–6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118705308.ch1

Zhang, S., Shen, J., Li, D., & Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics*, *11*(2), 614–648. https://doi.org/10.7150/thno.47007

Zhang, Y., Zhang, Y., Sun, K., Meng, Z., & Chen, L. (2019). The SLC transporter in nutrient and metabolic sensing, regulation, and drug development. *Journal of Molecular Cell Biology*, *11*(1), 1–13. https://doi.org/10.1093/jmcb/mjy052

Zhou, H., Zhu, Y., Qi, H., Liang, L., Wu, H., Yuan, J., & Hu, Q. (2021). Evaluation of the prognostic values of solute carrier (SLC) family 39 genes for patients with lung adenocarcinoma. *Aging*, *13*(4), 5312–5331. https://doi.org/10.18632/aging.202452

Zhou, K., Donnelly, L. A., Kimber, C. H., Donnan, P. T., Doney, A. S. F., Leese, G., Hattersley, A. T., McCarthy, M. I., Morris, A. D., Palmer, C. N. A., & Pearson, E. R. (2009). Reduced-function SLC22A1 polymorphisms encoding organic cation transporter 1 and glycemic response to

metformin: A GoDARTS study. *Diabetes*, *58*(6), 1434–1439. https://doi.org/10.2337/db08-0896

Zhu, G., Qian, M., Lu, L., Chen, Y., Zhang, X., Wu, Q., Liu, Y., Bian, Z., Yang, Y., Guo, S., Wang, J., Pan, Q., & Sun, F. (2019). O-GlcNAcylation of YY1 stimulates tumorigenesis in colorectal cancer cells by targeting SLC22A15 and AANAT. *Carcinogenesis*, *40*(9), 1121–1131. https://doi.org/10.1093/carcin/bgz010

Zhu, L. J. (2015). Overview of guide RNA design tools for CRISPR-Cas9 genome editing technology. *Frontiers in Biology*, *10*(4), 289–296. https://doi.org/10.1007/s11515-015-1366-y

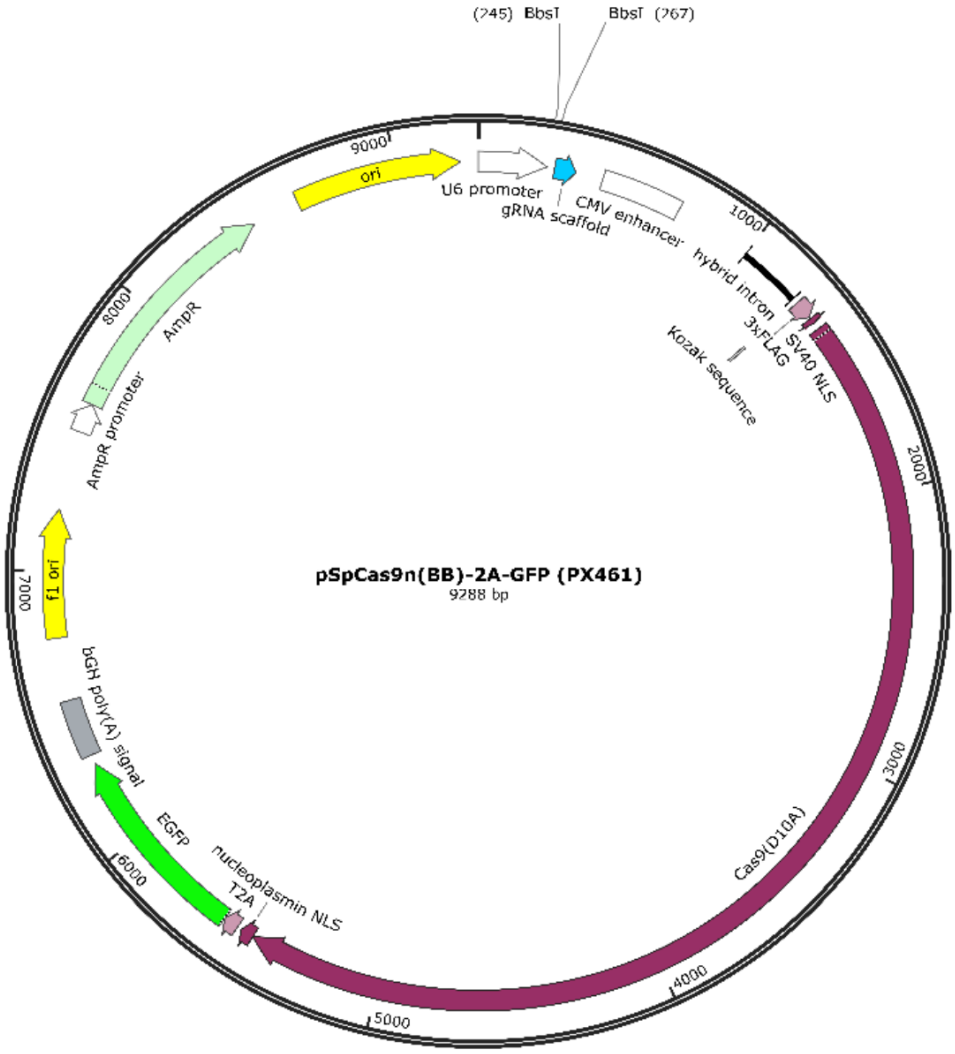
Zhu, L. J., Holmes, B. R., Aronin, N., & Brodsky, M. H. (2014). CRISPRseek: A Bioconductor Package to Identify Target-Specific Guide RNAs for CRISPR-Cas9 Genome-Editing Systems. *PLoS ONE*, *9*(9), e108424. https://doi.org/10.1371/journal.pone.0108424

Zhu, Q., Yu, L., Qin, Z., Chen, L., Hu, H., Zheng, X., & Zeng, S. (2019). Regulation of OCT2 transcriptional repression by histone acetylation in renal cell carcinoma. *Epigenetics*, *14*(8), 791–803. https://doi.org/10.1080/15592294.2019.1615354

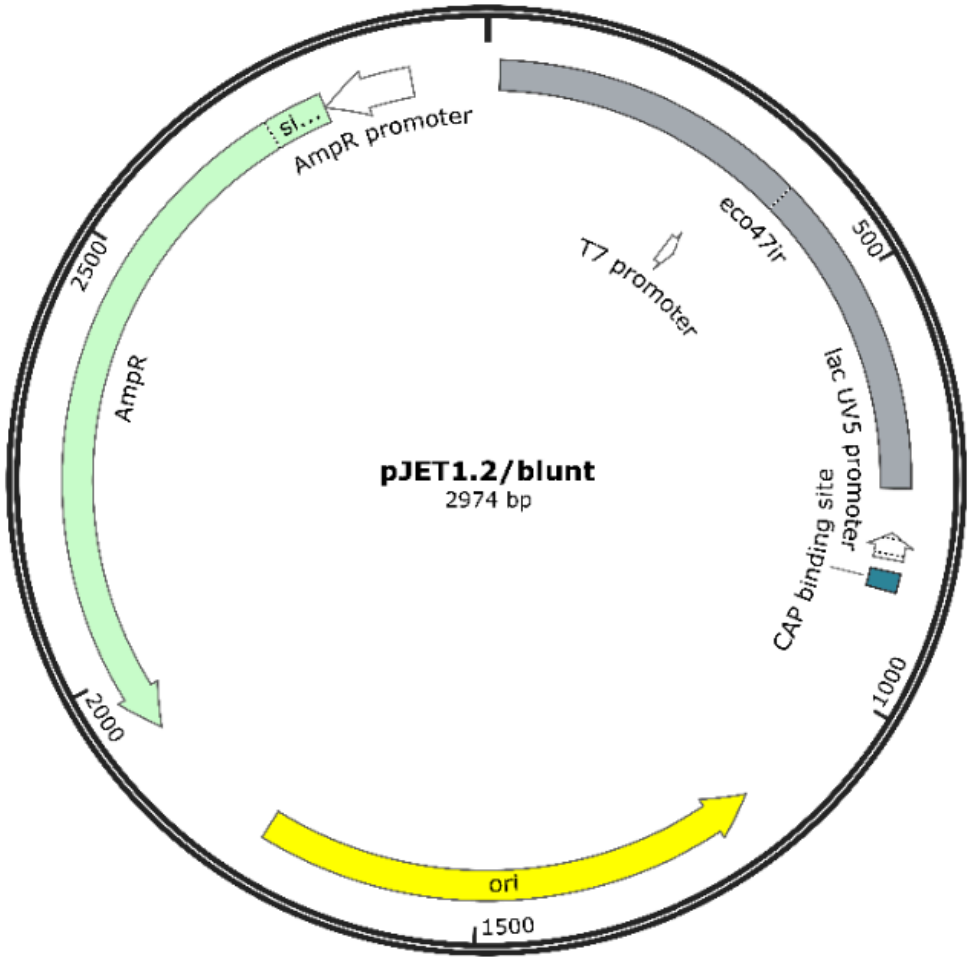
Zhu, Y., Chidekel, A., & Shaffer, T. H. (2010). Cultured Human Airway Epithelial Cells (Calu-3): A Model of Human Respiratory Function, Structure, and Inflammatory Responses. *Critical Care Research and Practice*, *2010*. https://doi.org/10.1155/2010/394578

APPENDIX

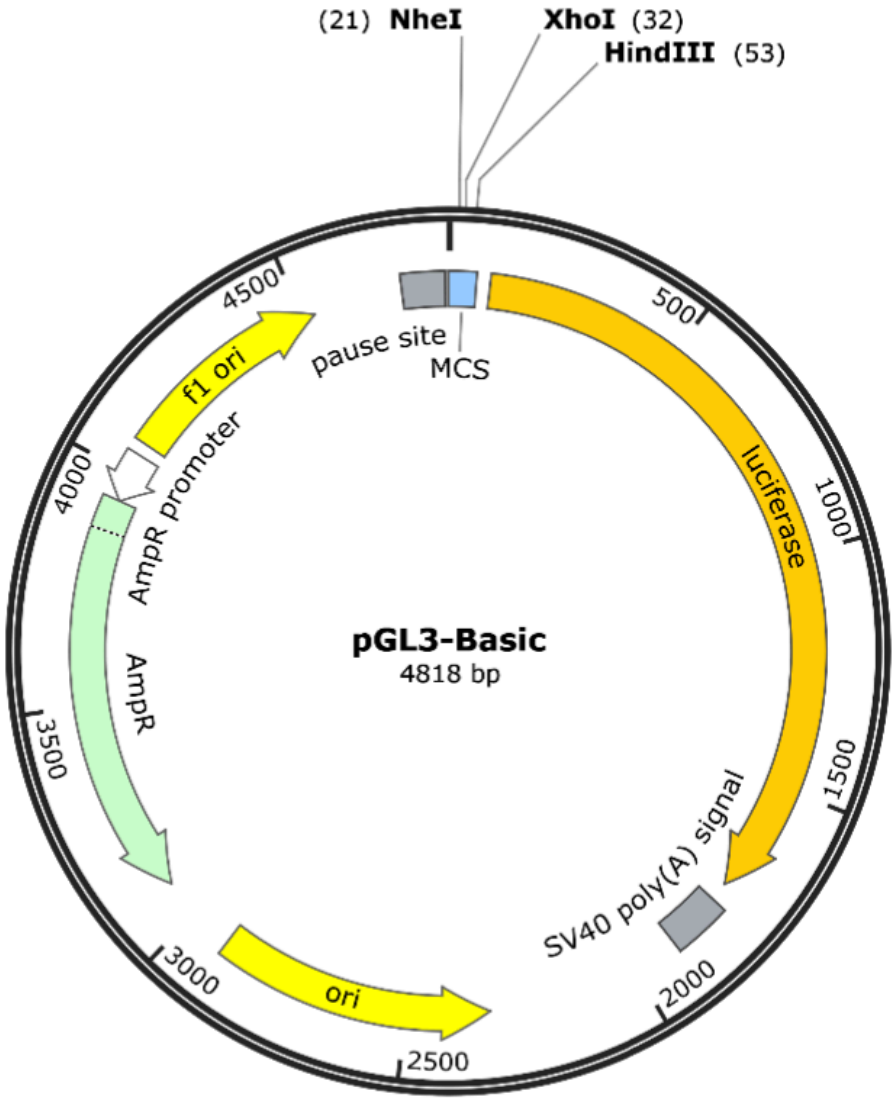
Appendix 1. Visual map of pX459



Appendix 2. Visual map of plasmid pX461



Appendix 3. Visual map pJET1.2/blunt



Appendix 4. Visual map of pGL3-Basic

Table 6-1. Differentially expressed genes (DEGs) in in asthma versus healthy controls from GEO dataset GSE67472. Genes are ranked by their Benjamini-Hochberg adjusted p-value. logFC: Log2 fold change, AveExpr: Average expression, t: t-statistic, P.Value: Uncorrected p-value, adj.P.Val: Benjamini-Hochberg adjusted p-value, B: B statistic.

Symbol	logFC	AveExpr	t	P.Value	adj.P.Val	B
CPA3	-2.1069678	7.55481922	-8.6366309	6.05E-14	1.26E-09	20.8644228
SERPINB2	-2.2151708	7.79949523	-8.0660123	1.13E-12	1.17E-08	18.1422035
CST4	-1.7282734	7.3863211	-7.5773983	1.33E-11	7.59E-08	15.8405346
PRR4	-2.2312421	7.60024631	-7.5594016	1.46E-11	7.59E-08	15.7564215
CLCA1	-3.0602064	6.4305941	-7.4275067	2.82E-11	9.60E-08	15.1416068
CDH26	-0.9332657	9.27255911	-7.4132409	3.03E-11	9.60E-08	15.0752864
TPSAB1	-0.9511781	7.14887268	-7.4004633	3.23E-11	9.60E-08	15.0159146
CST1	-2.4506792	6.07520662	-6.9211459	3.43E-10	8.93E-07	12.8115509
STEAP4	1.07838641	5.93156924	6.8439333	4.99E-10	1.09E-06	12.4610342
CEACAM5	-1.541956	9.65322194	-6.8336644	5.25E-10	1.09E-06	12.4145221
FHOD3	0.88980592	6.55584368	6.69272074	1.04E-09	1.97E-06	11.7787095
CST2	-0.8847663	5.54705168	-6.6150496	1.51E-09	2.62E-06	11.4304543
P2RY14	-1.5621379	4.92881259	-6.5579518	1.98E-09	3.18E-06	11.1754486
MUC5B	0.9865119	5.22253643	6.52452288	2.33E-09	3.46E-06	11.0265558
SCGB3A1	1.37274903	12.3226612	6.45273968	3.27E-09	4.54E-06	10.7078669
C12orf57	-0.4743093	7.33887216	-6.3173975	6.21E-09	8.08E-06	10.1109728
MREG	-0.5212472	7.2872829	-6.1230037	1.54E-08	1.89E-05	9.26323898
C3	1.19210996	9.52305695	6.07271974	1.95E-08	2.25E-05	9.04589172
TSPAN10	0.32618097	5.59918393	6.05008777	2.16E-08	2.37E-05	8.94833608
PCAT19	-1.2667129	8.34192693	-5.9830575	2.95E-08	3.07E-05	8.66039687
SCGB1A1	0.50361811	13.1772817	5.96072126	3.26E-08	3.24E-05	8.56478377
CYP2J2	0.39254416	8.39140797	5.93650665	3.65E-08	3.45E-05	8.46132211
POC1B	-0.3885324	8.31732456	-5.9218118	3.90E-08	3.53E-05	8.3986336
TFF3	-0.6763549	9.27869652	-5.8618478	5.13E-08	4.45E-05	8.14360518
TMEM45A	1.04689316	9.58561228	5.85173717	5.37E-08	4.48E-05	8.1007291
SVIP	-0.5527672	7.21859007	-5.7196414	9.78E-08	7.83E-05	7.54393078
TRIAP1	-0.3136237	9.62923943	-5.7010271	1.06E-07	7.93E-05	7.46598408
MRLN	-0.4843689	8.56456125	-5.6984117	1.08E-07	7.93E-05	7.45504276
TRIM9	0.20749108	3.65459052	5.69250678	1.10E-07	7.93E-05	7.43034892
EMC6	-0.3545545	5.38441765	-5.6777051	1.18E-07	8.05E-05	7.36850765
PXDN	-0.370876	6.37958055	-5.674399	1.20E-07	8.05E-05	7.35470569
NDUFB3	-0.2517868	9.32689047	-5.6610047	1.27E-07	8.28E-05	7.29883239
KIT	-0.6769657	6.96177166	-5.6496172	1.34E-07	8.45E-05	7.25138389
ARHGEF33	0.27774377	4.65485824	5.59360434	1.72E-07	1.05E-04	7.01871612
ATP6V1F	-0.3529378	8.42294771	-5.5682936	1.93E-07	1.12E-04	6.91397793

ARL1	-0.5420962	8.75487911	-5.5613928	1.99E-07	1.12E-04	6.88546544
SREK1IP1	-0.473088	9.72666336	-5.560122	2.00E-07	1.12E-04	6.8802169
TP53TG1	-0.4135026	7.54510367	-5.5515358	2.07E-07	1.13E-04	6.8447708
LDAH	-0.3346458	4.63080416	-5.5463572	2.12E-07	1.13E-04	6.82340608
IFI27L2	-0.3941348	7.40724694	-5.5123557	2.47E-07	1.28E-04	6.68339558
MUC12-AS1	-0.5755108	8.42252398	-5.5057781	2.54E-07	1.29E-04	6.65636351
FAM184A	0.60849631	7.89980752	5.48766955	2.75E-07	1.32E-04	6.58203204
CDC42EP5	-1.385674	7.53883745	-5.4854355	2.78E-07	1.32E-04	6.57287079
SH2D6	-0.2620289	5.44558045	-5.4839943	2.80E-07	1.32E-04	6.56696209
CISD3	-0.2572915	7.7319027	-5.4699644	2.98E-07	1.38E-04	6.50948426
MAP2K6	-0.536506	6.94978324	-5.4638832	3.06E-07	1.38E-04	6.48459544
RPL29	-0.4722763	8.8458631	-5.4573873	3.15E-07	1.39E-04	6.4580256
DNAH7	0.34886804	9.22050481	5.43188062	3.52E-07	1.52E-04	6.35386449
PCP4L1	-0.5166828	8.03219402	-5.4283673	3.57E-07	1.52E-04	6.33953818
UQCRB	-0.2563241	10.8011345	-5.4219389	3.68E-07	1.53E-04	6.31333763
H2BC12	-0.812346	8.2587309	-5.4044856	3.97E-07	1.62E-04	6.24228915
ZMAT2	-0.3281322	9.49607281	-5.3979145	4.08E-07	1.64E-04	6.21557209
HPGDS	-0.519376	6.27256201	-5.3887395	4.25E-07	1.67E-04	6.17829777
CCL26	-0.8156511	4.84668612	-5.371204	4.59E-07	1.77E-04	6.10715611
ROMO1	-0.3328719	9.98018151	-5.3429261	5.19E-07	1.97E-04	5.99270351
LOC101927499	-0.1650175	5.35729988	-5.3290942	5.51E-07	2.05E-04	5.93684266
ABHD17A	-0.2389623	4.70576044	-5.3210266	5.71E-07	2.09E-04	5.90429841
TTC7B	-0.6990798	7.38235962	-5.3134669	5.90E-07	2.12E-04	5.87382833
NACC1	-0.2928035	8.05381197	-5.297422	6.33E-07	2.23E-04	5.80923841
NDUFA12	-0.2943091	10.1153123	-5.2847281	6.68E-07	2.32E-04	5.75821651
FBXL14	0.2637177	4.4187131	5.27288521	7.04E-07	2.39E-04	5.71067716
SIGLEC6	-0.7129632	4.77160379	-5.2698835	7.13E-07	2.39E-04	5.69863748
CCNI	-0.2437446	11.0337475	-5.2668067	7.22E-07	2.39E-04	5.68630067
LRMDA	-0.2879539	6.42634132	-5.2113558	9.17E-07	2.99E-04	5.46466758
WNT5B	0.38563338	6.43874967	5.20267277	9.52E-07	3.05E-04	5.43008396
E2F2	-0.1989886	4.79743018	-5.1948195	9.85E-07	3.08E-04	5.39883373
ITLN1	-1.4239285	7.08662792	-5.1933602	9.91E-07	3.08E-04	5.39302983
FAM110C	-0.6176728	6.07384049	-5.187888	1.01E-06	3.11E-04	5.37127448
CYP2A13	0.63881644	6.38430908	5.18175164	1.04E-06	3.14E-04	5.34689406
IFI27	-0.6619313	8.97060115	-5.1768496	1.06E-06	3.16E-04	5.32742962
MCUB	-0.4765038	6.47719247	-5.17168	1.09E-06	3.16E-04	5.30691481
NDUFB6	0.33183621	4.2178935	5.17030638	1.09E-06	3.16E-04	5.30146569
KIAA1191	-0.3229366	8.91828104	-5.1491597	1.20E-06	3.41E-04	5.21768414
AVP	-0.2205011	4.6654235	-5.1463235	1.21E-06	3.41E-04	5.20646266
BCL10	0.28401693	3.87551713	5.14332774	1.23E-06	3.41E-04	5.19461369
CHPF	-0.3524327	5.92772036	-5.1386365	1.25E-06	3.43E-04	5.17606666
TMEM256	-0.3042127	8.57676426	-5.1330435	1.28E-06	3.47E-04	5.15396753
RPL27	-0.2182176	12.3061251	-5.1281322	1.31E-06	3.50E-04	5.13457362

DNAJC8	-0.2789469	7.75114681	-5.105895	1.44E-06	3.80E-04	5.04689892
ZNF467	-0.6986366	6.70431244	-5.0907841	1.54E-06	4.00E-04	4.98744923
TXNRD1	0.79424197	4.01115958	5.07768375	1.62E-06	4.17E-04	4.93599355
S100A16	-0.5686467	8.8438359	-5.073664	1.65E-06	4.19E-04	4.92022072
PRB1	-0.4953117	7.04121642	-5.0434153	1.88E-06	4.71E-04	4.80176651
PCOTH	-0.35732	5.80212263	-5.0320526	1.97E-06	4.84E-04	4.75737943
RAMP1	-0.365602	7.95875566	-5.030383	1.98E-06	4.84E-04	4.75086219
UCHL3	-0.4090075	7.59353056	-5.0254646	2.02E-06	4.84E-04	4.73167163
RABEP2	-0.3454193	4.98553156	-5.0229069	2.05E-06	4.84E-04	4.72169616
BATF	-0.2101484	6.05757559	-5.0228323	2.05E-06	4.84E-04	4.72140539
PYCR1	-0.2979136	6.64224612	-5.0198877	2.07E-06	4.85E-04	4.70992491
NOPCHAP1	-0.4183199	6.9084917	-5.0147632	2.12E-06	4.90E-04	4.68995513
PHYH	-0.4699233	9.16489106	-5.0085791	2.17E-06	4.95E-04	4.66587264
CHCHD10	-0.5753997	6.99954532	-5.0053296	2.20E-06	4.95E-04	4.65322546
DCTPP1	-0.3295527	8.67691334	-5.0046748	2.21E-06	4.95E-04	4.65067757
AP2S1	-0.2295274	6.32560853	-4.996435	2.29E-06	5.07E-04	4.61863257
LINC02145	0.39100354	5.24761708	4.99398256	2.31E-06	5.07E-04	4.60910082
MRPL34	-0.2387065	5.69988558	-4.9900673	2.35E-06	5.10E-04	4.59388967
NOP10	-0.4524209	8.28278858	-4.9837982	2.41E-06	5.18E-04	4.56954889
ALOX15	-0.485062	9.83648897	-4.9800566	2.45E-06	5.21E-04	4.55503048
PFDN1	-0.4130472	7.13107961	-4.9725138	2.53E-06	5.31E-04	4.52578201
LBH	-0.3415047	8.66311946	-4.9656705	2.60E-06	5.31E-04	4.49926937

Table 6-2. Differentially expressed genes (DEGs) in severe chronic obstructive pulmonary disease (COPD) lung tissue compared to controls with normal lung function from GEO dataset GSE76925. Genes are ranked by their Benjamini-Hochberg adjusted p-value. logFC: Log2 fold change, AveExpr: Average expression, t: t-statistic, P.Value: Uncorrected p-value, adj.P.Val: Benjamini-Hochberg adjusted p-value, B: B statistic.

ID	logFC	AveExpr	t	P.Value	adj.P.Val	B
LOC649214	-1.1385913	5.34367091	-9.9529295	2.64E-18	3.89E-14	30.9312986
KRT18P28	-1.0148232	6.15736133	-9.9167009	3.29E-18	3.89E-14	30.7177442
KRT8P9	-1.0825243	6.22100883	-9.72128	1.08E-17	8.52E-14	29.5685019
LOC645351	-0.6771051	5.25177287	-9.3816508	8.43E-17	4.98E-13	27.5832298
LOC643896	-0.7253207	5.09497156	-9.3361799	1.11E-16	5.24E-13	27.3187182
LOC643431	-1.3553169	5.84209967	-9.2698475	1.65E-16	6.50E-13	26.933439
LOC391359	-1.9176879	6.42082079	-9.2294836	2.10E-16	7.10E-13	26.6993422
ZNF143	0.49297554	8.62106728	9.11299916	4.22E-16	1.25E-12	26.0253053
LOC100130289	-1.4685319	6.0578756	-9.0880263	4.90E-16	1.29E-12	25.8811053
LOC391670	-0.8331726	5.56163904	-8.9221126	1.32E-15	3.02E-12	24.9259392
ANXA2P3	-0.8337259	6.7253921	-8.9113583	1.40E-15	3.02E-12	24.8642035
LOC648863	-0.9401184	5.59728083	-8.8863314	1.63E-15	3.21E-12	24.7206216
KRT18P17	-1.4215723	5.95010203	-8.8016613	2.69E-15	4.89E-12	24.2357678
NDUFA5	-1.238596	6.27099311	-8.7503759	3.64E-15	6.15E-12	23.9427841
LOC440563	-0.7680555	5.4801116	-8.6344011	7.21E-15	1.06E-11	23.2822505
LOC643550	-0.5975412	5.28520681	-8.629369	7.43E-15	1.06E-11	23.2536546
LIMS1	-0.5734106	7.33195906	-8.6243337	7.65E-15	1.06E-11	23.2250459
LOC388122	-0.4782599	5.21672171	-8.5944011	9.12E-15	1.20E-11	23.0550946
C1D	-1.0963203	5.81376864	-8.5620222	1.10E-14	1.37E-11	22.8714743
UBASH3B	-1.0619087	6.00654899	-8.5429611	1.23E-14	1.46E-11	22.7634871
VEZT	-0.4115634	7.62089428	-8.5128554	1.47E-14	1.66E-11	22.5930941
CENTD1	-0.8899427	5.30192404	-8.4995639	1.59E-14	1.71E-11	22.5179319
LOC727848	-0.9019739	5.92539286	-8.4868617	1.71E-14	1.76E-11	22.4461389
LOC149501	-1.303565	7.98254643	-8.4601711	2.00E-14	1.97E-11	22.2954047
LOC653324	-0.961903	5.44320894	-8.4255904	2.45E-14	2.26E-11	22.100355
LOC642076	-1.2512833	7.73891086	-8.4232984	2.48E-14	2.26E-11	22.0874368
TMEM65	-0.919706	5.49339188	-8.4041522	2.78E-14	2.43E-11	21.9795745
LOC100133770	-0.5214588	5.06369336	-8.3854411	3.10E-14	2.61E-11	21.8742466
CBWD1	-0.595553	5.45188125	-8.364039	3.51E-14	2.86E-11	21.7538721
C6orf62	-0.5342946	8.4231876	-8.3341692	4.17E-14	3.22E-11	21.5860551
GTF2H2	-1.1057797	5.38753368	-8.3274857	4.34E-14	3.22E-11	21.5485348
LOC651919	-1.4418014	6.00491218	-8.3267283	4.36E-14	3.22E-11	21.5442838
TOP1P1	-0.9648665	5.49074986	-8.2896923	5.40E-14	3.82E-11	21.336572
CARD16	-0.933119	6.95822791	-8.2868386	5.49E-14	3.82E-11	21.3205813

LOC649571	-0.6705585	6.12092383	-8.2076573	8.68E-14	5.82E-11	20.8776965
LOC729687	-1.5589154	6.69842021	-8.1999028	9.08E-14	5.82E-11	20.8344074
LOC100129499	-0.5840734	5.15343514	-8.199318	9.11E-14	5.82E-11	20.8311436
LOC347364	-0.2976874	4.80479936	-8.1882155	9.72E-14	6.04E-11	20.7691936
LOC388955	-0.5409132	5.17644391	-8.1762377	1.04E-13	6.31E-11	20.7023943
PEX3	-0.9203443	5.25194579	-8.1709869	1.07E-13	6.34E-11	20.673123
LOC641992	-1.6366022	6.64660441	-8.1664728	1.10E-13	6.35E-11	20.6479636
C4orf43	-1.1757111	5.70958408	-8.1361344	1.31E-13	7.39E-11	20.4790107
SLC26A2	-0.93846	7.57861886	-8.116722	1.47E-13	8.07E-11	20.3710296
LOC440926	-0.7522367	11.1955518	-8.1081045	1.54E-13	8.28E-11	20.3231262
LOC642585	-0.7757758	5.92989084	-8.1029796	1.59E-13	8.34E-11	20.2946468
LOC728602	-0.9109968	6.34340365	-8.0840352	1.77E-13	9.10E-11	20.1894324
LOC442153	-1.2011497	6.36892918	-7.9971962	2.91E-13	1.47E-10	19.7083677
KRT18P13	-0.4950436	6.96649395	-7.9878191	3.07E-13	1.51E-10	19.6565427
C1orf9	-0.5442106	5.34727162	-7.9826678	3.17E-13	1.52E-10	19.6280832
FRY	-1.6050767	6.43398064	-7.9805699	3.20E-13	1.52E-10	19.616495
LOC729255	-1.4457149	8.10581454	-7.9710486	3.38E-13	1.57E-10	19.5639171
ARL13B	-0.8561201	6.45735066	-7.9577563	3.65E-13	1.61E-10	19.4905572
P2RY14	-0.6342096	5.95970207	-7.9570065	3.67E-13	1.61E-10	19.4864202
LOC442727	-1.4663505	8.55585496	-7.9563398	3.68E-13	1.61E-10	19.4827424
XRCC4	-0.5774228	4.97959073	-7.9429418	3.97E-13	1.67E-10	19.4088536
CCDC82	-0.7539063	5.12016988	-7.9411809	4.01E-13	1.67E-10	19.3991461
RAP1B	-0.9746556	6.89596315	-7.9401475	4.04E-13	1.67E-10	19.3934492
LOC100130561	-0.5792552	8.13454822	-7.9221075	4.47E-13	1.82E-10	19.2940537
LOC731605	-1.4028477	5.90277243	-7.9082399	4.84E-13	1.94E-10	19.2177086
KCNK1	-1.0093825	7.4329246	-7.9010293	5.05E-13	1.99E-10	19.1780336
LOC100129067	-1.0692272	5.78483775	-7.8965739	5.18E-13	1.99E-10	19.1535256
LOC642443	-1.1675037	5.76815265	-7.895549	5.21E-13	1.99E-10	19.147889
SDAD1	0.30853347	11.0494631	7.88792127	5.44E-13	2.04E-10	19.1059464
SLC12A2	-1.1570967	8.37145335	-7.8820326	5.62E-13	2.08E-10	19.0735778
HSPCAL3	-1.0355343	5.7648621	-7.8670843	6.12E-13	2.20E-10	18.9914549
HSP90AB4P	-0.88912	5.39389794	-7.8664753	6.14E-13	2.20E-10	18.9881108
LOC730382	-1.3949455	7.47917049	-7.8608608	6.34E-13	2.21E-10	18.9572832
PPP3CB	0.39574196	10.7022873	7.86035795	6.36E-13	2.21E-10	18.9545225
LOC649260	-0.8398522	5.73624234	-7.8393337	7.17E-13	2.46E-10	18.8391682
SELT	-0.7565062	8.05239072	-7.8351588	7.34E-13	2.48E-10	18.8162769
UGP2	-0.5438133	8.14627102	-7.8085037	8.54E-13	2.84E-10	18.6702427
LOC391532	-1.0695498	8.5261417	-7.799108	9.01E-13	2.96E-10	18.6188166
LOC644380	-1.3771515	6.55284463	-7.7866958	9.66E-13	3.13E-10	18.550919
LOC100132086	-0.872605	5.81439924	-7.7750803	1.03E-12	3.30E-10	18.4874208
WDR85	0.57805362	5.72898134	7.76502915	1.09E-12	3.42E-10	18.432506
RNF214	0.41470958	7.16528867	7.76367381	1.10E-12	3.42E-10	18.4251034
PPIB	-1.7956332	6.70355861	-7.7608105	1.12E-12	3.43E-10	18.4094661

SNORD87	-0.9722856	5.59807608	-7.7456764	1.22E-12	3.69E-10	18.3268565
SNRPE	-0.6245152	5.11539972	-7.7425148	1.24E-12	3.71E-10	18.309607
LOC100133211	-1.6426416	6.28662169	-7.7343983	1.30E-12	3.84E-10	18.2653389
LOC646821	-1.098687	6.62986554	-7.7216575	1.40E-12	4.07E-10	18.1958884
PSMC6	-0.8306484	8.64609553	-7.7185859	1.42E-12	4.09E-10	18.179152
LOC645086	-0.6783421	5.49427892	-7.7144529	1.45E-12	4.14E-10	18.1566375
DNTTIP2	-1.6957111	6.12857783	-7.7058395	1.53E-12	4.30E-10	18.1097318
LOC653189	-0.5304586	5.02790941	-7.6946702	1.63E-12	4.52E-10	18.048941
LOC654074	-1.2143568	6.25654207	-7.6884537	1.68E-12	4.63E-10	18.0151228
CREM	-0.7851296	5.16680562	-7.6846547	1.72E-12	4.67E-10	17.9944621
ROBO2	-0.9486039	6.21735055	-7.6800901	1.76E-12	4.74E-10	17.9696431
LOC440525	-0.9180218	5.82951191	-7.6691851	1.88E-12	4.98E-10	17.9103751
KIAA1632	0.42892514	5.76987445	7.66682967	1.90E-12	4.99E-10	17.8975784
7A5	-0.3962819	4.96061309	-7.6418627	2.19E-12	5.64E-10	17.7620399
DACH1	-0.6908388	6.76615126	-7.6414639	2.19E-12	5.64E-10	17.7598763
LOC730535	-0.6336294	5.36997595	-7.6129412	2.57E-12	6.54E-10	17.6052733
PI4K2B	-0.7273776	7.57547655	-7.6073954	2.66E-12	6.68E-10	17.5752425
LOC642741	0.28566079	13.8186254	7.59705999	2.81E-12	7.00E-10	17.5193007
LOC728791	-1.2144804	7.6345934	-7.5892583	2.94E-12	7.17E-10	17.4770951
LOC100128291	-1.4162171	6.32153468	-7.5892033	2.94E-12	7.17E-10	17.4767976
ANXA2P1	-1.5946222	10.7015938	-7.5824661	3.05E-12	7.37E-10	17.4403666
C14orf125	-1.2655234	6.0629413	-7.5680652	3.31E-12	7.90E-10	17.3625418
MMRN2	-0.7492178	5.51016605	-7.5640644	3.38E-12	8.00E-10	17.3409321

Table 6-3. Top 100 Differentially Expressed Genes (DEGs) in TCGA-LUAD Ranked by Benjamini-Hochberg Adjusted P-Value. logFC: Log2 fold change. AveExpr: Average expression, t: t-statistic, P.Value: Uncorrected p-value, adj.P.Val: Benjamini-Hochberg adjusted p-value, B: B statistic.

gene_name	logFC	AveExpr	t	P.Value	adj.P.Val	B
AL606469.1	-5.9042668	-4.1159177	-41.095501	1.05E-176	2.53E-172	392.426171
SLC6A4	-7.690503	-0.0095352	-40.011772	1.35E-171	1.63E-167	381.929934
SERTM1	-7.3326534	-2.6249873	-37.76292	8.67E-161	6.97E-157	356.869312
AC095050.1	-4.8798601	-5.0514238	-37.317656	1.29E-158	7.76E-155	350.648206
ITLN2	-7.3146787	-1.9180069	-36.147855	7.31E-153	3.53E-149	338.811562
FABP4	-6.2193125	1.45718783	-34.823038	2.92E-146	1.17E-142	323.773086
LINC02016	-5.9744913	-3.9320254	-34.260376	1.97E-143	6.78E-140	316.621456
AC128709.3	-4.7224881	-4.856002	-33.616475	3.56E-140	1.07E-136	308.730456
GPM6A	-5.9567498	0.54793062	-33.507162	1.28E-139	3.42E-136	308.507317
AL354714.1	-5.1538428	-4.2232136	-33.109064	1.35E-137	3.26E-134	303.101338
LINC01996	-5.8380647	-3.3123436	-33.048362	2.76E-137	6.04E-134	302.709163
CD300LG	-6.3971418	-1.9258668	-32.808588	4.61E-136	9.27E-133	300.216723
RTKN2	-4.5514624	3.95977328	-32.042786	3.89E-132	7.21E-129	291.322549
STX11	-3.3467644	3.39477235	-31.423329	6.07E-129	1.05E-125	283.983402
RS1	-5.5956008	-3.0517193	-31.119959	2.25E-127	3.62E-124	280.055645
CLEC3B	-4.4519744	1.72490668	-31.081044	3.59E-127	5.40E-124	279.913248
SCUBE1	-4.3569521	1.4074433	-30.563302	1.75E-124	2.48E-121	273.736112
FAM107A	-5.0551684	3.14553684	-30.549834	2.05E-124	2.75E-121	273.573423
SH3GL3	-5.8319982	-2.7166617	-30.365577	1.87E-123	2.37E-120	271.162968
WNT3A	-5.4135441	-1.0317772	-30.273662	5.63E-123	6.79E-120	270.220575
TEK	-3.5141612	3.45714698	-30.143637	2.68E-122	3.08E-119	268.709077
MYOC	-5.4449814	-3.1897796	-29.965747	2.28E-121	2.50E-118	266.273341
LGI3	-6.8986136	0.2850376	-29.931513	3.44E-121	3.60E-118	266.170105
GPD1	-4.7533952	1.68950461	-29.80839	1.51E-120	1.52E-117	264.691084
ANGPT4	-4.8144365	-1.6485707	-29.76825	2.45E-120	2.37E-117	264.067502
UPK3B	-5.8704613	1.98827443	-29.717896	4.50E-120	4.17E-117	263.601594
HTR3C	-5.2939337	-4.1227727	-29.660078	9.03E-120	8.06E-117	262.446948
LANCL1-AS1	-4.3408264	-1.1930214	-29.605	1.75E-119	1.51E-116	262.125885
GPIHBP1	-4.582913	1.38022343	-29.546967	3.53E-119	2.94E-116	261.548687
AC093110.1	-3.6875541	0.24975409	-29.413771	1.76E-118	1.42E-115	259.905644
HSPA12B	-2.9140265	2.57450554	-29.208309	2.11E-117	1.64E-114	257.46891
KANK3	-3.1208001	1.82677415	-29.141412	4.73E-117	3.56E-114	256.661819
OTUD1	-2.1358316	4.31416195	-29.083278	9.55E-117	6.98E-114	255.947997
CHRM1	-5.585816	-3.4524128	-29.080658	9.85E-117	6.99E-114	255.642271
SGCG	-4.5957226	-1.7069937	-29.015048	2.18E-116	1.50E-113	254.995957
SPAAR	-2.7867856	-0.0335748	-28.956686	4.42E-116	2.96E-113	254.324275
STXBP6	-4.2436288	1.55113681	-28.917685	7.08E-116	4.61E-113	253.962176

CA4	-5.8052044	0.05389244	-28.868337	1.29E-115	8.17E-113	253.366601
FREM3	-5.0685337	-3.2462681	-28.850041	1.61E-115	9.93E-113	252.852233
AC104984.4	-4.4545069	-4.5234131	-28.770399	4.21E-115	2.54E-112	251.61178
NCKAP5	-4.4959745	1.9127037	-28.71745	8.00E-115	4.71E-112	251.538887
TAL1	-3.0073849	1.33213471	-28.621813	2.55E-114	1.47E-111	250.378266
ANKRD1	-5.6853901	0.53624299	-28.550866	6.03E-114	3.38E-111	249.527595
FENDRR	-4.5708173	1.61143109	-28.522764	8.48E-114	4.65E-111	249.185033
MYZAP	-3.6486497	0.42929262	-28.320872	9.84E-113	5.27E-110	246.724699
CYP1A2	-5.0650076	-4.0021166	-28.287205	1.48E-112	7.76E-110	245.959882
ST8SIA6	-4.176964	1.06119303	-28.176896	5.66E-112	2.90E-109	244.998712
AC135012.3	-4.5849843	-3.9607018	-28.11268	1.23E-111	6.20E-109	243.815033
CHRNA2	-4.5211144	-4.2545904	-28.083457	1.76E-111	8.67E-109	243.408431
ADCY8	-5.8314696	-3.5723085	-28.047687	2.72E-111	1.31E-108	243.202414
JAM2	-2.7304164	3.43731029	-27.968779	7.11E-111	3.36E-108	242.458592
CCM2L	-2.7525524	0.64831385	-27.804933	5.22E-110	2.42E-107	240.449525
RAMP2	-3.0618983	3.61750873	-27.792103	6.10E-110	2.78E-107	240.307479
PTPRQ	-5.33885	-1.8986044	-27.746265	1.07E-109	4.76E-107	239.699909
NECAB1	-3.771642	0.90952579	-27.698335	1.91E-109	8.38E-107	239.189387
EDNRB	-3.8905209	3.89765603	-27.687671	2.18E-109	9.38E-107	239.033524
MCEMP1	-5.0873878	1.93351568	-27.63867	3.96E-109	1.67E-106	238.452273
AC104237.3	-4.0979802	-4.6864035	-27.635955	4.09E-109	1.70E-106	237.869355
KCNA4	-5.2387924	-3.1535656	-27.606639	5.84E-109	2.39E-106	237.865979
LDB2	-2.7543457	3.88146319	-27.449968	3.95E-108	1.59E-105	236.144109
SEMA3G	-3.1763285	3.06570626	-27.292665	2.69E-107	1.06E-104	234.237495
TMEM100	-5.0663127	2.32299821	-27.102084	2.75E-106	1.07E-103	231.915946
PRX	-3.5159387	3.56118311	-26.873153	4.51E-105	1.73E-102	229.114748
C10orf67	-5.0834689	-1.9515848	-26.791114	1.23E-104	4.64E-102	228.078472
LIN7A	-3.1871665	1.73498769	-26.770087	1.59E-104	5.90E-102	227.889548
TNNC1	-4.4180676	2.25250204	-26.717113	3.04E-104	1.11E-101	227.223502
CD5L	-4.8210705	-3.0248185	-26.71485	3.13E-104	1.13E-101	227.033315
FCN3	-4.8867288	2.77383016	-26.673396	5.19E-104	1.84E-101	226.682845
BTNL9	-4.1013495	2.13775804	-26.648027	7.08E-104	2.47E-101	226.384269
MGAT3	-3.9200528	2.59849705	-26.646264	7.23E-104	2.49E-101	226.356939
ANGPTL7	-4.8935996	-2.9460621	-26.620867	9.87E-104	3.35E-101	225.902311
AL355499.1	-4.0156476	-4.3118297	-26.619446	1.00E-103	3.36E-101	225.6331
GYPE	-3.8076946	-1.3991726	-26.593873	1.37E-103	4.54E-101	225.64958
FOXF1	-2.9130198	3.01507858	-26.562186	2.02E-103	6.59E-101	225.33267
ADRB2	-3.548451	2.28293881	-26.431512	1.00E-102	3.22E-100	223.743825
AL590226.1	-3.8362543	-1.8194251	-26.316224	4.11E-102	1.30E-99	222.234582
RAMP3	-3.2859944	3.67976514	-26.256587	8.52E-102	2.67E-99	221.584414
S1PR1	-2.9927298	4.62082107	-26.229373	1.19E-101	3.68E-99	221.243245
ADAMTS7P3	-4.8112861	-0.9222654	-26.214612	1.43E-101	4.35E-99	221.100128
PLAC9	-2.8854667	1.96291814	-26.199106	1.72E-101	5.19E-99	220.91995
ARHGAP6	-2.6366636	2.10408122	-26.192412	1.87E-101	5.57E-99	220.838755

PTPN21	-2.5865703	4.20468876	-26.182323	2.12E-101	6.22E-99	220.673025
ADAMTS8	-4.5727061	1.61401838	-26.179062	2.20E-101	6.40E-99	220.660522
LIMS2	-3.0526433	4.02515091	-26.01459	1.65E-100	4.74E-98	218.621907
ROBO4	-3.0203364	4.08319923	-25.968118	2.92E-100	8.28E-98	218.052895
AC009093.3	-5.1053811	-0.8440732	-25.93965	4.14E-100	1.16E-97	217.74957
LYVE1	-3.5273125	2.67314274	-25.864703	1.04E-99	2.87E-97	216.807657
USHBP1	-2.4527306	1.12077965	-25.82566	1.67E-99	4.59E-97	216.356185
LINC02570	-3.9898203	-4.9754985	-25.71968	6.14E-99	1.66E-96	214.582521
ACVRL1	-2.6703078	5.09286936	-25.630298	1.84E-98	4.92E-96	213.911473
LINC00968	-4.0722098	-0.740628	-25.541339	5.47E-98	1.45E-95	212.865304
GRIA1	-5.2821015	-0.3569656	-25.516127	7.45E-98	1.95E-95	212.576828
CLEC1A	-2.4414253	1.56600767	-25.503571	8.69E-98	2.25E-95	212.424196
ARHGEF15	-2.6336552	3.23275255	-25.473155	1.26E-97	3.24E-95	212.014778
LRRC36	-4.0594302	1.04855111	-25.456867	1.54E-97	3.91E-95	211.84484
AL354714.3	-4.1614527	-4.6240297	-25.376813	4.11E-97	1.03E-94	210.502398
TCF21	-3.8864615	1.32773418	-25.355842	5.32E-97	1.31E-94	210.606183
LINC00656	-3.8399759	-3.9405713	-25.355693	5.33E-97	1.31E-94	210.31853
HBA2	-4.8620274	2.11192028	-25.344432	6.12E-97	1.49E-94	210.439432
CNTN6	-5.0577517	-0.4048305	-25.332658	7.07E-97	1.70E-94	210.333783

Table 6-4. Top 100 differentially expressed genes (DEGs) in Lung Adenocarcinoma (LUSC) identified by RNA-sequencing data analysis from The Cancer Genome Atlas (TCGA). Genes are ranked based on their Benjamini-Hochberg (BH) adjusted p-value. logFC: Log2 fold change, AveExpr: Average expression, t: t-statistic, P.Value: Uncorrected p-value, adj.P.Val: Benjamini-Hochberg adjusted p-value, B: B statistic

gene_name	logFC	AveExpr	t	P.Value	adj.P.Val	B
RS1	-6.7528219	-4.2248372	-51.497008	1.29E-213	2.67E-209	477.188643
AC008268.1	-8.3272938	-3.6330395	-49.392163	2.37E-205	2.45E-201	459.058559
CLEC3B	-5.4367913	0.64277828	-48.612402	3.11E-202	2.14E-198	452.465741
GPD1	-5.8737829	0.52497802	-48.010855	8.30E-200	4.29E-196	446.892918
CD300LG	-7.2957263	-2.8727356	-47.862578	3.31E-199	1.37E-195	445.092118
ADAMTS7P3	-5.7724502	-1.6794337	-47.686107	1.73E-198	5.94E-195	443.549777
AL606469.1	-5.9115459	-4.3780946	-47.375895	3.17E-197	9.36E-194	439.537909
LINC01996	-6.4185402	-4.2761652	-46.622307	3.91E-194	1.01E-190	432.683643
LINC02016	-6.3776305	-4.9227681	-45.159335	4.78E-188	1.10E-184	418.491025
AC236972.3	-6.3184325	-4.3786763	-45.050003	1.38E-187	2.84E-184	417.685805
NOSTRIN	-3.668571	1.85659472	-44.907037	5.50E-187	1.03E-183	417.444618
CHIAP2	-8.138534	-4.1194668	-44.592245	1.17E-185	2.02E-182	413.791063
CPB2	-7.04371	-1.6756629	-44.118223	1.20E-183	1.91E-180	409.647148
CA4	-7.0168439	-0.951823	-43.195903	1.07E-179	1.48E-176	400.645585
GPM6A	-6.9821305	-0.7728821	-43.195171	1.07E-179	1.48E-176	400.647798
C13orf46	-5.6157688	-1.3777627	-42.804805	5.20E-178	6.71E-175	396.65115
VEGFD	-5.978789	0.40540496	-42.758787	8.22E-178	9.99E-175	396.351863
GGTLC1	-7.696118	-1.8241993	-42.492325	1.18E-176	1.35E-173	393.611991
GPIHBP1	-6.1542238	0.18600611	-42.360493	4.40E-176	4.78E-173	392.376015
ADAMTS8	-6.1941885	0.23149918	-42.192907	2.36E-175	2.44E-172	390.700263
TNNC1	-6.1964608	0.40358132	-42.065171	8.51E-175	8.38E-172	389.422689
MCEMP1	-6.3436477	0.73203293	-42.044556	1.05E-174	9.84E-172	389.220363
TCF21	-5.4207442	-0.0330972	-41.684837	3.93E-173	3.53E-170	385.576325
HIGD1B	-3.7890266	0.15578773	-41.370385	9.48E-172	8.16E-169	382.331716
GKN2	-7.9537668	-1.5805108	-41.129632	1.09E-170	9.04E-168	379.927622
ECEL1P2	-6.56374	-3.0233313	-41.033393	2.91E-170	2.31E-167	378.645588
GPA33	-5.8085236	-1.5692418	-40.954412	6.51E-170	4.98E-167	378.058222
LRRC36	-4.9580285	-0.2644163	-40.884967	1.32E-169	9.75E-167	377.444451
CNTN6	-6.6121618	-2.1442423	-40.677843	1.10E-168	7.81E-166	375.228664
LANCL1-AS1	-4.3898255	-1.6912981	-40.649102	1.47E-168	1.01E-165	374.744988
SCUBE1	-5.3944526	0.30884659	-40.124766	3.20E-166	2.13E-163	369.708148
AGRP	-5.2594418	-3.071675	-40.084513	4.84E-166	3.13E-163	368.763808
EFCC1	-4.2311983	0.76258509	-39.856254	5.10E-165	3.19E-162	366.937706
AC027288.1	-5.223305	-3.6678213	-39.732084	1.84E-164	1.12E-161	364.960877

USHBP1	-3.3429758	0.45123642	-39.675544	3.30E-164	1.95E-161	365.011669
MYZAP	-4.7624864	-0.6474688	-39.612762	6.33E-164	3.63E-161	364.353419
KANK3	-3.8884873	1.09514466	-39.55995	1.09E-163	6.11E-161	363.880129
PLA2G1B	-6.5293211	-1.9181841	-39.547266	1.25E-163	6.79E-161	363.648753
HSD17B6	-4.363793	2.3329858	-39.530232	1.49E-163	7.89E-161	363.585315
GRIA1	-6.2619957	-1.3154951	-39.393627	6.15E-163	3.18E-160	362.107703
GDF10	-6.1839313	-0.8023191	-39.258979	2.49E-162	1.26E-159	360.740424
ARHGEF15	-3.6114396	2.53193698	-39.246934	2.83E-162	1.39E-159	360.64599
TEK	-4.3665138	2.67146618	-39.117369	1.09E-161	5.24E-159	359.298056
SLC6A4	-7.2867637	0.09945103	-39.063723	1.91E-161	8.95E-159	358.734636
AC027288.3	-4.6704959	-1.1385723	-39.060758	1.97E-161	9.03E-159	358.560548
FAM107A	-6.1780982	2.28681535	-38.653679	1.39E-159	6.23E-157	354.458442
AC112206.2	-4.6420198	-2.0017627	-38.584047	2.88E-159	1.27E-156	353.426306
F11	-6.5442441	-2.7984326	-38.569244	3.36E-159	1.45E-156	353.349386
RAMP2	-3.3745751	3.3606759	-38.426477	1.50E-158	6.35E-156	352.074067
RETN	-5.8193253	-1.59268	-38.378664	2.49E-158	1.03E-155	351.477704
INMT	-5.6839246	2.69484215	-38.285498	6.63E-158	2.69E-155	350.595711
PLAC9	-3.753092	1.17383306	-38.051579	7.80E-157	3.10E-154	348.128957
NPR1	-4.3283849	2.55184028	-37.981611	1.63E-156	6.37E-154	347.39453
AC093110.1	-4.1502485	-0.3348545	-37.978361	1.69E-156	6.47E-154	347.290804
AC116407.1	-4.5734035	-1.5970989	-37.786061	1.29E-155	4.85E-153	345.128631
MYOC	-6.1648462	-3.868114	-37.722993	2.52E-155	9.30E-153	344.169516
ROBO4	-3.8814608	3.5072897	-37.438157	5.18E-154	1.88E-151	341.637967
PRRT1B	-5.6225408	-2.8314145	-37.383179	9.29E-154	3.31E-151	340.757403
CLIC5	-6.2551515	2.81537091	-37.366483	1.11E-153	3.89E-151	340.881908
LINC00968	-4.9953766	-1.5199087	-37.325081	1.72E-153	5.94E-151	340.310882
FHL5	-4.6498763	-0.1135389	-37.238938	4.32E-153	1.46E-150	339.507731
AP001528.2	-4.0037358	-1.0314655	-37.158977	1.01E-152	3.37E-150	338.52364
CD5L	-5.6912564	-3.7414167	-37.116491	1.59E-152	5.23E-150	337.728208
MS4A15	-6.736197	-0.1829871	-37.015788	4.67E-152	1.51E-149	337.149186
CYP1A2	-5.4888918	-4.746722	-36.956945	8.75E-152	2.78E-149	335.715243
ANKRD1	-6.7348928	-0.2548462	-36.675922	1.77E-150	5.56E-148	333.516931
CTXND1	-5.0528698	-1.321098	-36.617715	3.31E-150	1.02E-147	332.80574
LDB2	-3.5341076	3.18744913	-36.412731	3.00E-149	9.11E-147	330.684784
PRX	-4.1079279	2.89614383	-36.327688	7.49E-149	2.24E-146	329.772133
EDNRB	-4.7852016	3.02648786	-36.264138	1.48E-148	4.38E-146	329.088487
ASPA	-4.4676178	-1.1222928	-36.216501	2.48E-148	7.22E-146	328.485896
ADRA1A	-5.8281285	-3.58627	-36.075973	1.13E-147	3.24E-145	326.69186
AL590226.1	-4.6465805	-2.8842151	-36.05209	1.46E-147	4.14E-145	326.43749
APOH	-6.5072599	-3.1393379	-35.990824	2.83E-147	7.92E-145	325.943685
C8B	-6.7519867	-3.027471	-35.941244	4.84E-147	1.33E-144	325.452377
MYCT1	-3.4005821	2.22214439	-35.864674	1.11E-146	3.02E-144	324.786752
ANGPT4	-5.0687633	-2.122187	-35.824251	1.72E-146	4.61E-144	324.174886

SPAAR	-3.6390589	-0.762253	-35.793873	2.39E-146	6.32E-144	323.907346
AL136452.1	-5.0551971	-4.3178775	-35.77435	2.95E-146	7.72E-144	323.159367
HSPA12B	-3.236437	2.27638643	-35.576996	2.51E-145	6.48E-143	321.672552
CACNA2D2	-5.4792629	2.81625582	-35.574272	2.58E-145	6.59E-143	321.637001
SLC46A2	-5.3882038	-0.9424782	-35.442074	1.09E-144	2.74E-142	320.191115
CHRM2	-5.3283121	-3.8381471	-35.288636	5.77E-144	1.44E-141	318.086673
SHE	-3.9830094	2.19644016	-35.220159	1.22E-143	2.99E-141	317.792136
VEPH1	-5.3908931	1.27425531	-35.141564	2.87E-143	6.97E-141	316.939645
TMEM100	-6.2599039	0.98109847	-35.135766	3.05E-143	7.34E-141	316.87754
ITLN2	-5.9149854	-1.1355449	-35.085454	5.29E-143	1.26E-140	316.320183
ARHGAP6	-3.6803368	1.20502044	-35.007686	1.24E-142	2.90E-140	315.486544
C1QTNF7	-4.7490481	0.00396361	-35.002984	1.30E-142	3.02E-140	315.432924
AL135999.3	-5.6065871	-3.5579643	-34.929367	2.91E-142	6.68E-140	314.284194
SFTA1P	-5.8725368	0.81687427	-34.826863	8.93E-142	2.03E-139	313.507185
AC104211.1	-4.6777248	-3.8435362	-34.824668	9.15E-142	2.06E-139	312.967091
RTKN2	-4.2876957	3.78750824	-34.76148	1.83E-141	4.06E-139	312.773935
EMCN	-4.2204937	2.50492477	-34.668514	5.06E-141	1.11E-138	311.764357
TNXB	-5.2682438	2.62397655	-34.653667	5.96E-141	1.30E-138	311.601872
NOVA2	-3.2827703	1.42352977	-34.628598	7.84E-141	1.69E-138	311.343324
AC108053.1	-4.0465271	-2.3653551	-34.59855	1.09E-140	2.32E-138	310.731504
CCM2L	-3.2908525	0.17592801	-34.523387	2.49E-140	5.25E-138	310.160287
FCN3	-5.6732599	2.26760533	-34.497217	3.32E-140	6.92E-138	309.889012
FREM3	-5.1087767	-3.8675978	-34.443574	5.98E-140	1.24E-137	308.867347