EVALUATION OF POTENTIAL PHOTODYNAMIC THERAPY AGENTS AND PATIENT-RELEVANT BIOMARKER COMBINATIONS FOR THE SELECTIVE TARGETING OF CANCER

José Ángel Rodríguez Corrales

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Chemistry

Brenda S. J. Winkel, Chair

John L. Robertson, Co-chair

Carla V. Finkielstein

Harry C. Dorn

August 2nd, 2018
Blacksburg, VA

Keywords: photodynamic therapy, excited state reactivity, DNA, heteromultivalency, melanoma, cytotoxicity, resazurin

Evaluation of potential photodynamic therapy agents and patient-relevant biomarker combinations for the selective targeting of cancer

José Ángel Rodríguez Corrales

ABSTRACT

Cancer, the second leading cause of death worldwide, is characterized by uncontrolled and abnormal cell growth. Even though researchers have made significant progress in its treatment over the past several decades, innovative therapeutic approaches that both improve patient survival and lessen the many debilitating side effects of conventional cancer treatments are vital. Accordingly, we first investigated the mechanism of interaction of a bimetallic complex, **Ru(II)-Rh(III)**, with DNA. Non-covalent binding of **Ru(II)-Rh(III)** is strong and involves electrostatic and, potentially, groove binding interactions. **Ru(II)-Rh(III)** photobinds and photocleaves DNA through an O_2-independent, metal-center mediated mechanism that could be beneficial in hypoxic tumors. Furthermore, the extent of covalent binding and cleavage of DNA, which inhibit PCR amplification, is dependent upon the strength of the non-covalent interactions. These results suggest that the toxicity of **Ru(II)-Rh(III)** could be selectively generated in tissues irradiated with light (e.g., a tumor). Secondly, we identified protein combinations selectively present in melanoma, which could be utilized in heteromultivalency. Heteromultivalent scaffolds display higher affinity towards cells that express a protein combination in comparison to those with only one of the proteins, which facilitates cell discrimination. Using an empirically-optimized threshold-based screening method and expression profiles of melanoma patients and normal

tissues, we identified surface proteins and protein combinations that are selectively found in melanoma patients and not in normal tissues. After a preliminary validation process using the scientific literature, we used immunofluorescence to confirm differential expression of some of these combinations in established melanoma cell lines in comparison to immortalized keratinocytes controls. Finally, we investigated the resazurin assay, a method used for the evaluation of proliferation and cytotoxicity in more than 2,000 publications. We found that only ~14% of these utilized validated assay conditions, while ~40% failed to report essential analytical parameters needed for their replication. We evaluated assay parameters needed for accurate estimation of cell number in eight cell lines, and found that these are highly variable and independent of tissue type, growth kinetics, and energetic parameters. Furthermore, we obtained some insights into the biochemical reduction of resazurin and proposed minimum reporting standards, along with a sample protocol for assay validation.

Evaluation of potential photodynamic therapy agents and patient-relevant biomarker combinations for the selective targeting of cancer

José Ángel Rodríguez Corrales

GENERAL ABSTRACT

Cancer, a group of diseases characterized by uncontrolled and abnormal cell growth, is the second-leading cause of death worldwide. Even though researchers have made significant progress in its treatment over the past several decades, innovative therapeutic approaches that both improve survival outcome and lessen the many debilitating side-effects of conventional cancer treatments are vital. First, we investigated the mechanism of interaction of a particular molecule, **Ru(II)-Rh(III)**, with DNA. We found that **Ru(II)-Rh(III)** is strongly attracted to DNA due to its charge and an interaction with the indentations along its helix. Upon light activation only, **Ru(II)-Rh(III)** binds to and cleaves DNA without the need for molecular oxygen, which is scarce in tumors and can limit the activity of other drugs, and to an extent that is affected by the concentration of ions in the solution. Thus, the cytotoxic effect of **Ru(II)-Rh(III)** might be selectively activated in those tissues that are irradiated with light (e.g., a tumor). Secondly, we identified protein combinations selectively present in melanoma, which could be utilized in heteromultivalency. Heteromultivalent scaffolds bind strongly to cells that express a combination of proteins rather than one protein at a time, making them excellent candidates for delivering a payload in a selective manner. Using expression profiles of melanoma and normal tissues, we identified surface proteins and protein combinations that are selectively found in melanoma patients and not in normal tissues. After a

preliminary validation process using the scientific literature, we used confirmed differences in the expression intensities of some of these combinations in melanoma cell lines in comparison to normal skin controls. Finally, we investigated the resazurin assay, a method used for the evaluation of cell growth and drug candidates in more than 2,000 publications. We found that only ~14% of these utilized validated assay conditions, while ~40% failed to report essential analytical parameters needed for their replication. We evaluated assay conditions for eight cell lines, and found that these are highly variable and independent of tissue type and some metabolic parameters. Furthermore, we obtained insights into the mechanism through which cells react with resazurin and proposed minimum reporting standards for publications, along with a protocol for assay validation.

ATTRIBUTIONS

Chapter 2 of this dissertation was written based on an accepted manuscript to which the author, Jose A. Rodriguez Corrales (J.A.R.C.), majorly contributed. Dr. Jing Wang synthesized the metal complex, performed lifetime and quantum yield measurements, and contributed intellectually during project design. J.A.R.C. performed all other experiments and conducted manuscript writing and editing. Dr. Brenda S.J. Winkel, co-principal investigator, contributed intellectually during the project and assisted in manuscript editing and revision. The late Dr. Karen J. Brewer, co-principal investigator, supervised the study and provided technical support during the experiments.

Chapter 3 was majorly written by J.A.R.C., who performed the experiments and bioinformatics screening using Python scripts of his own design, and conducted manuscript writing and editing. Dr. Jatinder S. Josan, principal investigator, supervised the study and provided feedback during experiment design and data analysis. Authors acknowledge the helpful contribution of Sadia Ahmed during screening validation and literature survey.

Chapter 4 was written based on a manuscript to which J.A.R.C. majorly contributed. J.A.R.C. participated in all experiments, and conducted manuscript writing and editing. Donald H. Clark, III was an undergraduate researcher on the project and contributed to several of the experiments. Kristýna Cagašová and Luke R. Vass were graduate and undergraduate researchers on the project, respectively, and optimized assay conditions in BT-474 and WM-266-4. Dr. Jatinder S. Josan, principal investigator, supervised the study, provided feedback during experiment design and participated in manuscript editing and revision.

Chapter 5 was written using a published manuscript to which the author, J.A.R.C., majorly contributed. Dr. Jatinder S. Josan, principal investigator, provided feedback and participated in manuscript editing and revision. Authors acknowledge the helpful editing work of Dr. Iulia M. Lazar.

PREFACE

During my tenure in graduate school, I have had the opportunity to participate in several projects, which focus lies beyond the scope of this dissertation. The following list contains the journal publications and oral and poster presentations that I have authored or coauthored.

1. Journal publications:

- Zhu, J.; Rodríguez-Corrales, J.Á.; Prussin, R.; Dominijanni, A.; Higgins, S.L.H.; Robertson, J.L.; Winkel, B.S.J.; Brewer, K.J. Exploring Cellular Activity of a Polyazine Bridged Ru(II)-Pt(II) Supramolecule in Rat Malignant Glioma F98. *Chem. Commun.* **2017**, *53*, 145-148.
- Xiao, L.; Li, T.; Ding, M.; Yang, J.; Rodríguez-Corrales, J.; Nacey, N.; Weiss, D.; Jin, L.; Dorn, H.; Li, X. Detecting chronic post–traumatic osteomyelitis of mouse tibia via an IL-13Rα2 targeted metallofullerene magnetic resonance imaging probe. *Bioconjugate Chem.* **2017**, *28* (2), 649-658.
- Zhu, J.; Dominijanni, A.; Rodríguez-Corrales, J.Á.; Prussin, R.; Zhao, Z.; Li, T.; Robertson, J. L.; Brewer, K. J., Visible light-induced cytotoxicity of Ru,Os–polyazine complexes towards rat malignant glioma. *Inorg. Chim. Acta.* **2017**, *454*, 155-161.
- Zigler, D.F., Morseth, Z.A.; White, T.A.; Canterbury, T.R.; Sayre, H.J.; Rodríguez-Corrales, J.A.; Brennaman, M.K.; Brewer, K.J.; Papanikolas, J.M. Ultrafast Kinetics of Supramolecules with a Ru(II)- or Os(II)-polypyridyl Light Absorber, cis-Rh(III)Cl_2-polypyridyl Electron Collector, and 2,3-bis(2-pyridyl)pyrazine Bridge. *Inorg. Chim. Acta.* **2017**, *454*, 266-274.

- Padilla, R.; Rodríguez-Corrales, J.Á.; Donohoe, L.E.; Winkel, B.S.J.; Brewer, K. J. A New Class of Ru(II) Polyazine Agents with Potential for Photodynamic Therapy. *Chem. Commun.* **2016**, *52*, 2705-2708.

2. Oral presentations:

- Ahmed, S.; Rodríguez Corrales, J.A.; Clark D.H.; Grams, R.J; Josan, J.S. In-situ nanoimprinting of cell surface receptors in cutaneous melanoma. Presentation given at the *Virginia Tech Graduate Student Assembly Annual Research Symposium and Exposition*, Blacksburg, Virginia, March 28, 2018.
- Rodríguez-Corrales, J.A.; Ahmed, S.; Josan, J.S. Beyond single targets: identification of protein combinations for heteromultivalent targeting of melanoma. Presentation given at the *1st Commonwealth of Virginia Cancer Research Conference*, Charlottesville, Virginia, Sep 22-23, 2017. *Presentation received the Trainee Award for the section.*
- Rodríguez-Corrales, J.A.; Yao, Z.; Clark, D.; Josan, J.S. Improving the synthesis of heteromultivalent scaffolds: towards precision medicine and melanoma theranostics. Presentation given at the *Graduate Student Assembly Annual Research Symposium and Exposition*, Blacksburg, Virginia, Mar 29, 2017.
- Padilla, R.; Rodríguez-Corrales, J.A.; Zhu, J.; Newman, J.; Prussin, R.; Brewer, K. J.; Naughton, E. Multifunctional Supramolecules for Interactions with DNA and Cancer Cells Exploiting Photochemical Activation. Presentation given at the *249th American Chemical Society National Meeting & Exposition*, Denver, Colorado, March 22-26, 2015.

- Rodríguez-Corrales, J.A.; Brewer, K.J. Polymetallic complexes for photodynamic therapy: a tribute to Karen J. Brewer. Presentation given at the *National Academy of Science*, San José, Costa Rica, Jan 12, 2015.
- Rodríguez-Corrales, J.A.; Brewer, K.J. Describing the Interaction of Supramolecular Systems and DNA. Presentation given at the *Journal Club for the Physical Sciences*, Virginia Tech, Blacksburg, USA, Oct 17, 2015.

3. Poster presentations:

- Rodríguez Corrales, J.A.; Ahmed, S.; Josan, J.S. Validation of surface receptor combinations for heteromultivalent targeting of melanoma. Poster presented at the *2017 Virginia Tech Center for Drug Discovery Poster Session*, Blacksburg, Virginia, Nov 2, 2017.
- Ahmed, S.; Rodríguez Corrales, J.A.; Josan, J.S. Investigation of in situ nanoimprinting–based phenotyping in cutaneous melanoma. Poster presented at the *2017 Virginia Tech Center for Drug Discovery Poster Session*, Blacksburg, Virginia, Nov 2, 2017.
- Staley, C.B.; Clark, D.H.; Grams, R.J.; Rodríguez-Corrales, J.A.; Josan, J.S. Synthesis of a Eu-bearing peptide for the versatile tagging of ligands. Poster at the *2017 Annual Biomedical Research Conference for Minority Students*, Phoenix, Arizona, Nov 1-4, 2017.
- Staley, C.B.; Clark, D.H.; Grams, R.J.; Rodríguez-Corrales, J.A.; Josan, J.S. Synthesis of a Eu-bearing peptide for the versatile tagging of ligands. Poster at the

Multicultural Academic Opportunity Program symposium, Blacksburg, Virginia, Jul 28, 2017.

- Rodríguez-Corrales, J.A.; Josan, J.S. Identification and validation of surface receptors for heteromultivalent targeting of melanoma. Poster at the *VirginiaCancerRx symposium*, Charlottesville, Virginia, May 30-31, 2017.
- Rodríguez Corrales, J.A.; Clark, D.H.; Kohnke, P.; Josan, J.S. Heteromultivalency in melanoma with GPCR ligands: Target validation and use of orthogonal chemistries towards a heterobivalent scaffold. Poster at the *253rd American Chemical Society National Meeting*, San Francisco, California, Apr 2-6, 2017.
- Clark, D.H.; Rodríguez Corrales, J.A.; Josan, J.S. Towards the synthesis of a heteromultivalent scaffold for cancer targeting. Poster at the *Virginia Tech's Department of Chemistry Undergraduate Research Symposium*, Blacksburg, Virginia, Dec 2, 2016.
- Clark, D.H.; Rodríguez Corrales, J.A.; Marron, D.; Josan, J.S. Metabolic variations, and Alamar Blue assay protocol standardization. Poster at the *Virginia Tech's Department of Chemistry Undergraduate Research Symposium*, Blacksburg, Virginia, Apr 22, 2016. *Poster was awarded 2nd place of the symposium.*
- Vass, L.R.; Rodríguez Corrales, J.A.; Josan, J.S. Photocytotoxicity of a Nanoencapsulated Bimetallic Compound and its Potential for Targeted Photodynamic Therapy of Melanoma. Poster at the *Virginia Tech's Department of Chemistry Undergraduate Research Symposium*, Blacksburg, Virginia, Dec 4, 2015.

- Rodríguez-Corrales, J.A.; Brewer, K.J. Non-covalent interactions of a Ru-Rh supramolecular complex with DNA and their effect on covalent modification of the biomolecule. Poster at the *249th American Chemical Society National Meeting*, Denver, Colorado, Mar 22-26, 2015. *Poster was selected for the Sci-Mix.*
- Rodríguez Corrales, J.A.; Knoll, J.D.; Newman, J.L., Brewer, K.J. Ru(II),Ru(II),Pt(II) Supramolecular Structures as Multifunctional Modifiers of DNA with Potential Application in Photodynamic Therapy. Poster at the *Annual Biomedical Research Conference for Minority Students*, Nashville, Tennessee, Nov 13-16, 2013.

ACKNOWLEDGMENTS

My tenure as a graduate student has been the most challenging and rewarding stage of my life. I have many people to thank for their help during the academic and personal challenges that I faced during my Ph.D.

First, I must thank God and my family. God gave me life and strength all the way to the end of my Ph.D., and I am thankful. Furthermore, I would not be who nor where I am without my parents' constant love, encouragement, and advice. My parents have worked hard their whole lives, sacrificing a lot for us, and I can only hope to make them proud with this accomplishment and my future life. I must also thank them for their unconditional understanding and support. They never complained when I could not call because I had to stay late in lab, nor did they ever forget to tell me how much they loved me and believed in me every time we talked.

Second, I must thank my advisors. My first advisor in graduate school, Dr. Karen Brewer, was a wonderful scientist, teacher, and human being who taught me how to become an independent and critical scientist. She constantly encouraged her group to help others become better scientists and to hold each other and ourselves to the highest standards. She had high expectations for all of us, but always remembered to compliment us on our accomplishments. Her untimely passing was a huge loss for the scientific community, academia, and the world at large.

My second advisor, Dr. Jatinder Josan is the epitome of an interdisciplinary scientist with an impressive background in synthetic organic and medicinal chemistry, pharmacology, and biology. He constantly challenged me to visualize multiple and complementary approaches to address caveats. On a personal level, Dr. Josan taught me several lessons, perhaps the most

important being that self-pity is one of our biggest enemies. Graduate school challenges us in many ways and feeling sorry about ourselves does not bring us any closer to a solution, or graduation.

Finally yet importantly is Dr. Brenda Winkel, whose generosity and mentorship skills are unparalleled. She actively contributed to my work in the Brewer group and made time out of her busy schedule as Department Head to discuss my research, both from the Brewer and Josan groups, and to answer my questions in molecular biology. Her willingness to join my committee and to revise my dissertation, especially Chapter 2, speak of her commitment to students. Dr. Winkel's constant encouragement and ideals with respect to education and science remind me of why I want to be a professor.

Third, I would like to acknowledge the faculty and staff at Virginia Tech. The former and current members of my committee have been contributed greatly to the development of my research project and dissertation. Dr. John Robertson provided constant encouragement throughout the complications of my graduate career, while his comments reminded me to look beyond chemistry and account for the needs of patients and clinicians. Dr. Carla Finkielstein always provided valuable and prompt feedback, her disposition and enthusiasm were reassuring even after heated discussions. Dr. Harry Dorn showed me the importance of collaboration and scientific curiosity for the development of new ideas. I would also like to thank my professors, Dr. Tijana Grove, Dr. Amanda Morris, Dr. Brian Tissue and Dr. Allan Esker, for teaching me new ways of approaching chemistry. Furthermore, Dr. James Tanko deserves a huge acknowledgement for his constant willingness to help me and other members of the Brewer Group. Even though we only had a few conversations, I always walked away knowing that he assist us in every way possible. Joli Huynh deserves not only my thanks, but those of every graduate student in the department

because we would be lost without her. Thanks to Dr. Mehdi Ashraf-Khorassani for his helpful encouragement and pleasant conversations.

I must also thank my Chemistry friends and family at Virginia Tech, present and past. I must start with my first mentor, Dr. Roberto Padilla, who taught me the basics of research. I have lost track of how many times I have quoted his teachings to graduate students under my mentorship. My friend, colleague, and Chinese brother, Dr. Jie Zhu, always came to my help in times of professional and personal need. I am grateful and proud of meeting such a talented, kind and clever individual whose achievements include an amazing number of publications. I thank previous members of the Brewer Group, particularly Dr. Elise Naughton, Dr. Theodore Canterbury, Dr. Jing Wang, Kristen Noble, and Diana Baumgardt. I also thank previous and current members of the Josan Group, including Justin Grams, Kristýna Cagašová, Ashley Gates, Sadia Ahmed, Donald Clark, Daniel Marron, and Luke Vass. Thanks to my friends throughout the Department of Chemistry, Dr. Paula Celis-Salazar, José Santiago-Rivera, Dr. Marwa Abdel-Latif, Dr. Tinghui Li, and Dr. Hadi Mohammadi. For those that I did not mention, please forgive my faulty memory (and limited space) and trust that I am thankful for your friendship and help.

Finally, I want to thank all the friends who supported me during graduate school. This was a mentally taxing experience and I would not have made it without you all. Special thanks to all my coffee friends: Ashley Gates, Kristýna Cagašová, Odiney Álvarez Campos, Dr. Lily Virgüez, José Santiago-Rivera, Dr. Paula Celis-Salazar, Justin Grams, and Danny Marron. Thank you to all the wonderful people I have met from all around the world, including Brian Bess, Oxana Brenes Angulo, Didier Mena-Aguilar, Dr. Óscar Galagarza, Ariana Umaña, Ricardo Cruz, Dr. Carlos Landaverde Alvarado, Dr. Ana Lisa Valenciano Murillo, Melissa Brenes Bastos, Dr. Pilar Madrigal Quesada, and Giacomo Alloca.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Cancer

Cancer is the general term used for a series of diseases characterized by abnormal and uncontrolled cell proliferation. From a biological perspective, the transformation of a normal cell into a cancer cell occurs through three basic steps: initiation, promotion, and progression. Initiation involves the irreversible modification of one or more genes that regulate cellular growth and differentiation. Promotion is defined as the proliferation of cells with the aforementioned mutation(s), thus allowing their stabilization in the genome and replication into new cells. Finally, progression is the buildup of four to five mutations and the subsequent expression of the typical behavior attributed to cancer, including abnormal cell growth and the continued evolution of increasingly aggressive phenotypes.[1-2]

Despite the high heterogeneity found among different cancers, there are ten general organizing principles, or hallmarks, that characterize this disease, as presented in Fig. 1.1. Hanahan and Weinberg proposed these hallmarks in two seminal publications describing the processes through which normal cells become cancer cells.[3-4] These principles address both primarily intracellular processes (i.e., genome instability and mutation, sustaining proliferative signaling, deregulating cellular energetics, evading growth suppressors, enabling replicative immortality, and resisting cell death) and the interaction of cancer cells with their environment (i.e., inducing angiogenesis, activating invasion and metastasis, promoting inflammation, and avoiding immune destruction). As can be inferred from these hallmarks, cancer behavior can be expected to maximize cell proliferation *in* or *ex situ* while promoting access to nutrients and evading the immune system.

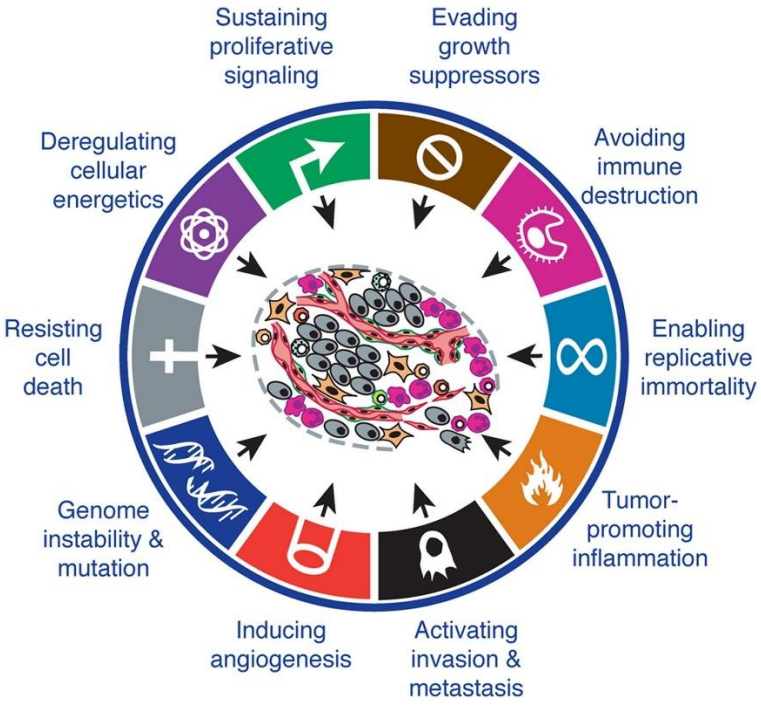


Figure 1.1. The hallmarks of cancer. Figure was adapted from Hanahan, D.; Weinberg, R. A., The Hallmarks of Cancer. *Cell* **2000,** *100* (1), 57-70, with permission from Elsevier.

1.2. Research motivation

The aggressive phenotype commonly observed in cancer has made it a feared and prevalent disease. In 2012, cancer accounted for 8.4 million deaths worldwide, whereas malignant neoplasms were consistently ranked as the second leading cause of death throughout the globe.[5] The American Cancer Society estimated that more than 1.7 million cancer diagnoses and close to 600,000 deaths will occur in the United States in 2018.[6] In terms of its financial implications, a cancer diagnosis can be extraordinarily burdensome. The National Cancer Institute of the US assessed that the total cost of care in 2010 was approximately $125 billion and projected a 27% increase by 2020.[7] Indirect costs associated with cancer, which include time lost from work and lost productivity due to premature death, were estimated in the order of €75 billion in Europe in

2008.[8] Unquestionably, the human component of this disease should not be overlooked: patients and their families face costly and aggressive treatments and, in some cases, unfavorable prognoses that negatively impact their well-being and lifestyle.

Anti-cancer drugs interact with intra- or extracellular epitopes of the cancer cells or the immune system in order to address one or more of the hallmarks of cancer.[3] For instance, *cis*-diamminedichloroplatinum (II), or cisplatin, was one of the first cancer drugs to be discovered and remains widely used due to its efficacy in treating several types of cancer, most notably testicular cancer.[9-11] This drug triggers apoptosis by generating covalent adducts with DNA, a mechanism that is shared with other platinum derivatives in addition to a series of molecules known as alkylating agents (e.g., chlorambucil and temozolomide).[12-14] Although these drugs are efficacious towards the tumor, they often generate off-target effects ranging from nausea to nephrotoxicity, neutropenia, and anemia.[15-17] In summary, notwithstanding the considerable advances in cancer treatment over the last decades, there is a pressing need for the development of drugs that selectively target cancer cells, and/or can be selectively activated in the tumor tissue, and that are less likely to result in unwanted and often-debilitating side effects.

1.3. Photodynamic therapy

Photodynamic therapy (PDT), which is an alternative to traditional chemotherapy, is currently used in the treatment of cancer. PDT drugs, also known as photosensitizers, are administered as nontoxic compounds, whereby cytotoxicity is triggered upon photoexcitation, thus restricting the toxic effect to those areas where an external light source is applied.[18-19] Effective photosensitizers should display strong absorption in the phototherapeutic window, a range of wavelengths from 600 to 800 nm that are not readily absorbed by endogenous skin pigments (e.g.,

melanin and hemoglobin), thereby enabling them to penetrate deeper into the skin (see Fig. 1.2).[20-21]

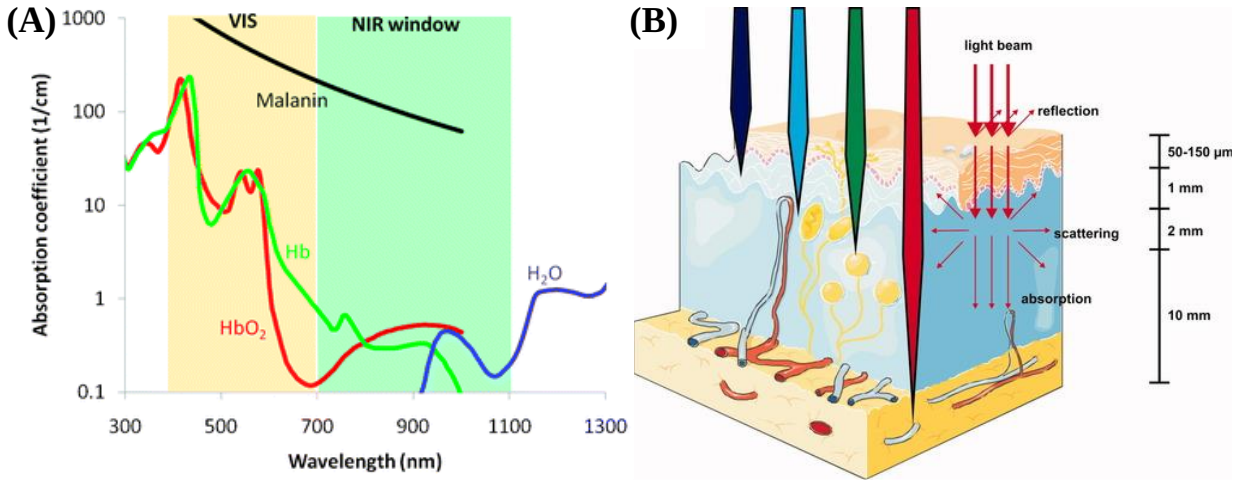


Figure 1.2. Phototherapeutic window. Absorbance spectra of endogenous skin pigments (**A**), including melanin, hemoglobin (Hb) and oxyhemoglobin (HbO_2), and penetration depth at different wavelengths (**B**). Figure adapted from Qin, Z.; Bischof, J. C., Thermophysical and biological responses of gold nanoparticle laser heating. *Chem. Soc. Rev.* **2012,** *41* (3), 1191-1217 with permission from the Royal Society of Chemistry, and Agostinis *et al* Photodynamic therapy of cancer: an update. *CA – Cancer J. Clin.* **2011,** *61* (4), 250-281 with permission from John Wiley and Sons.

Photodynamic therapy agents are classified into three distinct categories based on the mechanism through which they exert cytotoxicity, as described in Fig. 1.3. Photosensitizer activation begins when a molecule in the singlet ground state (^{1}GS) absorbs a photon to populate a singlet excited state (^{1}ES), followed by decay into a triplet excited state (^{3}ES) through intersystem crossing. The excited molecule can then relax through several different pathways, including

radiative decay, non-radiative decay, energy transfer, or electron transfer.[18, 22] In Type I PDT drugs, the ^{3}ES undergoes electron transfer to water, thus generating radicals that can oxidize biological targets; in contrast, Type III drugs are able to react directly with biomolecules. Type II photosensitizers experience triplet-triplet annihilation with ground state molecular oxygen to generate singlet oxygen, a strong oxidant that reacts with cellular targets. This form of oxygen is not generated during normal cell growth; thus, the cell has no effective mechanism to defend against it.[23-24] The two PDT drugs approved for the treatment of cancer, Photofrin® and Levulan®, undergo this Type II mechanism.[18-19]

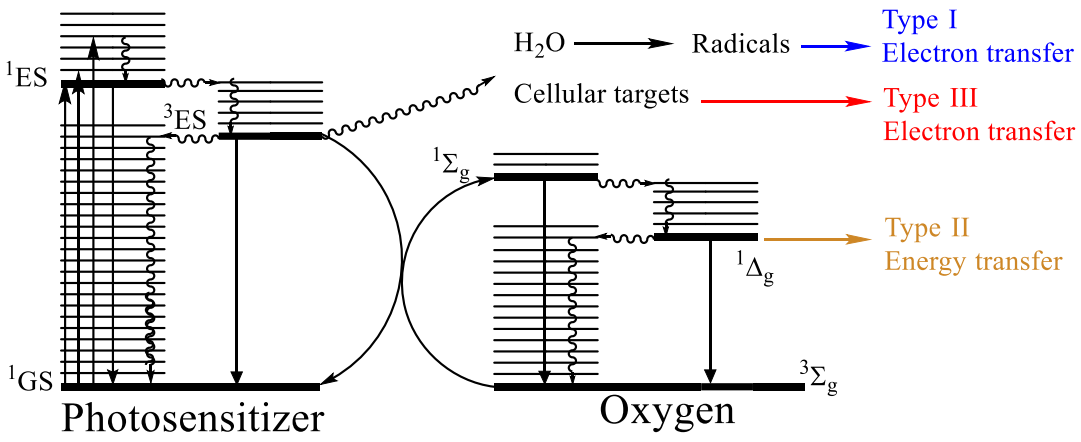


Figure 1.3. Mechanisms of action of PDT agents. ^{1}GS, ^{1}ES, and ^{3}ES correspond to the singlet ground state, singlet excited state, and triplet excited of the photosensitizer. $^3\Sigma_g$ is the triplet ground state of molecular oxygen, whereas $^1\Sigma_g$ and $^1\Delta_g$ are its two singlet excited states. Figure was adapted from references 18 and 25.

Photodynamic therapy has proven to be successful in the treatment of several types of cancer. Photofrin® and Levulan® are currently approved in several countries (including the US) for use in lung, esophagus, bile duct, bladder, brain and ovarian cancer.[21] However, these drugs

possess a Type II mechanism that relies on molecular oxygen, which is not readily available in the hypoxic regions commonly found in tumors. Thus, their efficacy is limited by the oxygen diffusion rates to the tumor.[26-27] The design of novel PDT agents with oxygen-independent activity and long wavelength activation within the phototherapeutic window is desirable and currently under investigation.[28-30]

1.4. Heteromultivalency

Multivalent interactions, defined as the formation of multiple simultaneous molecular contacts, are common in nature.[31] Perhaps the best example is double-stranded DNA, where each pair of complementary nucleobases is held together through hydrogen bonding, thus generating a biopolymer whose secondary structure is stabilized by a myriad of individual interactions.[32] Multivalent scaffolds display an overall or apparent binding affinity that is stronger than that of the individual binding events, allowing biological systems to strengthen the interaction between ligands and receptors. In particular, multivalency plays a critical role in cell surface adhesion and viral entry, as do other events that rely on non-covalent protein-carbohydrate and protein-protein complexes, as reviewed elsewhere.[31, 33-35]

Heteromultivalency represents a novel approach for developing precision medicines that relies on combination-based targeting. Heteromultivalent scaffolds contain ligands for multiple heterologous receptors. Therefore, their apparent binding is higher in those cells that express the cognate receptor combination on their surface. This principle, known as expression pattern recognition, allows for cell discrimination.[36-38] In contrast to several targeting approaches that rely on receptor overexpression in tumors,[39-41] heteromultivalency aims to targets protein combinations that are exclusively coexpressed in the neoplasm, leaving normal (i.e., mono-expressing) tissues unaffected.

The effectiveness of heterobivalent scaffolds has been explored *in vitro* for different receptor combinations. These include Melanocortin receptor 4 (MC4R)/δ-opioid receptor (δ-OR),[37] D_2-Like Receptor (D_2-likeR)/μ-Opioid Receptor (μ-OR)[42] and adenosine A_{2A} receptor (A_{2A}R)/dopamine D_2 receptor (D_2R), among others.[43-44] A scaffold bearing ligands towards the MCR4 and Cholecystokinin-2 receptor (CCK-2R), bridged through a semirigid Pro-Gly linker, displayed up to an 81-fold higher affinity towards dual-expressing cells compared to a mono-expressing control.[36, 45] The level of selectivity was dependent on the chemical structure and length of the linker and the identity of the receptor. Furthermore, imaging of cells with different expression profiles showed that a fluorescently-tagged heteromultivalent ligand bound to the cytoplasmic membrane and displayed higher selectivity towards those cells that expressed both receptors (Fig. 1.4).[36]

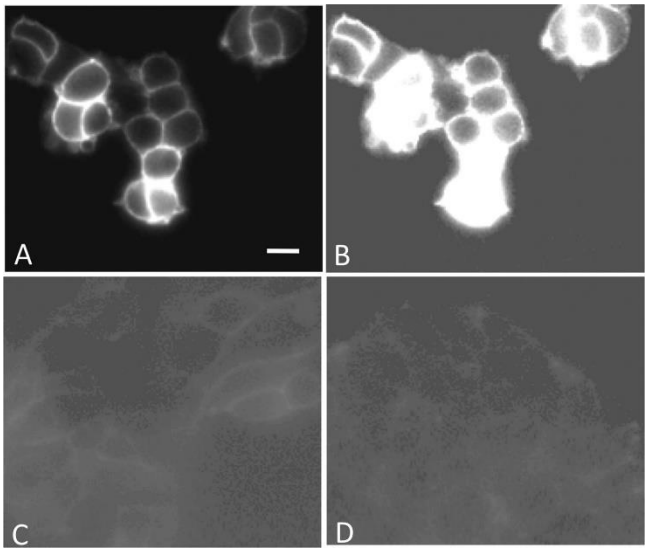


Figure 1.4. Binding of a heterobivalent ligand with a fluorescent tag in cells with different expression profiles. MC4R/CCK-2R dual-expressing cells incubated with ligand (**A** and **B**), in comparison to cells expressing CCK-2R (**C**) or MC4R (**D**) only. Image **B** corresponds to the same cells as in **A**, but captured under the imaging conditions used in **C** and **D**. Reproduced with permission from Josan *et al.* Cell-Specific Targeting by Heterobivalent Ligands. *Bioconjugate Chem.* **2011,** *22* (7), 1270-1278. Copyright (2011) American Chemical Society.

Similar results have been observed with other scaffolds at the cell and tissue levels. A heterobivalent ligand towards $A_{2A}R$ and D_2R was evaluated in brain striatal membranes, which expresses both receptors, and showed enhanced affinity (a 7 to 18-fold increase) compared to monovalent ligands. Mono-expressing cell controls showed comparable binding affinity for the bivalent and monovalent ligands.[44] Similarly, another heterobivalent ligand with affinity towards a protein combination natively expressed in β-cells, Glucagon-Like Peptide 1 (GLP-1)/Yohimbine (Yhb), was shown to selectively accumulate in pancreatic islets while being excluded from other tissues in the abdominal area.[46] In turn, a scaffold targeted towards the Melanocortin 1 receptor (MC1R) and CCK-2R was injected into mice with induced tumors that expressed one or both

receptors, and was selectively found in the dual-expressing tumor (see Fig. 1.5, right flank). This scaffold was retained in the tumor for more than 24 h, thus affirming its potential for imaging and delivery purposes.[38]

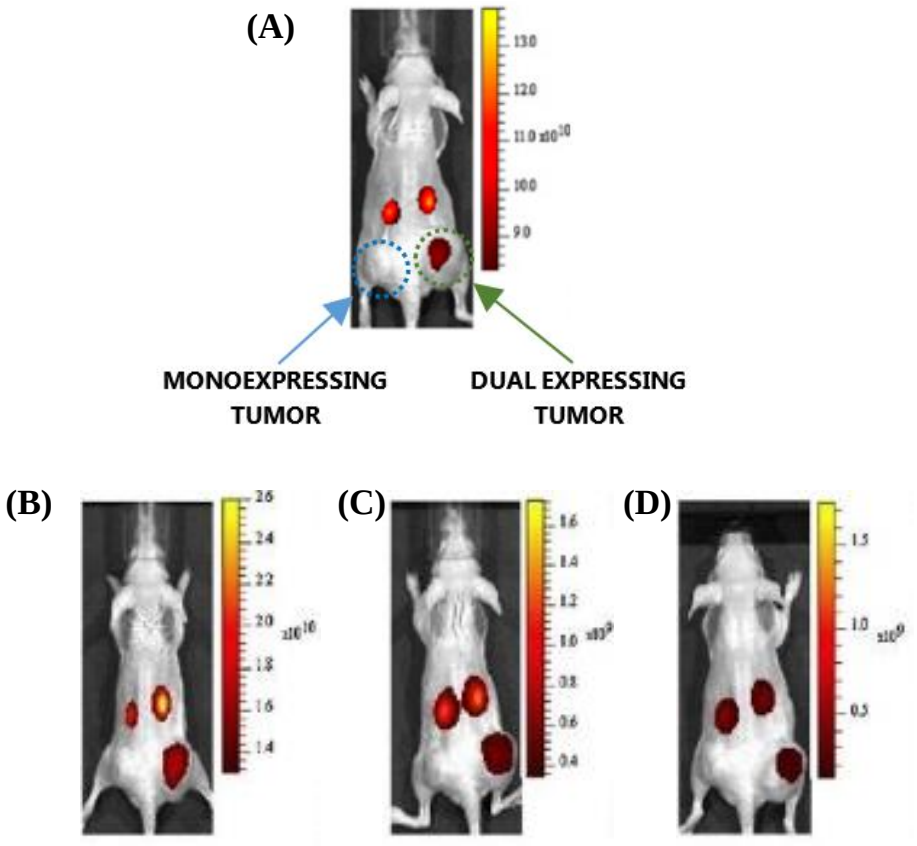


Figure 1.5. *In vivo* imaging of fluorescently labeled heteromultivalent ligand shows selective accumulation in a dual expressing MC1R/CCK-2R tumor compared to a MC1R-expressing tumor at 30 min (**A**), 1 h (**B**), 4 h (**C**), and 10 h (**D**) after ligand injection. Figure was adapted from Xu *et al.* Heterobivalent ligands target cell-surface receptor combinations in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (52), 21295-21300.

Another application for heteromultivalent ligands pertains to the activation or suppression of biological responses. A series of heterobivalent ligands with affinity towards the D_2R/μ-OR

combination demonstrated high potency and efficacy for the activation of the two individual signaling pathways involving these receptors, which was confirmed by the phosphorylation and/or recruitment of downstream proteins.[42] Interestingly, the affinities of heterobivalent and monovalent ligands were comparable, with no significant differences between dual- and mono-expressing cells. In less successful cases, heterobivalent ligands only remained functional towards one of the cognate receptors, presumably due to the covalent modifications performed on the individual ligands, which are required to incorporate them into the scaffolds.[44, 47]

1.5. References

1. Martinez, J. D.; Parker, M. T.; Fultz, K. E.; Ignatenko, N. A.; Gerner, E. W., Molecular Biology of Cancer. In *Burger's Medicinal Chemistry and Drug Discovery*, 6th ed.; Abraham, D. J., Ed. John Wiley & Sons, Inc.: Hoboken, NJ, 2003; Vol. 5, pp 1-50.

2. Bertram, J. S., The molecular biology of cancer. *Mol. Aspects Med.* **2000,** *21* (6), 167-223.

3. Hanahan, D.; Weinberg, R. A., Hallmarks of cancer: The next generation. *Cell* **2011,** *144* (5), 646-674.

4. Hanahan, D.; Weinberg, R. A., The hallmarks of cancer. *Cell* **2000,** *100* (1), 57-70.

5. American Cancer Society *Global cancer facts & figures*; American Cancer Society: Atlanta, GA, 2015; pp 1-64.

6. American Cancer Society *Cancer facts & figures 2018*; American Cancer Society: Atlanta, GA, 2018; pp 1-76.

7. Mariotto, A. B.; Yabroff, K. R.; Shao, Y.; Feuer, E. J.; Brown, M. L., Projections of the cost of cancer care in the United States: 2010–2020. *JNCI, J. Natl. Cancer Inst.* **2011,** *103* (2), 117-128.

8. Paul, H.; Isabelle, S.; Linda, S., Measuring the societal burden of cancer: The cost of lost productivity due to premature cancer-related mortality in Europe. *Int. J. Cancer* **2015,** *136* (4), E136-E145.

9. Rosenberg, B.; Van Camp, L.; Krigas, T., Inhibition of cell division in *Escherichia coli* by electrolysis products from a platinum electrode. *Nature (London)* **1965,** *205*, 698-699.

10. Rosenberg, B.; Vancamp, L.; Trosko, J. E.; Mansour, V. H., Platinum compounds: A new class of potent antitumour agents. *Nature (London)* **1969,** *222*, 385-386.

11. Kelland, L., The resurgence of platinum-based cancer chemotherapy. *Nat. Rev. Cancer* **2007,** *7* (8), 573-584.

12. Dasari, S.; Tchounwou, P. B., Cisplatin in cancer therapy: Molecular mechanisms of action. *Eur. J. Pharmacol.* **2014,** *740*, 364-378.

13. Gibson, N. W.; Hickman, J. A.; Erickson, L. C., DNA cross-linking and cytotoxicity in normal and transformed human cells treated *in vitro* with 8-Carbamoyl-3-(2-chloroethyl)imidazo[5,1-*d*-1,2,3,5-tetrazin-4(3*H*)-one. *Cancer Res.* **1984,** *44* (5), 1772-1775.

14. Ramirez, M. L.; Shepherd, C. R.; Pinsky, S. D.; McGhee, K. G.; Abel, G.; Harrap, K. R., Chlorambucil-induced changes in chromatin methylation and their modulation by sodium butyrate. *Biochem. Biophys. Res. Commun.* **1982,** *108* (4), 1734-1742.

15. Miller, R. P.; Tadagavadi, R. K.; Ramesh, G.; Reeves, W. B., Mechanisms of cisplatin nephrotoxicity. *Toxins* **2010,** *2* (11), 2490-2518.

16. Rabik, C. A.; Dolan, M. E., Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat. Rev.* **2007,** *33* (1), 9-23.

17. Teimouri, F.; Nikfar, S.; Abdollahi, M., Efficacy and side effects of dacarbazine in comparison with temozolomide in the treatment of malignant melanoma: A meta-analysis consisting of 1314 patients. *Melanoma Res.* **2013,** *23* (5), 381-389.

18. Miller, J. B., Photodynamic therapy: The sensitization of cancer cells to light. *J. Chem. Ed.* **1999,** *76* (5), 592-594.

19. Detty, M. R.; Gibson, S. L.; Wagner, S. J., Current clinical and preclinical photosensitizers for use in photodynamic therapy. *J. Med. Chem.* **2004,** *47* (16), 3897-3915.

20. Qin, Z.; Bischof, J. C., Thermophysical and biological responses of gold nanoparticle laser heating. *Chem. Soc. Rev.* **2012,** *41* (3), 1191-1217.

21. Agostinis, P.; Berg, K.; Cengel, K. A.; Foster, T. H.; Girotti, A. W.; Gollnick, S. O.; Hahn, S. M.; Hamblin, M. R.; Juzeniene, A.; Kessel, D.; Korbelik, M.; Moan, J.; Mroz, P.; Nowis, D.; Piette, J.; Wilson, B. C.; Golab, J., Photodynamic therapy of cancer: An update. *CA – Cancer J. Clin.* **2011,** *61* (4), 250-281.

22. Wang, J.; Zigler, D. F.; Hurst, N.; Othee, H.; Winkel, B. S. J.; Brewer, K. J., A new, bioactive structural motif: Visible light induced DNA photobinding and oxygen independent photocleavage by Ru^{II}, Rh^{III} bimetallics. *J. Inorg. Biochem.* **2012,** *116*, 135-139.

23. Redmond, R. W.; Kochevar, I. E., Spatially resolved cellular responses to singlet oxygen. *Photochem. Photobiol.* **2006,** *82* (5), 1178-1186.

24. DeRosa, M. C.; Crutchley, R. J., Photosensitized singlet oxygen and its applications. *Coord. Chem. Rev.* **2002,** *233–234*, 351-371.

25. Wang, J.; Brewer, K. J.; Winkel, B. S. J.; Tissue, B. M.; Madsen, L. A. The design, synthesis and study of mixed-metal Ru,Rh and Os,Rh complexes with biologically relevant reactivity. PhD dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 2012.

26. Henderson, B. W.; Busch, T. M.; Vaughan, L. A.; Frawley, N. P.; Babich, D.; Sosa, T. A.; Zollo, J. D.; Dee, A. S.; Cooper, M. T.; Bellnier, D. A.; Greco, W. R.; Oseroff, A. R., Photofrin photodynamic therapy can significantly deplete or preserve oxygenation in human basal cell carcinomas during treatment, depending on fluence rate. *Cancer Res.* **2000,** *60* (3), 525-529.

27. Sitnik, T. M.; Hampton, J. A.; Henderson, B. W., Reduction of tumour oxygenation during and after photodynamic therapy in vivo: Effects of fluence rate. *Br. J. Cancer* **1998,** *77* (9), 1386-1394.

28. Knoll, J. D.; Turro, C., Control and utilization of ruthenium and rhodium metal complex excited states for photoactivated cancer therapy. *Coord. Chem. Rev.* **2015,** *282-283*, 110-126.

29. Fan, W.; Huang, P.; Chen, X., Overcoming the Achilles' heel of photodynamic therapy. *Chem. Soc. Rev.* **2016,** *45* (23), 6488-6519.

30. Cheng, L.; Wang, C.; Liu, Z., Upconversion nanoparticles and their composite nanostructures for biomedical imaging and cancer therapy. *Nanoscale* **2013,** *5* (1), 23-37.

31. Mammen, M.; Choi, S. K.; Whitesides, G. M., Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem. Int. Ed.* **1998,** *37* (20), 2754-2794.

32. Nakamoto, K.; Tsuboi, M.; Straham, G. D., *Drug-DNA interactions: Structures and spectra*. 1st ed.; John Wiley & Sons: Hoboken, NJ, 2008.

33. Kiessling, L. L.; Lamanna, A. C. In *Multivalency in biological systems*, Chemical Probes in Biology, Dordrecht, 2003; Schneider, M. P., Ed. Springer Netherlands: Dordrecht, 2003; pp 345-357.

34. Evans, S. V.; MacKenzie, C. R., Characterization of protein–glycolipid recognition at the membrane bilayer. *J. Mol. Recognit.* **1999,** *12* (3), 155-168.

35. Lindhorst, T. K., Oligosaccharides and glycoconjugates in recognition processes. In *Carbohydrate-Modifying Biocatalysts*, Pan Stanford Publishing Pte. Ltd.: New York, NY, 2012; pp 119-182.

36. Josan, J. S.; Handl, H. L.; Sankaranarayanan, R.; Xu, L.; Lynch, R. M.; Vagner, J.; Mash, E. A.; Hruby, V. J.; Gillies, R. J., Cell-specific targeting by heterobivalent ligands. *Bioconjugate Chem.* **2011,** *22* (7), 1270-1278.

37. Vagner, J.; Xu, L.; Handl, H. L.; Josan, J. S.; Morse, D. L.; Mash, E. A.; Gillies, R. J.; Hruby, V. J., Heterobivalent ligands crosslink multiple cell-surface receptors: The human Melanocortin-4 and δ-Opioid Receptors. *Angew. Chem. Int. Ed.* **2008,** *47* (9), 1685-1688.

38. Xu, L.; Josan, J. S.; Vagner, J.; Caplan, M. R.; Hruby, V. J.; Mash, E. A.; Lynch, R. M.; Morse, D. L.; Gillies, R. J., Heterobivalent ligands target cell-surface receptor combinations in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (52), 21295-21300.

39. Danhier, F.; Feron, O.; Préat, V., To exploit the tumor microenvironment: Passive and active tumor targeting of nanocarriers for anti-cancer drug delivery. *J. Controlled Release* **2010,** *148* (2), 135-146.

40. Low, P. S.; Henne, W. A.; Doorneweerd, D. D., Discovery and development of folic-acid-based receptor targeting for imaging and therapy of cancer and inflammatory diseases. *Acc. Chem. Res.* **2008,** *41* (1), 120-129.

41. Sridhar, S. S.; Hedley, D.; Siu, L. L., Raf kinase as a target for anticancer therapeutics. *Mol. Cancer Ther.* **2005,** *4* (4), 677-685.

42. Qian, M.; Vasudevan, L.; Huysentruyt, J.; Risseeuw Martijn, D. P.; Stove, C.; Vanderheyden Patrick, M. L.; Van Craenenbroeck, K.; Van Calenbergh, S., Design, synthesis, and biological

evaluation of bivalent ligands targeting Dopamine D_2-Like Receptors and the μ-Opioid Receptor. *ChemMedChem* **2018,** *13* (9), 944-956.

43. Hiller, C.; Kühhorn, J.; Gmeiner, P., Class A G-protein-coupled receptor (GPCR) dimers and bivalent ligands. *J. Med. Chem.* **2013,** *56* (17), 6542-6559.

44. Soriano, A.; Ventura, R.; Molero, A.; Hoen, R.; Casadó, V.; Cortés, A.; Fanelli, F.; Albericio, F.; Lluís, C.; Franco, R.; Royo, M., Adenosine A_{2A} Receptor-antagonist/Dopamine D_2 Receptor-agonist bivalent ligands as pharmacological tools to detect A_{2A}-D_2 Receptor heteromers. *J. Med. Chem.* **2009,** *52* (18), 5590-5602.

45. Xu, L.; Vagner, J.; Josan, J.; Lynch, R. M.; Morse, D. L.; Baggett, B.; Han, H.; Mash, E. A.; Hruby, V. J.; Gillies, R. J., Enhanced targeting with heterobivalent ligands. *Mol. Cancer Ther.* **2009,** *8* (8), 2356-2365.

46. Steyn, L. V.; Ananthakrishnan, K.; Anderson, M. J.; Patek, R.; Kelly, A.; Vagner, J.; Lynch, R. M.; Limesand, S. W., A synthetic heterobivalent ligand composed of Glucagon-Like Peptide 1 and Yohimbine specifically targets β cells within the pancreas. *Mol. Imaging Biol.* **2015,** *17* (4), 461-470.

47. Hart, N. J.; Chung, W. J.; Weber, C.; Ananthakrishnan, K.; Anderson, M.; Patek, R.; Zhang, Z.; Limesand, S. W.; Vagner, J.; Lynch, R. M., Hetero-bivalent GLP-1/Glibenclamide for targeting pancreatic β-cells. *ChemBioChem* **2014,** *15* (1), 135-145.

2. MECHANISTIC INVESTIGATION INTO DNA MODIFICATION BY A Ru^{II},Rh^{III} BIMETALLIC COMPLEX

José Á. Rodríguez Corrales, Jing Wang, Brenda S.J. Winkel, Karen J. Brewer

This chapter has been adapted from a manuscript accepted for publication in the journal *ChemBioChem* under the "Full paper" format.

2.1. Abstract

Despite significant progress in the treatment of cancer, there remains an urgent need for more effective therapies that also have less impact on patient wellbeing. Photodynamic therapy employs targeted light activation of a photosensitizer in selected tissues, thereby reducing off-target toxicity. Our group previously reported a Ru^{II},Rh^{III} bimetallic architecture that displays multifunctional covalent photomodification of DNA in the therapeutic window through an oxygen-independent manner, features essential for treating deep and hypoxic tumors. Herein, we explored the mechanism by which a new analog, $[(phen)_2Ru(dpp)Rh(phen)Cl_2]^{3+}$, or **Ru(II)-Rh(III)**, interacts with DNA. We established that **Ru(II)-Rh(III)** exhibits "light switch" behavior in the presence of DNA, undergoing strong electrostatic interactions that may involve groove binding. Furthermore, these non-covalent interactions play a major role in covalent photobinding and photocleavage of DNA, which occur via an oxygen-independent mechanism. Polymerase chain reaction (PCR) revealed that covalent modification of DNA by **Ru(II)-Rh(III)**, especially photobinding, is critical to inhibit amplification, suggesting that the complex could exert a toxic activity by interfering with DNA replication in cells. This new structural motif, with

phenanthroline at all three terminal ligand positions, displays a number of properties that are promising for the continued refinement of photodynamic therapy strategies.

2.2. Introduction

Cancer is a complex group of diseases that are the second leading cause of death worldwide.[1] In the United States alone, it is predicted that more than 1.7 million cases of cancer will be diagnosed during 2018, and that 600,000 will die from the disease.[2] Although many antineoplastic agents have been developed since the first anticancer drug was introduced 70 years ago, there remains an urgent need to develop new, more effective treatments.

One of the most widely-used drugs in cancer treatment is *cis*-diamminedichloroplatinum (II), or cisplatin. This monometallic compound is effective against a variety of cancers, including carcinomas, lymphomas, and sarcomas. Cisplatin promotes apoptosis after generating covalent adducts with DNA, either though the formation of inter- or intra-strand crosslinks, or mono-adducts.[3] Efforts to develop platinum-based analogs with reduced side effects have resulted in only one compound to date that has clear advantages over cisplatin, i.e. carboplatin. Although these drugs effectively diminish the tumor burden, lack of selectivity, other than preferentially affecting rapidly-dividing cells, results in significant off-target effects that range from nausea to nephrotoxicity, neurotoxicity, and neutropenia.[4-5]

Photodynamic therapy (PDT) is a therapeutic approach in which drugs known as photosensitizers are selectively activated with light, thus minimizing toxicity to non-irradiated cells and tissues. Photofrin® and Levulan®, the two PDT drugs approved to date by the FDA for cancer treatment, generate an excited and highly oxidizing form of O_2, known as singlet oxygen ($^1O_2$), that can rapidly result in cell death.[6-7] However, the efficacy of these drugs is negatively

impacted by the hypoxic nature of many tumors, the result of an imbalance between O_2 consumption and supply.[8-9]

For this reason, there is significant interest in photosensitizers that can generate toxicity through oxygen-independent mechanisms. Ruthenium complexes accomplish this goal through two pathways: ligand photolabilization (e.g., ejection of toxic ligands that are inactive while bound to Ru, or covalently binding to biological targets through vacant coordination sites) and generation of a highly oxidizing excited state for direct electron transfer with biomolecules.[10-11] Rhodium complexes have displayed photoinduced DNA binding or cleavage in an anaerobic environment under UV or violet light excitation.[12-13] Interestingly, $[Rh(phen)_2(phi)]^{3+}$ (phen=1,10-phenanthroline, phi=9,10-phenanthrenequinone diimine) preferentially cleaves 5'-pyrimidine-pyridimidine-purine-3' sequences, while $[Rh(phi)_2(bpy)]^{3+}$ (bpy=2,2'-bipyridine) is not selective. This difference in behavior is attributed to the specific size and shape of the complexes which alter the non-covalent interactions established with DNA prior to scission.[13]

Recently, our group reported the synthesis and characterization of multimetallic supramolecular complexes with oxygen-independent multifunctional DNA modification capabilities. In particular, DNA binding and cleavage has been reported for a series of complexes with the general structure $[(bpy)_2Ru(BL)Rh(phen)Cl_2]^{3+}$ (BL= bridging ligand: dpp=2,3-bis(2-pyridyl)pyrazine or bpm=2,2′-bipyrimidine).[14] These supramolecules display intense Ru-based metal-to-ligand charge transfer (MLCT) transitions in the visible range, as well as generation of a Ru→Rh metal-to-metal charge transfer (^{3}MMCT) state. Unlike other rhodium complexes, these molecules are activated with low energy light closer to the phototherapeutic window due to the Ru-based absorption, and display dual covalent binding and oxygen-independent cleavage of DNA, with minimum reactivity under dark conditions.

Herein we report a new analog from the Ru^{II},Rh^{III} series that bears 1,10-phenanthroline ligands on the Ru center with the aim of enhancing reactivity with DNA. We establish that $[(phen)_2Ru(dpp)Rh(phen)Cl_2]^{3+}$, henceforth referred to as **Ru(II)-Rh(III),** displays a strong electrostatic interaction with DNA that appears to involve groove binding, as evidenced by competitive binding experiments. Furthermore, we show that these non-covalent interactions play a critical role in the covalent photobinding and photocleavage of DNA, the latter proceeding through a collision-based oxygen-independent mechanism. Finally, we show that the formation of covalent adducts is essential for the inhibition of DNA amplification, as assessed by polymerase chain reaction (PCR), with electrostatic interactions again playing a major role in promoting the activity of the complex.

2.3. Experimental section

2.3.1. Materials

Tris, agarose, boric acid, Bu_4NCl, calf-thymus DNA, and Hoechst 33342 were procured from Sigma Aldrich. $RuCl_3 \cdot 3H_2O$, 1,10-phenanthroline monohydrate, dpp, Bu_4NPF_6, and NH_4PF_6 were procured from Alfa Aesar. pUC18 and pUC19 plasmid DNA were purchased from Bayou Biolabs. λ DNA/*Hin*dIII markers and 1 kb DNA molecular weight ladder were obtained from Promega. DNA primers were procured from Integrated DNA Technologies. Biomix™ was purchased from Bioline. $[(phen)_2Ru(dpp)](PF_6)_2$[15] and $[RhCl_3(phen)(DMF)]$[16] (DMF=N,N-dimethylformamide) were prepared according to previously reported methods.

2.3.2. Synthesis of Ru(II)-Rh(III)

$[(phen)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$ was synthesized through a building block approach (see Scheme S2.1) based on previously reported methods for Ru^{II},Rh^{III} bimetallics modified as follows.[17-18] $[RhCl_3(phen)(DMF)]$ (60.9 mg, 0.132 mmol) was dissolved in warm ethanol:water 2:1 (10 mL) and a solution of $[(phen)_2Ru(dpp)](PF_6)_2$ (110 mg, 0.112 mmol) in the same solvent mixture (10 mL) was added slowly. The resulting solution was refluxed for 2 h and then added dropwise into a concentrated solution of NH_4PF_6 (~50 mL) to induce precipitation of the complex. Solid was collected by vacuum filtration, redissolved in ethanol:acetonitrile 2:1 (~2-4 mL), syringe filtered and purified by size exclusion chromatography using a Sephadex® LH-20 column. Elution was performed using ethanol:acetonitrile 2:1 and the product was collected as the leading dark red band. Fractions were examined by emission spectroscopy ($\lambda_{excitation}$ = 450 nm, 550 nm < $\lambda_{emission}$ < 750 nm) to detect the highly emissive $[(phen)_2Ru(dpp)](PF_6)_2$. The fractions free of the latter compound were combined and the solvent was removed by vacuum evaporation. The resulting solid was redissolved in a small volume of acetonitrile and flash precipitated into 500 mL of ether. The final solid was collected by vacuum filtration, analyzed by electronic absorption and emission spectrometry and electrochemistry as described below. To metasthesize the product to its chloride salt, $[(phen)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$ was redissolved in a small volume of dry acetone and added into a saturated solution of Bu_4NCl, thus affording $[(phen)_2Ru(dpp)RhCl_2(phen)]Cl_3$ as a precipitate. The resulting solid was collected by vacuum filtration and rinsed extensively with ethyl ether:dry acetone 9:1. The chloride salt of **Ru(II)-Rh(III)** was used in all experiments except where noted otherwise.

2.3.3. Electronic absorption spectroscopy

Spectra were collected using an Agilent 8453 diode array spectrophotometer with 2 nm resolution in a 1 cm quartz cell at room temperature.

2.3.4. Steady state emission spectroscopy

Spectra were collected with a modified QuantaMaster Model QM-200-45E emission spectrophotometer from Photon Technology, Inc. (PTI). A water-cooled 150 W xenon arc lamp was used as excitation source and emission was detected using a thermoelectrically-cooled Hamamatsu 1527 photomultiplier tube (PMT) in photon counting mode with 0.25 nm resolution, located at a 90° angle of the excitation source. All spectra were collected at least three times and averaged to minimize background noise.

2.3.5. Excited state lifetime

PF_6^- salts of the metal complexes were dissolved in spectrophotometric grade acetonitrile, deoxygenated through argon purge. Sample excitation was performed using a PL 2300 nitrogen laser from PTI, coupled to a PL 201 dye laser with a Coumarin 500 laser dye. Detection was conducted using a single photon counting direct output wired Hamamatsu R928 photomultiplier tube grounded to a resistor. Voltage from the photomultiplier tube was measured with a LeCroy 9361 Dual 300 MHz oscilloscope. In turn, air-equilibrated aqueous samples of Hoechst 33342 were prepared as described in the competitive binding assays section. Lifetime was measured in the QuantaMaster instrument described above using a 350 nm LED and a Becker & Hickl GmbH PMH-100 PMT detector at 465 nm. Reported lifetimes correspond to the average of three different samples measured over 300 pulses each.

2.3.6. Emission quantum yield

Samples were prepared as described for the lifetime measurements and their emission spectra corrected for photomultiplier tube response. Emission quantum yield (Φ^{em}) was calculated using $[Os(bpy)_3](PF_6)_2$ as standard ($\Phi^{em} = 4.6\times10^{-3}$)[19] according to Eq. 2.1,

$$\Phi^{em}_{sample} = \Phi^{em}_{standard} \times \frac{Area_{sample}}{Area_{standard}} \times \frac{A_{standard}}{A_{sample}} \quad \text{(Eq. 2.1)}$$

where A and Area correspond to the absorbance at the excitation wavelength and the area under the emission spectra, respectively, of the sample or standard.

2.3.7. Electrochemistry

Cyclic voltammograms were determined using a Bioanalytical Systems (BAS) Epsilon electrochemical analyzer with a three-electrode, single-compartment cell. Potentials were measured using a Ag/AgCl reference electrode, a platinum disk working electrode, and a platinum wire counter electrode. The supporting electrolyte was a 0.1 M solution of Bu_4NPF_6 in deoxygenated and dry acetonitrile at room temperature. The cyclic voltammograms were obtained at a scan rate of 100 mV/s. Argon was used to purge all solutions prior to measurement and to blanket the cell during data collection.

2.3.8. Non-covalent binding constant

Concentrations of CT-DNA solutions were calculated using the reported extinction coefficient at 260 nm (6,600 $M^{-1}\,cm^{-1}$ per base).[20] The ratio of absorbance at 260 and 280 nm was always greater than 1.8, indicating that the CT-DNA was substantially free of protein.[21] A 30 μM aqueous solution of **Ru(II)-Rh(III)** was prepared in 5 mM Tris buffer (pH = 7.0) and 10, 50 or 500 mM NaCl, and titrated with increasing concentrations of CT-DNA from 10^{-5} to 10^{-3} M. After

each addition, the solution was incubated at room temperature for 5 min protected from light to allow equilibration. Binding constants were determined using Eq. 2.2,[22-23]

$$\frac{C_N - C_T}{|\varepsilon_a - \varepsilon_f|} = \frac{C_N - C_T}{|\varepsilon_b - \varepsilon_f|} + \frac{1}{K_b|\varepsilon_b - \varepsilon_f|} \quad \text{(Eq. 2.2)}$$

where C_N is the DNA concentration, C_T is the concentration of complex, K_b is the apparent binding constant, ε_a is the apparent extinction coefficient (A/C_T), and ε_b and ε_f are the extinction coefficients for the bound and free forms, respectively.

2.3.9. Reverse salt titration

The titration was performed based on a previously reported method.[24-25] A solution (2 mL) of **Ru(II)-Rh(III)** (30 μM total concentration, C_t) with DNA (234 μM in N-bp, C_N) in NaCl (50 mM) and Tris buffer (5 mM, pH 7.0) was prepared and its emission spectrum was measured. This solution was titrated with NaCl (5 M) up to a final concentration of 100 mM, the emission spectrum was measured after each addition, and the area under the curve (F) was integrated. The integrated emission for the free compound (F_0) was calculated from a **Ru(II)-Rh(III)** (30 μM) solution with Tris buffer (5 mM, pH=7.0) and NaCl (50 mM). The molar integrated emission for the bound compound (F_b) was obtained from a titration with DNA until the emission spectra reached a constant value. The concentration of bound **Ru(II)-Rh(III)** (C_b) was calculated using Eq. 2.3,

$$C_b = \frac{F - F_b C_t}{F_b - F_0} \quad \text{(Eq. 2.3)}$$

The binding constant (K_b) at each [Na^+] was calculated using Eq. 2.4.

$$K_b = \frac{C_b}{(C_t - C_b)(C_N - C_b)} \quad \text{(Eq. 2.4)}$$

The parameters SK and Z were determined from Eq. 2.5,

$$SK = -Z\psi = \frac{\delta \log K_b}{\delta \log [Na^+]} \quad \text{(Eq. 2.5)}$$

where ψ=0.88 for DNA,[25] yielding SK=−2.2 and Z=2.5.

The "total" or "observed" change in Gibbs free energy (ΔG_{obs} at 50 mM Na^+) and its electrostatic contribution (ΔG_{pe}) were calculated Eq. 2.6 and 2.7, respectively,

$$\Delta G_{obs} = -RT \ln(K_b) \quad \text{(Eq. 2.6)}$$

$$\Delta G_{pe} = SK\ RT \ln[Na^+] \quad \text{(Eq. 2.7)}$$

where R=1.987 cal mol^{-1} K^{-1} and T is the temperature in K.

2.3.10. Competitive binding assays

H stock solution was prepared in water and quantified by dilution into methanol and use of an extinction coefficient of 4.7×10^4 $M^{-1}cm^{-1}$ at 343 nm as reported by the supplier. Solutions were prepared in 10 mM phosphate buffer (pH 7.4) using the substances specified in each figure at 22 μM for **H** or **Ru(II)-Rh(III**), and 110 μM for CT-DNA in order to generate the 1:5 binder:DNA bp ratio used in DNA gel shift assays. The effect of ionic strength in emission was evaluated under the same buffer, CT-DNA, and complex concentrations and using a 1 M NaCl stock as the source of any additional Na^+. All solutions were incubated at room temperature and protected from light for at least 5 min prior to measuring their emission spectra. The order of mixing (e.g., adding **Ru(II)-Rh(III)** before **H** or vice versa) did not affect the final emission profile.

2.3.11. Estimation of Förster radius

Förster radius was calculated using the emission and absorption spectra of the DNA-bound forms of **H** and **Ru(II)-Rh(III)**, respectively (see Fig. S2.3). First, the overlap integral, $J(\lambda)$, was determined using Eq. 2.8,[26]

$$J(\lambda) = \int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \qquad \text{(Eq. 2.8)}$$

where $F_D(\lambda)$ is the emission intensity of the donor at each wavelength (λ), corrected such that the area of the emission spectrum is unity, and $\varepsilon_A(\lambda)$ is the extinction coefficient of the acceptor in units of $M^{-1}cm^{-1}$. The Förster radius (in Å) was then calculated using Eq. 2.9,[26]

$$R_0 = 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{1/6} \qquad \text{(Eq. 2.9)}$$

where κ is the orientation factor (assumed to be 2/3), n is the refractive index (assumed to be 1.4 for biomolecules), and Q_D is the quantum yield of the donor in the absence of acceptor (reported as 0.38 for Hoechst 33342 bound to DNA).[27]

2.3.12. DNA gel shift assay

pUC18 or pUC19 DNA-metal complex solutions were prepared in a 5:1 BP (base pairs):MC ratio in NaH_2PO_4/Na_2HPO_4 phosphate buffer (pH = 7.4, concentration is specified in each figure or experiment) with 11.7 μM metal complex to ensure an absorbance of at least 0.2 at the excitation wavelength (455 nm). Solutions were deoxygenated by bubbling argon for 15 min and then kept closed to maintain an argon blanket. Aliquots of these solutions were taken, covered with aluminum foil and stored in closed cabinets to serve as "dark" controls. Photolysis of samples was performed at 455 nm using a previously-described LED array, and aliquots were collected at different irradiation times.[28] Gel electrophoresis was conducted in a model B1A stage Owl

Separation Systems using a 35 mL 0.8% w/w agarose gel immersed in 300 mL of TB buffer (90 mM tris base, 90 mM boric acid). DNA molecular weight ladder λ DNA-*Hin*dIII was diluted according to the manufacturer's instructions; untreated DNA was diluted to match the concentration of the treated samples. A 5 μL aliquot of each sample was mixed with 5 μL of 10 mM buffer and 2 μL of loading dye and then loaded into individual wells. Electrophoresis was at 100 V for 1.5 h, and gels were stained in 0.5 μg/mL ethidium bromide for 0.5 h, destained in deionized water for 0.5 h, and visualized on a Fisher Biotech UV-transilluminator. Images were captured with an Olympus SP-320 camera fitted with an ethidium bromide filter.

2.3.13. ROS scavenger assay

Solutions of pDNA and **Ru(II)-Rh(III)** were prepared in 10 mM phosphate buffer (pH 7.4) using the concentrations described for the DNA gel shift assay. Solutions were spiked with 1 mM ROS scavengers prior to photolysis. Sodium azide was utilized as singlet oxygen scavenger,[29-30] while potassium iodide[31-32] and sodium benzoate[33-35] were used as hydroxyl radical scavengers. A separate solution spiked with sodium chloride (1 mM) served as an ionic strength control.

2.3.14. DNA precipitation assay

This assay was adapted from a previously reported method.[36] A solution of 1 mM CT-DNA with 0.2 mM **Ru(II)-Rh(III)** was prepared in 10 mM phosphate buffer (pH 7.4) through the slow addition of a 0.4 mM **Ru(II)-Rh(III)** stock into a 2 mM CT-DNA stock. A separate solution was prepared similarly except NaCl was added to a final concentration of 90 mM. For each time period of irradiation at 455 nm or incubation in darkness, 250 μL aliquots were collected and 10 μL of 5 M NaCl and 800 μL of 95% ethanol added immediately. The samples were incubated in the freezer

(−20 °C) overnight to allow complete precipitation of DNA. After centrifugation at 13000 rpm (14,000 × g) in a VWR Galaxy 14D centrifuge for 15 min, electronic absorption spectra of the supernatants were measured and A/A_0 at 508 nm was calculated. In some cases, precipitated DNA was redissolved in 10 mM phosphate buffer (pH 7.4) and its absorbance spectrum was measured.

2.3.15. PCR amplification

To prevent sample contamination, the work area was cleaned with 75% ethanol and dedicated reaction tubes and pipette tips were used. Photolyzed samples and dark controls were prepared as described for the DNA gel assay using pUC18, which only differs from pUC19 with regard to orientation of the multiple cloning site, 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer, and 90 mM NaCl for the high ionic strength sample only. After photolysis or incubation in the dark, samples were diluted to 10^{-3} μg/μL using nuclease-free water. A 1 μL aliquot of the resulting solutions was mixed with 2 μL of 5 μM forward primer, 2 μL of 5 μM reverse primer and 5 μL of Biomix™ 2X reaction mix. Primer sequences are given in Table S2. Positive and negative controls consisted of solutions containing diluted pUC18 DNA and nuclease-free water, respectively, but no complex. Samples were amplified in a MJ Research PTC-100TM Thermal Controller using 35 cycles of 94°C for 30 s, 58 °C for 30 s, and 72 °C for 2 min. Afterwards, samples were diluted 2-fold in 10 mM phosphate buffer (pH 7.4), 4 μL of dye was added, and 12 μL of the resulting solutions was loaded in each well of a 0.8% w/w agarose gel. A 1 kb molecular weight ladder was mixed with loading dye and used as recommended by the supplier. Electrophoresis was performed as described above, except voltage was applied for 1 h.

2.4. Results and discussion

Ru(II)-Rh(III) was synthesized through a modified version of a previously-reported building block scheme (Scheme S2.1).[18] Characterization through electronic absorption and emission spectroscopy revealed that this complex and the previously-described bpy analog, $[(bpy)_2Ru(dpp)RhCl_2(phen)]^{3+}$,[14] share several common features, including strong MLCT transitions in the visible spectrum that tail into the phototherapeutic window,[37] and a red-shifted ^{3}MLCT emission (Fig. S2.1A and Table S2.1). While the quantum yield of these two complexes was within experimental error, at $2.3x10^{-4}$ and $2.4x10^{-4}$ for **Ru(II)-Rh(III)** and the bpy analog, respectively, the lifetime of the ^{3}MLCT state of **Ru(II)-Rh(III)** was approximately 2-fold higher, at 64 ns at room temperature in argon-purged acetonitrile versus 30 ns for the bpy analog.[18] Furthermore, cyclic voltammetry confirmed the presence of a Ru-based HOMO and a Rh-based LUMO for the new complex (Fig. S2.1B and Table S2.1). The spectroscopic and electrochemical characterization of **Ru(II)-Rh(III)** also showed population of a low-lying Ru→Rh ^{3}MMCT excited state, which has been implicated as the reactive state that leads to oxygen-independent DNA photocleavage by related bi- and trimetallic systems.[14, 38]

Non-covalent interaction of **Ru(II)-Rh(III)** with calf thymus (CT) DNA was characterized through absorption spectroscopy using a previously-reported method.[22] Briefly, **Ru(II)-Rh(III)** was incubated with different concentrations of DNA in the absence of light and the absorbance of the resulting solutions was recorded. The difference between the extinction coefficients of DNA-bound and free forms of the complex was used to determine the apparent binding constant (K_b) for the formation of the non-covalent adduct. **Ru(II)-Rh(III)** was found to bind to CT-DNA with a binding constant of $3.7(\pm0.3)\times10^4$ M^{-1} in the presence of 50 mM Na^+. This is one order of

magnitude greater than previously reported for the monometallic $[Ru(phen)_3]^{2+}$ (5.5×10^3 M^{-1}),[39] which we attribute to the higher positive charge displayed by **Ru(II)-Rh(III)**.

Moreover, the K_b of **Ru(II)-Rh(III)** was found to be highly dependent on ionic strength. By decreasing $[Na^+]$ from 50 mM to 10 mM, the K_b increased approximately one order of magnitude, to $2.2(\pm0.5)\times10^5$ M^{-1}, whereas no binding was detected at high ionic strength (i.e., 500 mM Na^+). This behavior is consistent with electrostatic interaction of **Ru(II)-Rh(III)** with DNA. Similarly, a series of Ru(II) bimetallic systems, $[(phen)_2Ru(Mebpy)\text{-}(CH_2)_n\text{-}(bpyMe)Ru(phen)_2]^{4+}$ (Mebpy=4-methyl-2,2'-bipyridyl-4'-, and n=5, 7 or 10) that bind electrostatically to DNA displayed binding constants sensitive to ionic strength. The K_b of these compounds, $2\text{-}4\times10^6$ M^{-1} at 10 mM Na^+, is one order of magnitude higher than **Ru(II)-Rh(III)**, again consistent with their higher charge.[40]

The interaction of **Ru(II)-Rh(III)** with DNA was further characterized by emission spectroscopy. As previously reported for other ruthenium-based complexes,[24, 41] in the presence of CT-DNA **Ru(II)-Rh(III)** exhibited both a hypsochromic shift (from 790 nm in aqueous buffer to 770 nm in buffered DNA solution) and a 6-fold increase in emission intensity (Fig. 2.1A; No DNA vs. 10 mM Na^+ or red vs. black traces, respectively). Although "light-switch" behavior has been observed in several other ruthenium complexes,[24, 41-42] this is only the second such report for a mixed-metal compound.[43] The enhanced emission in the presence of DNA has been attributed to several factors, including solvatochromic effects and increased motion restraint, which hinder relaxation of the excited state through non-radiative pathways.[24, 44]

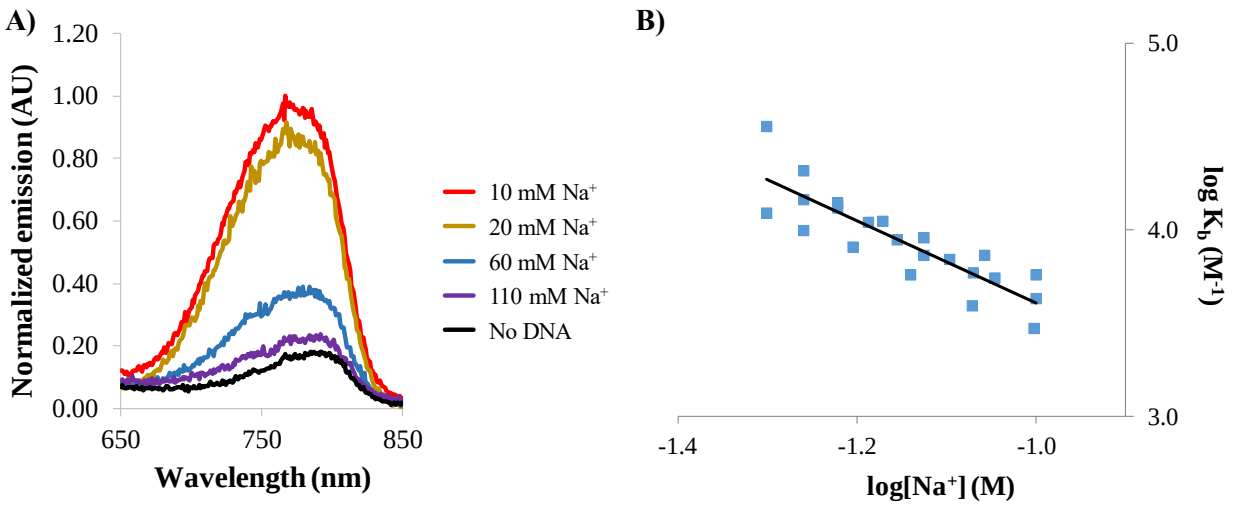


Figure 2.1. Effect of ionic strength on the non-covalent interaction of **Ru(II)-Rh(III)** with DNA. Normalized emission of **Ru(II)-Rh(III)** solutions with CT-DNA (**A**) in NaH_2PO_4/Na_2HPO_4 buffer (10 mM, pH=7.4) at various [Na^+]. Log plot of K_b and [Na^+] (**B**) obtained from reverse salt titration.

Furthermore, emission of **Ru(II)-Rh(III)** progressively decreased with increasing ionic strength of the solution, which we attribute to a reduction in the fraction of **Ru(II)-Rh(III)** non-covalently bound to DNA due to ion screening. At 110 mM Na^+, the emission spectrum approached that exhibited by the complex in the absence of DNA, both with respect to absorbance maximum and emission intensity. This again suggests the existence of a non-covalent electrostatic interaction that is largely abolished at high ionic strength, consistent with the effects on the measured K_b.

The relative contributions of electrostatic and non-electrostatic interactions in the non-covalent binding of **Ru(II)-Rh(III)** to DNA were explored through a reverse salt titration (Fig. 2.1B).[24-25] Briefly, a 30 μM solution of **Ru(II)-Rh(III)** in buffered DNA (234 μM), where ~88% of the compound was bound, was titrated with NaCl and the emission profile of the resulting

solutions was recorded. The K_b at each point was calculated from the emission intensity and using the molar emission intensities of free and DNA-bound forms and used to estimate Gibbs free energies as described in the methods section. Binding of **Ru(II)-Rh(III)** to DNA exhibited a ΔG_{obs} of −6.1 kcal/mol at 50 mM Na^+, with a 63% and 37% contribution from the electrostatic (−3.8 kcal/mol) and non-electrostatic (−2.3 kcal/mol) components. Similar values were reported for a bimetallic Ru(II) complex that binds electrostatically to DNA,[24] thus supporting the proposed binding mode for **Ru(II)-Rh(III)**.

To further probe the nature of the non-covalent **Ru(II)-Rh(III)**-DNA interaction, emission spectroscopy was used in competitive binding experiments with **Ru(II)-Rh(III)** and Hoechst 33342 (**H**), a well-established minor groove binder.[45-46] All incubations were performed in 10 mM [Na^+] at room temperature and in the absence of light. Under these conditions, both molecules exhibited strong emission in the presence of DNA, with emission in the absence of DNA approximately five-fold lower for **Ru(II)-Rh(III)** and at close to undetectable levels for **H** (Fig. 2.2).

Due to the distinctly different spectroscopic properties of **H** and **Ru(II)-Rh(III)**, the emission from each compound can be measured separately by using different excitation and emission wavelengths. As shown in Fig. 2.2B, the emission of **Ru(II)-Rh(III)** in a solution of CT-DNA was ~25% lower in the presence of **H** (solid purple line) than in its absence (solid red line), which we attribute to competition for binding in the minor groove, similar to previous reports for $[Ru(phen)_3]^{2+}$.[47-48] A control prepared with **Ru(II)-Rh(III)** and **H** in buffered aqueous solution without DNA (dashed purple line) showed emission comparable to that of the metal complex alone (dashed red line), thus confirming that free **H** does not quench the emission of **Ru(II)-Rh(III)**.

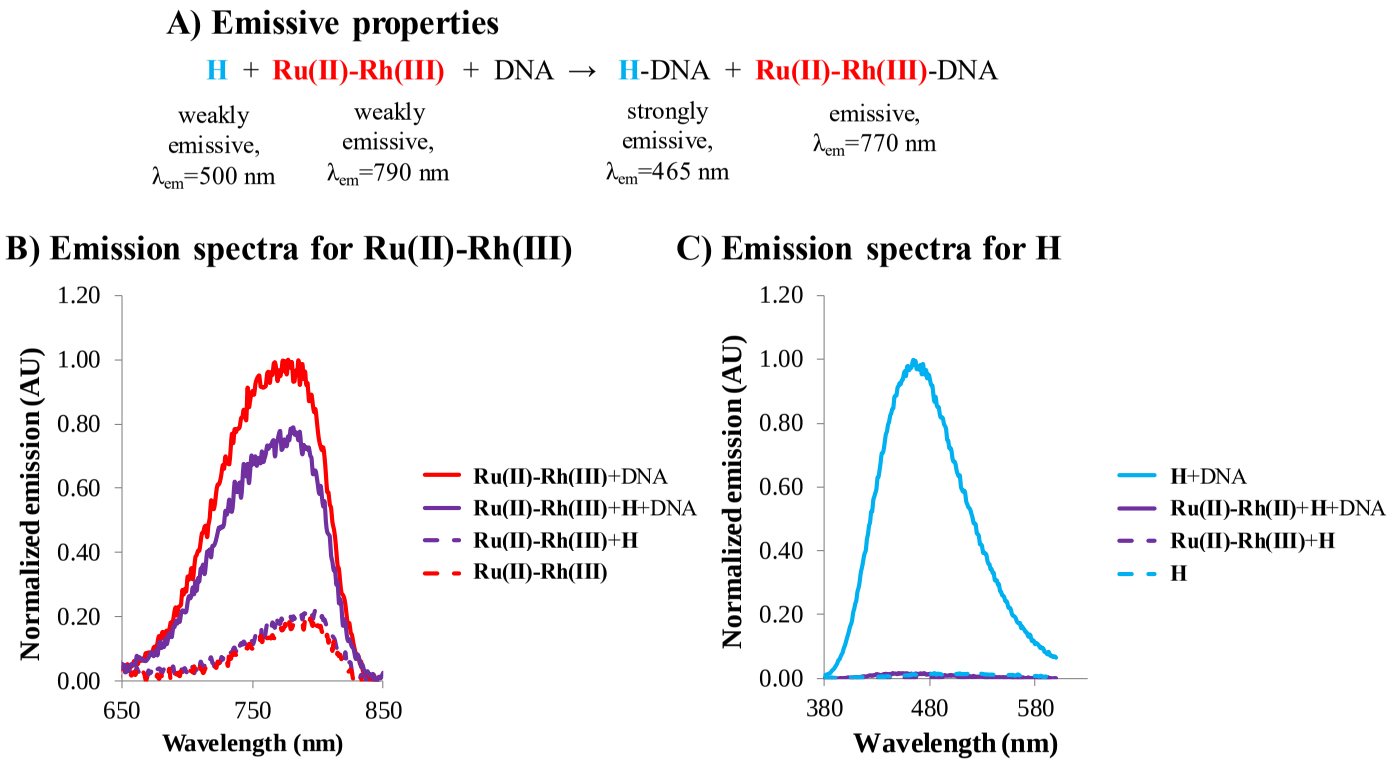


Figure 2.2. Competitive binding assays with DNA minor groove binder Hoechst 33342 (**H**). General photophysical properties of DNA-bound and free forms of **Ru(II)-Rh(III)** and **H** (**A**) and emission spectra of their solutions in the presence or absence of DNA upon excitation at 513 (**B**) or 355 (**C**) nm, corresponding to **Ru(II)-Rh(III)** and **H**, respectively.

The emission spectra of **H** in the same solutions displayed a more striking behavior. Fluorescence of **H** non covalently bound to DNA (Fig. 2.2C, pink line) was almost completely abolished in the presence of **Ru(II)-Rh(III)** (solid purple line, Fig. 2.2C and S2.2). It is unlikely that the observed decrease is exclusively due to competition for binding sites in the DNA since the binding constant of **H** is significantly higher ($K_b \sim 10^7\ M^{-1}$ in 100-150 mM ionic strength[49-50]) than that of **Ru(II)-Rh(III)**. Instead, we propose that these two molecules bind to DNA in close proximity to each other and thus can undergo Förster resonant energy transfer (FRET). This is supported by the overlap between the emission spectrum of **H** and the absorbance spectrum of **Ru(II)-Rh(III)** (Fig. S2.3), with a calculated Förster radius (R_0) of 37 Å. Furthermore, the fluorescence lifetime of the **H**-DNA adduct was substantially decreased in the presence of **Ru(II)-Rh(III)** (from 3.0 ns to almost undetectable at 465 nm; Fig. S2.4), further evidence for the occurrence of FRET between these molecules. FRET between chromophore pairs non-covalently bound to DNA has been previously reported, including for donor/acceptor pairs such as an intercalating Ru(II) complex with BO-PRO-3,[51] and for the minor groove binder Hoechst 33258 with the intercalator ethidium bromide.[52]

The electrochemical and photophysical properties of **Ru(II)-Rh(III)** are similar to those of previously-reported Ru^{II},Rh^{III} bimetallic complexes, thus suggesting the potential for light-promoted, oxygen-independent covalent modification of DNA.[14] To test this, gel shift assays (Fig. 2.3A) were performed on samples of supercoiled pUC19 plasmid DNA (pDNA, 2686 bp) treated with the complex under a variety of conditions; in all cases purged with argon to remove dissolved O_2. Under dark conditions (lanes D1-D4), pDNA did not show a significant difference in migration from control samples that did not contain the complex (lane C). However, blue light photolysis (lane L2) promoted covalent binding of the bimetallic complex to DNA, as evidenced by the slower

migration of supercoiled (SC) plasmid, attributed to increased mass, charge neutralization and disruption of DNA tertiary structure upon binding of the complex. In turn, photolysis lead to conversion of supercoiled DNA into the open circular (OC) form of the plasmid, which is attributed to single strand cleavage. Interestingly, we did not observe formation of linear DNA even after 6 h of irradiation (shown in Fig. S2.5), suggesting that double strand cleavages were not generated.

This experiment also indicated that covalent binding and cleavage occur in the absence of oxygen because all samples were purged with argon prior to photolysis. To test this further, ROS scavengers (i.e., benzoate, iodide and azide) were added to the samples prior to photolysis (Fig. 2.3B). These treatments also had no effect on the cleavage activity of **Ru(II)-Rh(III)** (lanes S1-4 vs L). These results are consistent with an oxygen-independent and metal-mediated DNA photobinding and photocleavage mechanism, as previously proposed for other multimetallic complexes with rhodium bioactive centers.[14, 38, 53]

The effect of photolyzing pDNA with **Ru(II)-Rh(III)** in solutions of different ionic strengths was also examined (Fig. 2.3A). We observed that an increase in the ionic strength of the solution from 0.8 mM (L1) to 8 mM (L2) and 80 mM (L3) progressively decreased the extent of DNA photobinding and photocleavage, with L3 exhibiting a substantial proportion of unreacted SC DNA. We attribute this behavior to the reciprocal relationship between K_b and ionic strength, which restricts the amount of **Ru(II)-Rh(III)** non-covalently bound to DNA and thus readily available to react with this biomolecule upon photoactivation. Substitution of phosphate for chloride (L3 and L4, respectively) in ionic strength-matched samples led to the same DNA profile after photolysis, thus showing that the activity of **Ru(II)-Rh(III)** is unconstrained by the chemical identity of the anion, and further supporting the critical effect of ionic strength and non-covalent interactions in DNA photomodification.

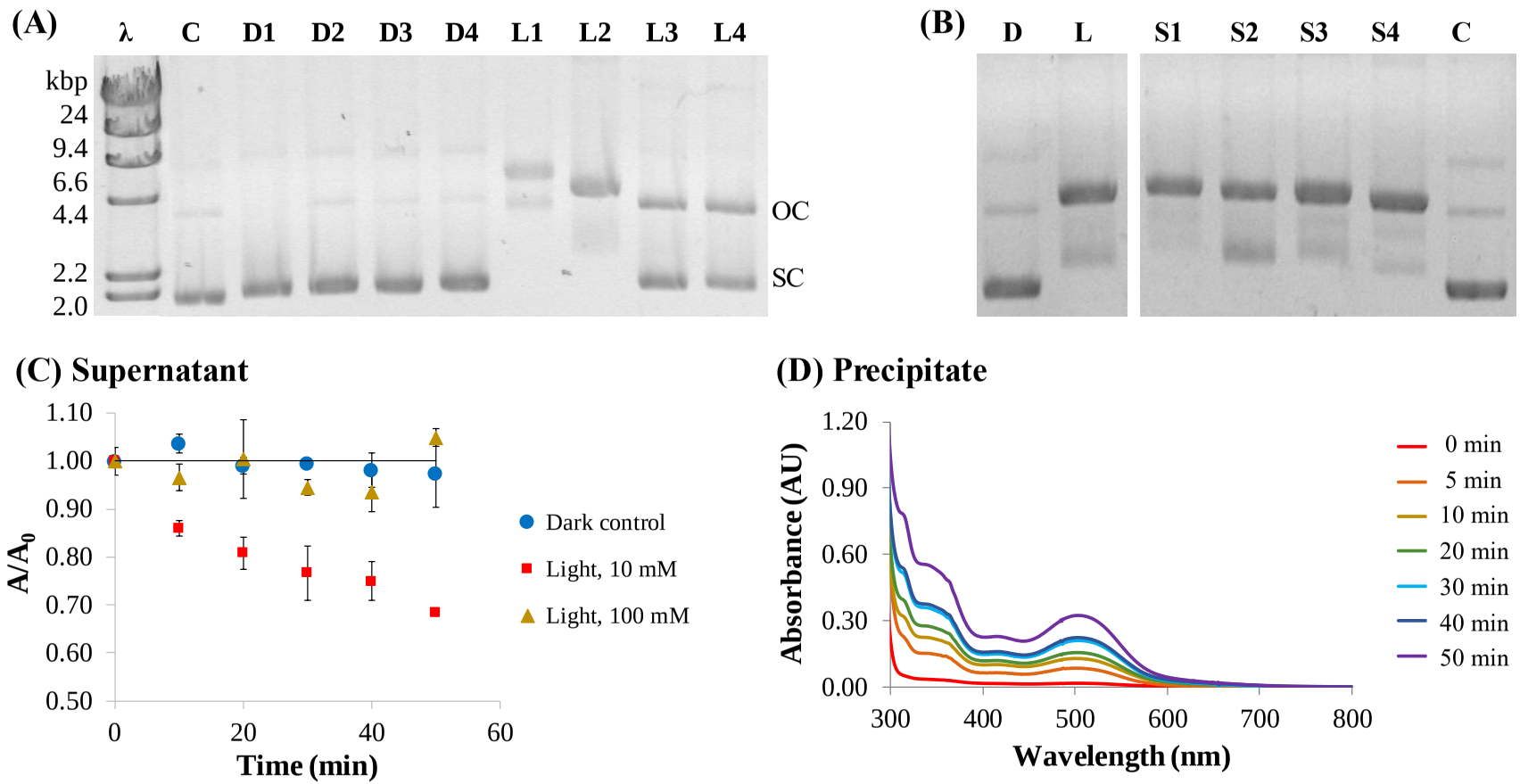


Figure 2.3. Analysis of covalent modification of DNA by **Ru(II)-Rh(III).** (**A**) DNA gel shift assay of supercoiled plasmid DNA (pUC19, 2686 bp). Lanes are: molecular weight ladder (λ), untreated pDNA control (C), and argon-purged solutions of pDNA incubated with **Ru(II)-Rh(III)** in the dark (D samples) or photolyzed at 455 nm (L samples) for 1 h in 0.8 mM (D1 and L1), 8 mM (D2, D4, L2 and L4) or 80 mM (D3 and L3) NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer, and 72 mM NaCl (D4 and L4). (**B**) Gel shift assay

of DNA photolyzed in the presence of ROS scavengers. Lane C (control) contains untreated DNA; D (dark) and L (light) samples were treated under the same conditions as samples D2 and L2, respectively, in panel A; S1, S2, S3 and S4 correspond to DNA and **Ru(II)-Rh(III)** solutions spiked with 1 mM sodium benzoate, KI, NaCl or NaN_3, respectively, and photolyzed (λ_{irr}=455 nm) for 1 h. S3 served as ionic strength control. 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer was used in all samples. (**C**) Relative absorbance (A/A_0) of the supernatant of solutions of CT-DNA and **Ru(II)-Rh(III)** incubated in the dark in 10 mM NaH_2PO_4/Na_2HPO_4 buffer or photolyzed at two ionic strengths (10 mM and 100 mM) as a function of incubation time, after selective precipitation of DNA with ethanol/NaCl. (**D**) Absorbance spectra of CT-DNA pellets redissolved in 10 mM NaH_2PO_4/Na_2HPO_4 as a function of photolysis time.

Gel electrophoresis allows a sensitive detection of cleavage events due to the wide difference in mobility between OC and SC DNA. However, it is less sensitive for the evaluation of covalent binding given the substantially smaller molecular weight of **Ru(II)-Rh(III)** (~1.17 kDa) relative to pDNA (~1.66 MDa). Thus, a selective precipitation assay was used to assess the proportion of complex that was covalently bound to DNA under various conditions. In this assay, CT-DNA was photolyzed in the presence of **Ru(II)-Rh(III)** and precipitated by the addition of ethanol/NaCl, while unbound **Ru(II)-Rh(III)** remained in solution and could be detected by absorbance spectroscopy. The absorbance at each time point (A) was then normalized to that of a sample precipitated without any treatment (A_0).

As shown in Fig. 2.3C, the relative absorbance (A/A_0) of DNA samples incubated with **Ru(II)-Rh(III)** in 10 mM buffer in the dark (blue circles) did not change significantly over the 50 min time course. However, photolysis resulted in a substantial decrease in complex remaining in solution, with some 30% binding covalently to DNA after 50 min of photolysis (red squares). Photobinding of **Ru(II)-Rh(III)** to DNA was slower but also continued much longer than reported for the bpy analog, which plateaus at 30 min at ~60% binding.[14] Binding of **Ru(II)-Rh(III)** proceeded for up to 22 h of photolysis, with its concentration in the supernatant decreasing by ~80% (Fig. S2.6). The exact nature of the difference in behavior is unclear but could involve several factors related to the structural features of the molecules, including sterics and non-covalent binding strength, and excited state properties, such as the quantum yields and kinetics for the population of the ^{3}MMCT state and ligand dissociation.

Further evidence that the complex depleted from the solution was covalently bound to the CT-DNA came from absorbance spectroscopy of the precipitated material (i.e., pellet) after it was redissolved in buffer. The resulting solutions displayed an absorbance spectrum with two local

maxima in the visible range (417 and 506 nm), which are consistent with **Ru(II)-Rh(III)**-based transitions (Fig. 2.3D). Moreover, absorbance intensity increased as a function of irradiation time, in agreement with the proposed light-promoted binding mechanism.

Interestingly, solutions photolyzed with 100 mM [Na^+] showed minimal to undetectable covalent binding (Fig. 2.3C, gold triangles) in agreement with the results of gel shift assays. Given that some cleavage of DNA was observed at this and similar ionic strengths (Fig. S2.7, L1, and Fig. 2.2, L3 and L4), we hypothesize that cleavage and covalent binding are independent events, and that nicking might occur through a collision-based process not involving covalent binding. Such behavior is similar to that of Ru^{II},Rh^{III},Ru^{II} trimetallic systems, which display DNA photocleavage but not photobinding,[38] presumably due to steric effects with the bulky [$(TL)_2$Ru(dpp)] subunits around the Rh center. Furthermore, this suggests that the lifetime of the Ru→Rh ^{3}MMCT state is long enough to allow diffusion of free excited complex from solution and collision with DNA prior to its decay.

The potential for the complex to interfere with the replication of DNA was examined by polymerase chain reaction (PCR). Solutions of pUC18 (pDNA*, 2686 bp) DNA and **Ru(II)-Rh(III)** under two different ionic strengths (10 mM and 100 mM)were incubated under dark conditions or photolyzed at 455 nm for different periods of time (Fig. 2.4). After these treatments, the samples were diluted and the DNA was amplified by PCR using two different pairs of forward and reverse primers (Table S2.2) that target different regions of the plasmid. The reactions were then analyzed by gel electrophoresis. Samples photolyzed at low ionic strength (i.e., 10 mM Na^+, Fig. 2.4A-B), where both covalent binding and cleavage occur, displayed progressively decreased amplification with increased irradiation time for both primer sets, with photolysis for 2 to 3 h almost completely abolishing amplification. This is consistent with covalent binding of the

complex and/or damage to DNA that might affect the denaturation of the template, block the extension step or interfere with the binding of the primers and/or the polymerase to DNA . Samples incubated in the dark (D) for 6 h did not display any inhibition of amplification compared to the untreated positive (+) controls, indicating that little binding or damage occurs in the absence of photolysis.

Amplification of samples that were photolyzed at high ionic strength (i.e., 100 mM Na^+, Fig. 2.4C-D), where covalent binding is widely hindered as discussed above, were also not significantly inhibited compared to the untreated positive (+) controls. This behavior was consistent throughout all of the photolysis times tested, including 24 h, even though the DNA was progressively cleaved under these conditions (see Fig. S2.7). Although DNA cleavage could potentially interfere with amplification, nicking events generated at high ionic strength might not be frequent enough or occur within short enough distances to have a detectable effect under the conditions used here. We attribute this behavior to a decrease in non-covalent interactions between **Ru(II)-Rh(III)** and DNA at higher ionic strength, thus hindering covalent photobinding and leaving sufficient unmodified template to allow amplification by PCR.

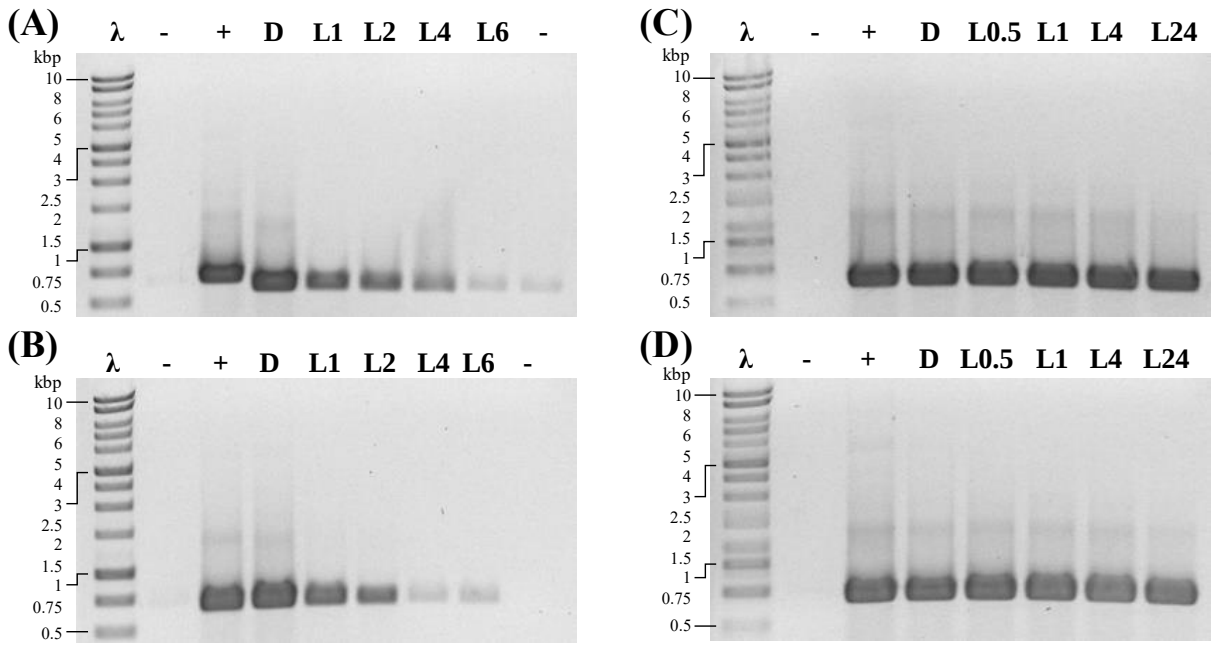


Figure 2.4. Effects of treatment with **Ru(II)-Rh(III)** on amplification of pUC18 plasmid DNA by PCR**.** Two different sets of forward and reverse primers (set 1 in gels **A-C**, set 2 in gels **B-D**) were used in these experiments, each targeting a different region of pDNA. Gels contain 1 kb molecular weight marker (λ), negative control (-) prepared using nuclease-free water instead of template, positive control (+) prepared with untreated pDNA*, dark (D) control that was incubated protected from light with **Ru(II)-Rh(III)** for 6 h in 10 mM buffer (gels **A-B**) or 24 h in 100 mM Na^+ (gels **C-D**) prior to amplification; whereas L0.5, L1, L2, L4, L6 and L24 samples were photolyzed in the presence of **Ru(II)-Rh(III)** and 10 mM (gels **A-B**) or 100 mM Na^+ (gels **C-D**) for 0.5, 1, 2, 4, 6 or 24 h, respectively.

2.5. Conclusions and future work

Despite substantial progress in recent years in developing novel anti-cancer drugs, including platinating, alkylating and photodynamic therapeutic agents, there remains an urgent need for options that overcome the limitations of current treatments. Ru^{II},Rh^{III} bimetallic systems

display an unusual mechanism of action, where light promotes population of a highly reactive ^{3}MMCT that photoreacts with DNA without the need for molecular oxygen, whereas the ground state molecule remains inactive.[14] Such molecules could allow for selective tuning of toxicity to those areas where the tumor is located, even under hypoxic conditions, using excitation sources with deeper tissue penetration than other rhodium-based bioactive molecules that have absorbance maximum at shorter wavelengths.[12-13]

In this study, we have aimed to better understand the nature of the interaction of one such bimetallic system with DNA. We found that **Ru(II)-Rh(III)** displayed enhanced emission intensity in the presence of DNA, i.e. "light-switch" behavior, and engaged in strong electrostatic interactions that were substantially influenced by ionic strength. Competitive DNA binding experiments with **H** suggest that these interactions may involve groove binding, further evidenced by an apparent FRET interaction between non-covalently bound **H** and **Ru(II)-Rh(III)**. The noncovalent interactions serve as a potent preassociation step that substantially facilitates the covalent modification of DNA by **Ru(II)-Rh(III)** upon photoactivation in the visible range, as evidenced in both gel shift assays and effects on amplification of DNA by PCR. These behaviors are highly desirable for photodynamic therapy drugs, given that any molecule that associates non-covalently with DNA in off-target tissues would not signficantly inhibit its replication, whereas its activity could be selectively triggered in the tumor through photoactivation. Moreover, the observed multifunctional covalent interactions with DNA proceeded in an oxygen-independent manner, a further essential property for the treatment of hypoxic tumors. Both **Ru(II)-Rh(III)** and the previously-reported bpy analog photobind and photocleave DNA, supporting the proposed role of a ^{3}MMCT state as the species responsible for this activity. The slower kinetics and higher extent of binding exhibited by **Ru(II)-Rh(III)** suggest that other factors (e.g., sterics and ligand

photolabilization) also play a role in photomodification of DNA, a consideration for the design and evaluation of new molecular architectures.

The current findings set the stage for future mechanistic studies to better understand the interaction of this class of potential anti-cancer agents with biological macromolecules, including proteins, and to examine the efficiency of uptake into cells and entry into the nucleus. Reports of ruthenium and rhodium complexes with sequence selective binding of DNA, indicate that this is another intriguing area for future investigation.[54-55]

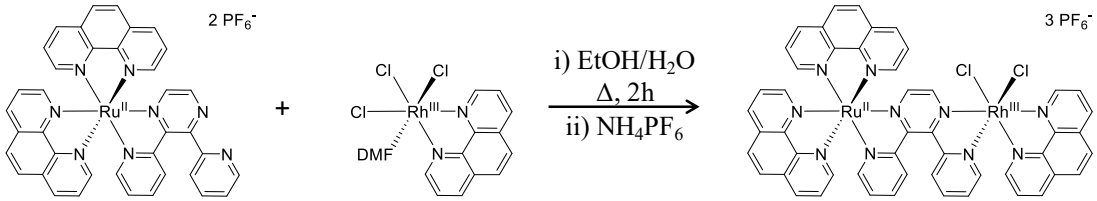
Together with prior studies of related complexes, this work provides further evidence of the utility of bimetallic systems for the design of novel PDT agents. In addition, it demonstrates that a complex with phenanthroline in all three terminal ligand positions retains many of the positive features of related mixed-metal architectures, while displaying several unique properties that offer the potential for further improvement of these systems, including functionalization of the phenanthroline moiety (e.g., by incorporation of electron withdrawing or donating substituents, or aromatic or aliphatic groups).

2.6. Acknowledgments

The authors acknowledge the generous support of this work by grants from the National Science Foundation (CHE-1301131) and Virginia Tech's Institute for Critical Technology and Applied Science (119552). We also thank Drs. Tijana Z. Grove, John L. Robertson, Jie Zhu, and Roberto Padilla for helpful discussions and Dr. Jatinder Josan for his support. We extend a special thank you to Dr. Roberto Padilla for developing the DNA structures presented in the graphical abstract. We particularly recognize the immense talent and contribution to science of the late Dr. Karen J. Brewer, who oversaw and inspired this work.

2.7. Supporting information

2.7.1. Synthesis of Ru(II)-Rh(III)



Scheme S2.1. Synthetic scheme for $[(phen)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$.

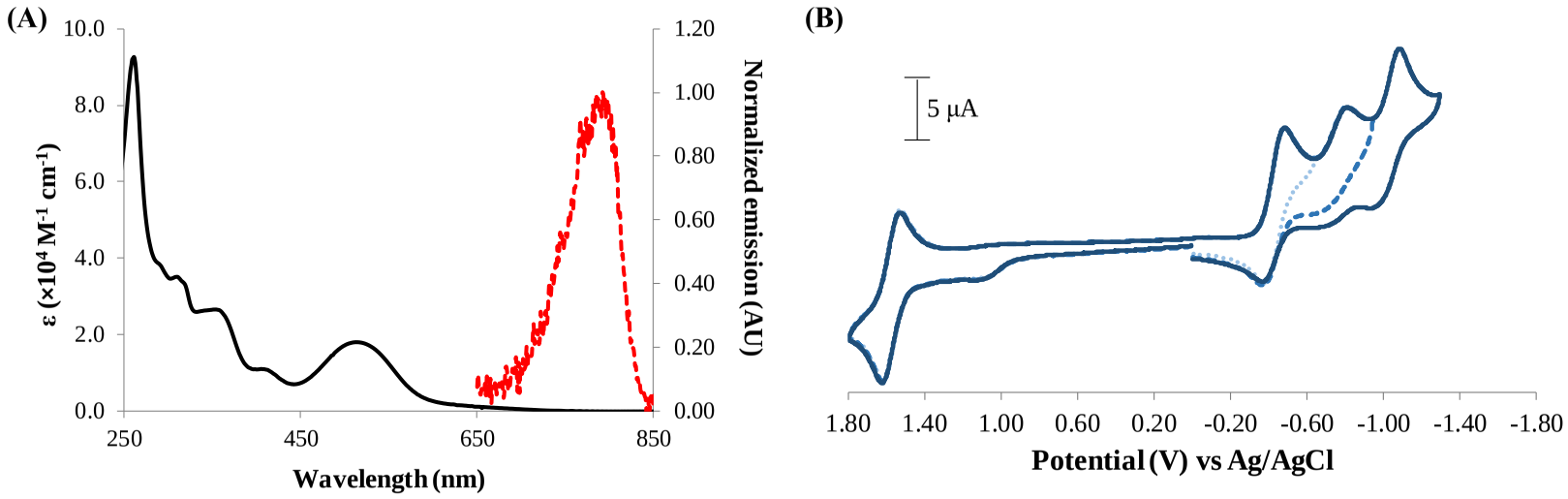


Figure S2.1. Spectrophotometric and electrochemical characterization of $[(phen)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$ (phen=1,10-phenanthroline and dpp=2,3-bis(2-pyridyl)pyrazine). (**A**) Electronic absorption (solid black line) and emission (dashed red line, $\lambda_{excitation}$ = 540 nm) spectra, and (**B**) cyclic voltammogram in 0.1M Bu_4NPF_6 acetonitrile, E vs. Ag/AgCl, ν=100 mV/s.

Table S2.1. Comparison of the spectrophotometric and electrochemical properties of $[(TL)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$, where TL=phen or 2,2'-bipyridine (bpy)).

Technique	Parameter	Assignment	Complex	
			$[(phen)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$	**$[(bpy)_2Ru(dpp)RhCl_2(phen)](PF_6)_3$ [d]**
Electronic absorption spectrometry	λ_{max} [a]	TL $\pi \rightarrow \pi^*$	261 nm	279 nm
		dpp $\pi \rightarrow \pi^*$	357 nm	338 nm
		$Ru(d\pi) \rightarrow TL\ (\pi^*)$ ^{1}MLCT	415 nm	418 nm
		$Ru(d\pi) \rightarrow dpp\ (\pi^*)$ ^{1}MLCT	513 nm	509 nm
Electronic emission spectrometry	$\lambda_{emission}$	$^1GS \leftarrow {}^3MLCT$	750 nm (acetonitrile), 790 nm (water)‡	786 nm (acetonitrile)
	$\Phi^{em,c}$		2.4×10^{-4}	2.3×10^{-4}
	τ^c		64 ns	30 ns
Cyclic voltammetry[b]	$E_{1/2}$	Reversible $Ru^{II/III}$	+1.58 V	+1.61
	E_c	Irreversible $Rh^{III/II}Cl_2$ and $Rh^{II/I}Cl$	−0.48 V	−0.39 V
	E_c	Irreversible $Rh^{II/I}Cl_2$	−0.81 V	−0.74 V
	E_c	$dpp^{0/-}$	−1.08 V	−0.98 V

‡Corresponds to the chloride salt of **Ru(II)-Rh(III)**.

[a]Measured at room temperature in acetonitrile.

[b]Measured at room temperature in 0.1M Bu_4NPF_6 acetonitrile, E vs. Ag/AgCl, ν=100 mV/s.

[c]Measured at room temperature in acetonitrile following deoxygenation with argon.

[d]Correspond to values reported in reference 18.

2.7.3. Emission spectra of solutions of Ru(II)-Rh(III) and H

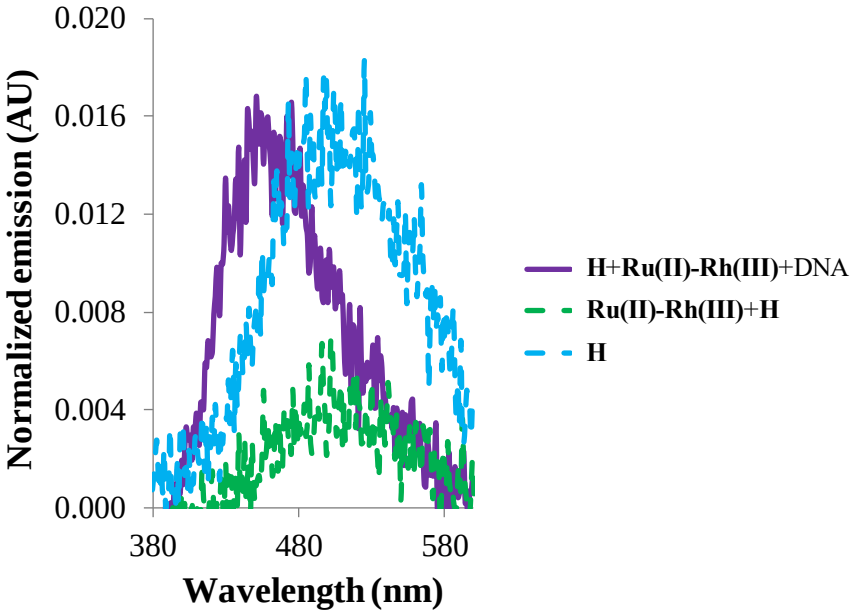


Figure S2.2. Emission spectra of 22 μM Hoechst 33342 (**H)** in the presence or absence of DNA (110 μM) and **Ru(II)-Rh(III)** (22 μM), measured in 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer upon excitation at 355 nm.

2.7.4. Spectra of H and Ru(II)-Rh(III) non-covalently bound to DNA

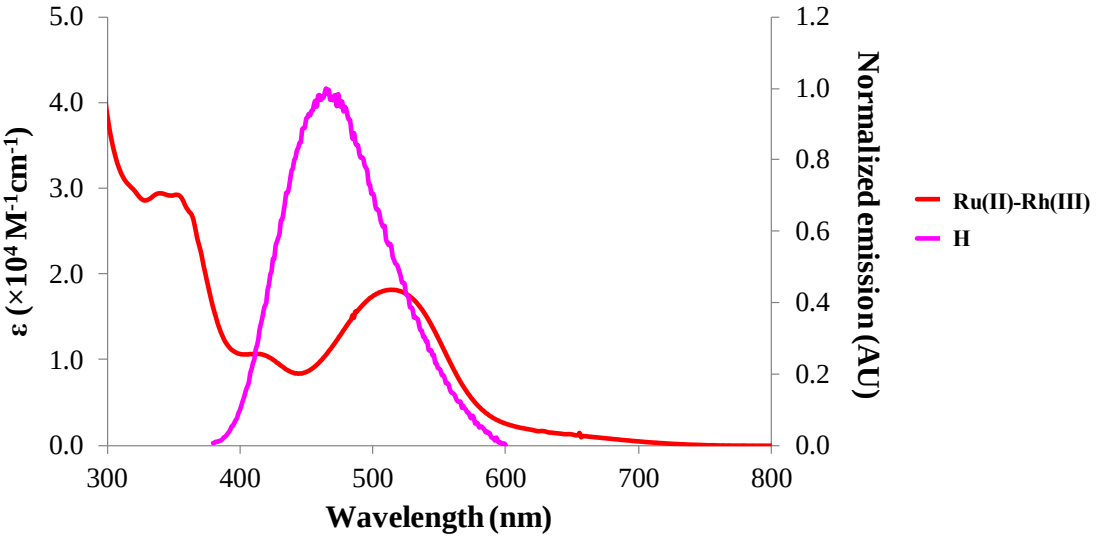


Figure S2.3. Overlay of the emission spectrum of **H** (pink, excitation at 355 nm) and the absorption spectrum of **Ru(II)-Rh(III)** (red), both measured in buffered DNA solutions.

2.7.5. Excited state lifetime of DNA-bound H in the absence and presence of Ru(II)-Rh(III)

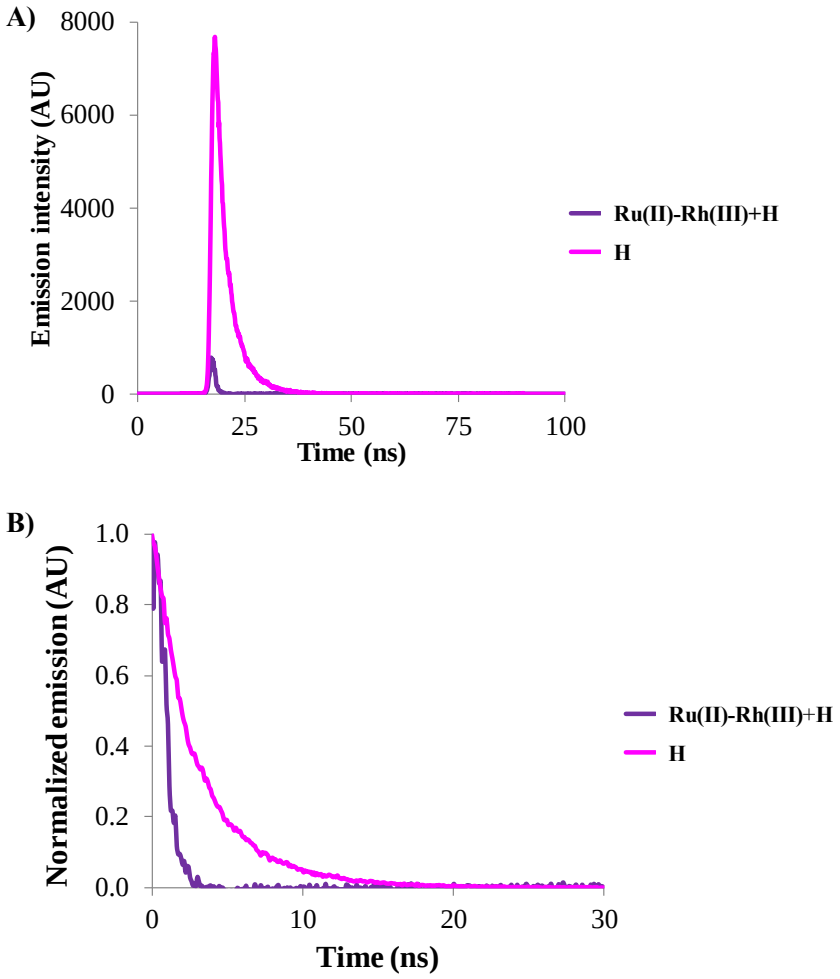


Figure S2.4. Excited state lifetime of **H** ($\lambda_{emission}$= 465 nm) in buffered aqueous DNA solutions in the absence (pink) and presence (purple) of **Ru(II)-Rh(III)**. Raw (**A**) and normalized (**B**) emission intensity are shown.

2.7.6. Photolysis of Ru(II)-Rh(III) with pDNA* at low ionic strength

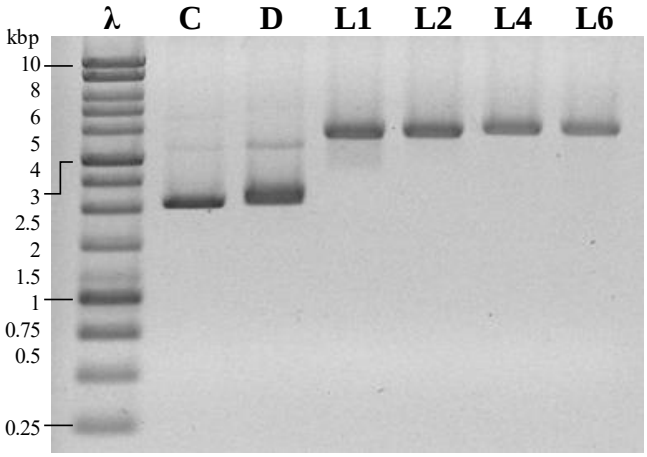


Figure S2.5. Gel shift assay of pDNA* photolyzed with **Ru(II)-Rh(III)** in 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer. **λ** = molecular weight marker, C (control)= untreated DNA, D (dark) = DNA and **Ru(II)-Rh(III)** incubated protected from light for 6 h, L1, L2, L4 and L6 = DNA and **Ru(II)-Rh(III)** photolyzed (λ_{irr}=455 nm) for 1, 2, 4 or 6 h, respectively.

2.7.7. DNA precipitation assay at extended photolysis times

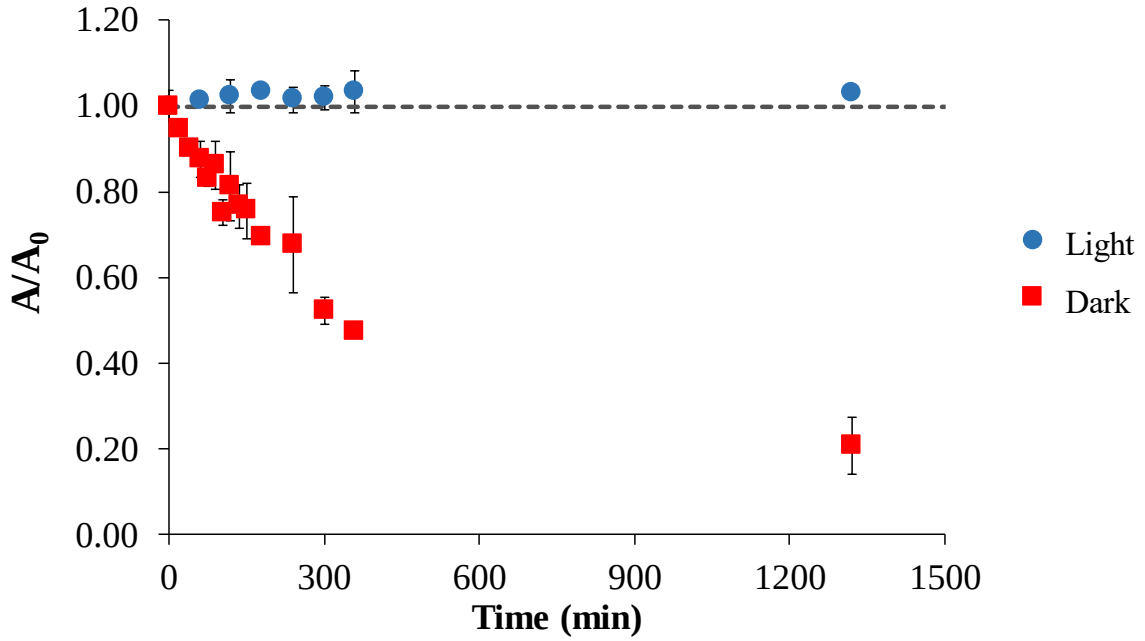


Figure S2.6. Relative absorbance (A/A_0) of the supernatant of solutions of CT-DNA and **Ru(II)-Rh(III)** in 10 mM NaH_2PO_4/Na_2HPO_4 incubated in the dark or photolyzed as a function of incubation time after precipitation of DNA with ethanol/NaCl.

2.7.8. PCR primer sequences

Table S2.2. Sequence of forward and reverse primers used in PCR amplification of pUC18.

Primer	Sequence
Forward Primer 1	5'-GCG TAT TGG GCG CTC TTC CG-3'
Reverse Primer 1	5'-TGT AGC CGT AGT TAG GCC ACC A-3'
Forward Primer 2	5'-CCC AAC GAT CAA GGC GAG-3'
Reverse Primer 2	5'-CGA AAG GGC CTC GTG ATA CG-3'

2.7.9. Photolysis of Ru(II)-Rh(III) with pDNA* at high ionic strength

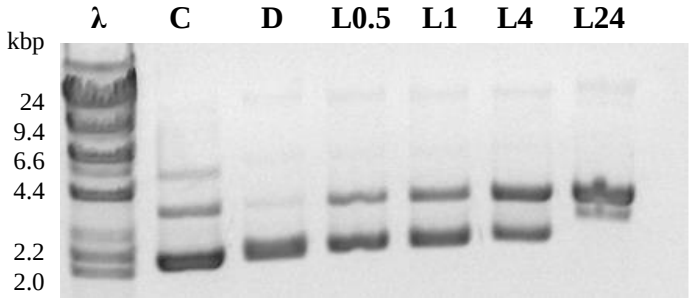


Figure S2.7. Gel shift assay of DNA photolyzed with **Ru(II)-Rh(III)** in 10 mM NaH_2PO_4/Na_2HPO_4 (pH 7.4) buffer and 90 mM NaCl. λ = molecular weight marker, C (control)= untreated DNA, D (dark) = DNA and **Ru(II)-Rh(III)** incubated protected from light for 24 h, L0.5, L1, L2 and L24 (light samples) = DNA and **Ru(II)-Rh(III)** photolyzed (λ_{irr}=455 nm) for 0.5, 1, 4 and 24 h, respectively.

2.8. References

1. World Health Organization, *World Health Statistics 2017: Monitoring health for the SDGs, Sustainable Development Goals*. World Health Organization: Geneva, Switzerland, 2017.

2. American Cancer Society *Cancer facts & figures 2018*; American Cancer Society: Atlanta, GA, 2018; pp 1-76.

3. Dasari, S.; Tchounwou, P. B., Cisplatin in cancer therapy: Molecular mechanisms of action. *Eur. J. Pharmacol.* **2014,** *740*, 364-378.

4. Miller, R. P.; Tadagavadi, R. K.; Ramesh, G.; Reeves, W. B., Mechanisms of cisplatin nephrotoxicity. *Toxins* **2010,** *2* (11), 2490-2518.

5. Rabik, C. A.; Dolan, M. E., Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat. Rev.* **2007,** *33* (1), 9-23.

6. Biel, M., Advances in photodynamic therapy for the treatment of head and neck cancers. *Lasers Surg. Med.* **2006,** *38* (5), 349-55.

7. Detty, M. R.; Gibson, S. L.; Wagner, S. J., Current clinical and preclinical photosensitizers for use in photodynamic therapy. *J. Med. Chem.* **2004,** *47* (16), 3897-3915.

8. Henderson, B. W.; Busch, T. M.; Vaughan, L. A.; Frawley, N. P.; Babich, D.; Sosa, T. A.; Zollo, J. D.; Dee, A. S.; Cooper, M. T.; Bellnier, D. A.; Greco, W. R.; Oseroff, A. R., Photofrin photodynamic therapy can significantly deplete or preserve oxygenation in human basal cell carcinomas during treatment, depending on fluence rate. *Cancer Res.* **2000,** *60* (3), 525-529.

9. Vaupel, P.; Harrison, L., Tumor hypoxia: Causative factors, compensatory mechanisms, and cellular response. *The Oncologist* **2004,** *9* (suppl 5), 4-9.

10. Knoll, J. D.; Turro, C., Control and utilization of ruthenium and rhodium metal complex excited states for photoactivated cancer therapy. *Coord. Chem. Rev.* **2015,** *282-283*, 110-126.

11. Poynton, F. E.; Bright, S. A.; Blasco, S.; Williams, D. C.; Kelly, J. M.; Gunnlaugsson, T., The development of ruthenium(II) polypyridyl complexes and conjugates for *in vitro* cellular and *in vivo* applications. *Chem. Soc. Rev.* **2017,** *46* (24), 7706-7756.

12. Mahnken, R. E.; Billadeau, M. A.; Nikonowicz, E. P.; Morrison, H., Development of photo *cis*-platinum reagents. Reaction of *cis*-dichlorobis(1,10-phenanthroline)rhodium(III) with calf thymus DNA, nucleotides and nucleosides. *J. Am. Chem. Soc.* **1992,** *114* (24), 9253-9265.

13. Sitlani, A.; Long, E. C.; Pyle, A. M.; Barton, J. K., DNA photocleavage by phenanthrenequinone diimine complexes of rhodium(III): Shape-selective recognition and reaction. *J. Am. Chem. Soc.* **1992,** *114* (7), 2303-2312.

14. Wang, J.; Zigler, D. F.; Hurst, N.; Othee, H.; Winkel, B. S. J.; Brewer, K. J., A new, bioactive structural motif: Visible light induced DNA photobinding and oxygen independent photocleavage by Ru^{II}, Rh^{III} bimetallics. *J. Inorg. Biochem.* **2012,** *116*, 135-139.

15. Wallace, A. W.; Rorer Murphy, W.; Petersen, J. D., Electrochemical and photophysical properties of mono- and bimetallic ruthenium(II) complexes. *Inorg. Chim. Acta.* **1989,** *166* (1), 47-54.

16. Sau, Y.-K.; Chan, K.-W.; Zhang, Q.-F.; Williams, I. D.; Leung, W.-H., Alkyl and aryl compounds of rhodium(III) and iridium(III) supported by 4,4'-di-*tert*-butyl-2,2'-bipyridyl. *Organometallics* **2007,** *26* (25), 6338-6345.

17. White, T. A.; Mallalieu, H. E.; Wang, J.; Brewer, K. J., Mechanistic insight into the electronic influences imposed by substituent variation in polyazine-bridged ruthenium(II)/rhodium(III) supramolecules. *Chem – Eur. J.* **2014,** *20* (27), 8265-8268.

18. Zigler, D. F.; Wang, J.; Brewer, K. J., Ruthenium(II)-polyazine light absorbers bridged to reactive *cis*-dichlororhodium(III) centers in a bimetallic molecular architecture. *Inorg. Chem.* **2008,** *47* (23), 11342-11350.

19. Caspar, J. V.; Kober, E. M.; Sullivan, B. P.; Meyer, T. J., Application of the energy gap law to the decay of charge-transfer excited states. *J. Am. Chem. Soc.* **1982,** *104* (2), 630-632.

20. Reichmann, M. E.; Rice, S. A.; Thomas, C. A.; Doty, P., A further examination of the molecular weight and size of desoxypentose nucleic acid. *J. Am. Chem. Soc.* **1954,** *76* (11), 3047-3053.

21. Marmur, J., A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* **1961,** *3* (2), 208-218.

22. Schmechel, D. E. V.; Crothers, D. M., Kinetic and hydrodynamic studies of the complex of proflavine with poly A·poly U. *Biopolymers* **1971,** *10* (3), 465-480.

23. Wolfe, A.; Shimer, G. H.; Meehan, T., Polycyclic aromatic hydrocarbons physically intercalate into duplex regions of denatured DNA. *Biochemistry* **1987,** *26* (20), 6392-6396.

24. Lutterman, D. A.; Chouai, A.; Liu, Y.; Sun, Y.; Stewart, C. D.; Dunbar, K. R.; Turro, C., Intercalation is not required for DNA light-switch behavior. *J. Am. Chem. Soc.* **2008,** *130* (4), 1163-1170.

25. Haq, I.; Lincoln, P.; Suh, D.; Norden, B.; Chowdhry, B. Z.; Chaires, J. B., Interaction of Δ- and Λ-$[Ru(phen)_2DPPZ]^{2+}$ with DNA: A calorimetric and equilibrium binding study. *J. Am. Chem. Soc.* **1995,** *117* (17), 4788-4796.

26. Lakowicz, J. R., Energy Transfer. In *Principles of Fluorescence Spectroscopy*, Springer US: Boston, MA, 1999; pp 367-394.

27. Sabnis, R. W., Hoechst 33342. In *Handbook of Fluorescent Dyes and Probes*, John Wiley & Sons: Hoboken, NJ, 2015; pp 238-239.

28. Prussin, A. J., II; Zigler, D. F.; Jain, A.; Brown, J. R.; Winkel, B. S. J.; Brewer, K. J., Photochemical methods to assay DNA photocleavage using supercoiled pUC18 DNA and LED or xenon arc lamp excitation. *J. Inorg. Biochem.* **2008,** *102* (4), 731-739.

29. Hasty, N.; Merkel, P. B.; Radlick, P.; Kearns, D. R., Role of azide in singlet oxygen reactions: Reaction of azide with singlet oxygen. *Tetrahedron Lett.* **1972,** *13* (1), 49-52.

30. Harbour, J. R.; Issler, S. L., Involvement of the azide radical in the quenching of singlet oxygen by azide anion in water. *J. Am. Chem. Soc.* **1982,** *104* (3), 903-905.

31. Thomas, J. K., Rates of reaction of the hydroxyl radical. *Trans. Faraday Soc.* **1965,** *61*, 702-707.

32. Ishimitsu, S.; Mishima, I.; Tsuji, S.; Shibata, T., Formation of a hydroxyl radical from riboflavin sodium phosphate by photo-illumination. *Chem. Pharm. Bull.* **1997,** *45* (12), 2107-2109.

33. Cohen, G.; Heikkila, R. E., The generation of hydrogen peroxide, superoxide radical, and hydroxyl radical by 6-hydroxydopamine, dialuric acid, and related cytotoxic agents. *J. Biol. Chem.* **1974,** *249* (8), 2447-2452.

34. Neta, P.; Dorfman, L. M., Pulse radiolysis studies. XIII. Rate constants for the reaction of hydroxyl radicals with aromatic compounds in aqueous solutions. In *Radiation Chemistry*, Hart, E. J., Ed. American Chemical Society: Washington, DC, 1968; Vol. 81, pp 222-230.

35. Tzeng, D. D.-S.; Lee, M.-H., Production of hydroxyl radicals in photodynamic action of methionine riboflavin mixture: A consequence of iron catalyzed Haber–Weiss reaction. *Bot. Bull. Academia Sinica* **1989,** *30*, 171-178.

36. Barton, J. K.; Lolis, E., Chiral discrimination in the covalent binding of bis(phenanthroline)dichlororuthenium(II) to B-DNA. *J. Am. Chem. Soc.* **1985,** *107* (3), 708-709.

37. Higgins, S. L. H.; Tucker, A. J.; Winkel, B. S. J.; Brewer, K. J., Metal to ligand charge transfer induced DNA photobinding in a Ru(II)-Pt(II) supramolecule using red light in the therapeutic window: A new mechanism for DNA modification. *Chem. Commun.* **2012,** *48* (1), 67-69.

38. Holder, A. A.; Swavey, S.; Brewer, K. J., Design aspects for the development of mixed-metal supramolecular complexes capable of visible light induced photocleavage of DNA. *Inorg. Chem.* **2004,** *43* (1), 303-308.

39. Pyle, A. M.; Rehmann, J. P.; Meshoyrer, R.; Kumar, C. V.; Turro, N. J.; Barton, J. K., Mixed-ligand complexes of ruthenium(II): Factors governing binding to DNA. *J. Am. Chem. Soc.* **1989,** *111* (8), 3051-3058.

40. O'Reilly, F. M.; Kelly, J. M., Binding of bimetallic 1,10-phenanthroline ruthenium(II) complexes to DNA. *New. J. Chem.* **1998,** *22* (3), 215-217.

41. Friedman, A. E.; Chambron, J. C.; Sauvage, J. P.; Turro, N. J.; Barton, J. K., Molecular light switch for DNA: $Ru(bpy)_2(dppz)^{2+}$. *J. Am. Chem. Soc.* **1990,** *112* (12), 4960-4962.

42. Zhang, A.-G.; Zhang, Y.-Z.; Duan, Z.-M.; Wang, K.-Z.; Wei, H.-B.; Bian, Z.-Q.; Huang, C.-H., Dual molecular light switches for pH and DNA based on a novel Ru(II) complex. A non-intercalating Ru(II) complex for DNA molecular light switch. *Inorg. Chem.* **2011,** *50* (14), 6425-6436.

43. P., F. S.; Tim, P.; R., G. M.; Michael, T.; W., P. A.; Michelle, W.; A., T. J., A multifunctional light switch: DNA binding and cleavage properties of a heterobimetallic ruthenium–rhenium dipyridophenazine complex. *Angew. Chem. Int. Ed.* **2007,** *46* (20), 3686-3688.

44. Moon, S. J.; Kim, J. M.; Choi, J. Y.; Kim, S. K.; Lee, J. S.; Jang, H. G., $[Ru(phen)_2DPPZ]^{2+}$ is in contact with DNA bases when it forms a luminescent complex with single-stranded oligonucleotides. *J. Inorg. Biochem.* **2005,** *99* (5), 994-1000.

45. Squire, C. J.; Baker, L. J.; Clark, G. R.; Martin, R. F.; White, J., Structures of *m*-iodo Hoechst–DNA complexes in crystals with reduced solvent content: Implications for minor groove binder drug design. *Nucleic Acids Res.* **2000,** *28* (5), 1252-1258.

46. Latt, S. A.; Stetten, G., Spectral studies on 33258 Hoechst and related bisbenzimidazole dyes useful for fluorescent detection of deoxyribonucleic acid synthesis. *J. Histochem. Cytochem.* **1976,** *24* (1), 24-33.

47. Eriksson, M.; Leijon, M.; Hiort, C.; Norden, B.; Graeslund, A., Minor groove binding of $[Ru(phen)_3]^{2+}$ to $[d(CGCGATCGCG)]_2$ evidenced by two-dimensional NMR. *J. Am. Chem. Soc.* **1992,** *114* (12), 4933-4934.

48. Eriksson, M.; Leijon, M.; Hiort, C.; Norden, B.; Graeslund, A., Binding of Δ- and Λ-$[Ru(phen)_3]^{2+}$ to $[d(CGCGATCGCG)]_2$ studied by NMR. *Biochemistry* **1994,** *33* (17), 5031-5040.

49. Bazhulina, N. P.; Nikitin, A. M.; Rodin, S. A.; Surovaya, A. N.; Kravatsky, Y. V.; Pismensky, V. F.; Archipova, V. S.; Martin, R.; Gursky, G. V., Binding of Hoechst 33258 and its derivatives to DNA. *J. Biomol. Struct. Dyn.* **2009,** *26* (6), 701-718.

50. Harapanhalli, R. S.; McLaughlin, L. W.; Howell, R. W.; Rao, D. V.; Adelstein, S. J.; Kassis, A. I., [^{125}I/^{127}I]IodoHoechst 33342: Synthesis, DNA binding, and biodistribution. *J. Med. Chem.* **1996,** *39* (24), 4804-4809.

51. Kang, J. S.; Piszczek, G.; Lakowicz, J. R., Enhanced emission induced by FRET from a long-lifetime, low quantum yield donor to a long-wavelength, high quantum yield acceptor. *J. Fluoresc.* **2002,** *12* (1), 97-103.

52. Verma, P. K.; Pal, S. K., Ultrafast resonance energy transfer in bio-molecular systems. *Eur. Phys. J. D.* **2010,** *60* (1), 137-156.

53. Wang, J.; Higgins, S. L. H.; Winkel, B. S. J.; Brewer, K. J., A new Os,Rh bimetallic with O_2 independent DNA cleavage and DNA photobinding with red therapeutic light excitation. *Chem. Commun.* **2011,** *47* (35), 9786-9788.

54. Zeglis, B. M.; Pierre, V. C.; Barton, J. K., Metallo-intercalators and metallo-insertors. *Chem. Commun.* **2007,** *44*, 4565-4579.

55. Wachter, E.; Moyá, D.; Glazer, E. C., Combining a Ru(II) "building block" and rapid screening approach to identify DNA structure-selective "light switch" compounds. *ACS Comb. Sci.* **2017,** *19* (2), 85-95.

3. GENETIC AND MOLECULAR CHARACTERIZATION APPROACHES FOR IDENTIFYING HETEROBIVALENT AND TRIVALENT TUMOR FINGERPRINTS THAT CAN DISTINGUISH MELANOMA FROM NORMAL TISSUES

José Á. Rodríguez Corrales, Jatinder S. Josan

This chapter has been adapted from a manuscript for future submission to *Cancer Research* under the "Research article" format.

3.1. Abstract

The incidence of melanoma, the most deadly form of skin cancer, has increased throughout the world. Although melanoma has been the subject of oncogene-targeted (e.g., *BRAF* and *MEK)* therapies and immune checkpoint inhibitors (e.g., PD-1 and CTLA-4), novel therapies that improve patient survival are desirable, especially for those in the late stages of the disease. In particular, heteromultivalency represents a promising concept that relies on the selective accumulation of the ligands in the cancer cells based on the expression of a combination of surface proteins not found in any other tissues. This report represents the first study on the identification and validation of protein combinations selectively expressed in melanoma with the goal of guiding future research in heteromultivalent targeting of this disease. Gene expression profiles of melanoma biopsies and normal tissues were obtained from The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx), respectively. Python scripts, developed by the authors, screened these datasets for melanoma-selective protein targets and combinations. We successfully identified several melanoma surface receptor fingerprints with low or minimum presence in normal tissues, which we validated in established melanoma cell lines by immunofluorescence. A

detailed analysis of available literature with respect to the presence of these proteins in melanoma or other cancers and the existence of small molecule or peptide ligands (among several factors), was used to further hone the selection of potential protein targets for heteromultivalent ligands.

3.2. Introduction

Melanoma, the neoplasm generated from melanocyte transformation, is the most deadly form of skin cancer. Overall, the incidence of melanoma has continued to increase worldwide despite decades-long public health efforts promoting sun avoidance and the early detection of invasive melanoma.[1-3] In the United States alone, 91,270 melanoma diagnoses and 9,320 deaths are expected in 2018.[4] Although tumor resection during initial stages of the disease can be effective, the later stages and residual infiltrating cells compromise long-term survival. Indeed, recurrence has been observed in about 30% of stage-II patients and 50% of stage-III patients two years after resection.[5] Sadly, stage-IV melanoma patients can anticipate a 5-year survival rate of only 16%.[6]

On a positive note, clinical approaches for treating melanoma have progressed immensely during the last two decades. For example, two oncogene-targeted drugs, Vemurafenib® and Dabrafenib®, were developed and approved for human use after the discovery of a mutated form of *BRAF* in more than half of melanoma patients.[7-8] Although these drugs can decrease tumor size and improve progression-free survival by approximately five months, patients often develop resistance, thus negating their long-term efficacy.[9] Additionally, two classes of immune checkpoint inhibitors, namely Cytotoxic T-lymphocyte Associated Protein 4 (CTLA-4) and Programmed Cell Death (PD-1), are known to improve patient survival beyond one year, but can generate a severely adverse immune response in some patients given that these markers are also

present in normal tissues.[10-11] Although a small number of melanoma-specific antigens have been identified, their patient coverage is relatively low.[12] Thus, novel therapies directed towards surface targets that are present in a wider subset of melanoma patients could potentially overcome these and other vexing limitations.

Accordingly, heteromultivalency represents an innovative and promising technique that is based on the recognition of a combination of surface proteins. Heteromultivalent scaffolds contain heterologous ligands with tethered payloads (e.g., diagnostic or therapeutic agents, see Fig. 3.1A) that accumulate selectively in cells that express the cognate receptor combination, while leaving monoexpressing cells largely unaffected (see Fig. 3.1B). Proof-of-concept studies have confirmed up to an 80-fold higher affinity of a heterobivalent ligand towards transfected dual-expressing cells *in vitro*[13-14] and more than a 10-fold higher retention in induced tumors *in vivo*.[15] More modest affinity enhancements were observed in native pancreatic β-cells using a heterobivalent ligand,[16] presumably due to differences in expression intensities. Furthermore, the total number of possible targets can be increased by several orders of magnitude using protein combinations over individual proteins (see Fig. 3.1B). However, heteromultivalent ligands for the targeting of melanoma have yet to be reported.

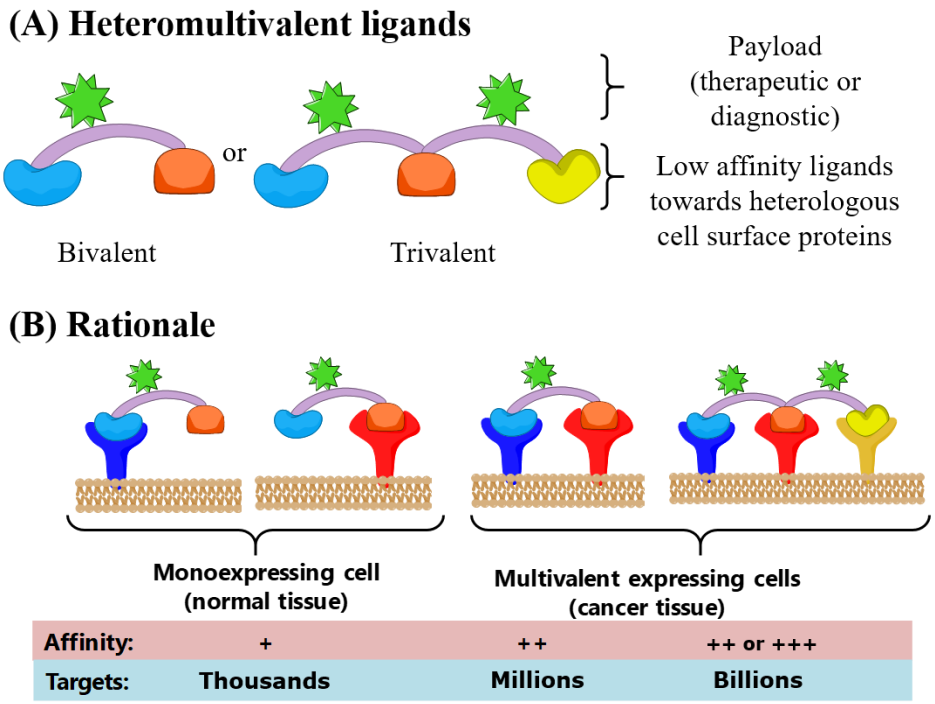


Figure 3.1. Rationale for the use of heteromultivalent ligands. Representation of heterobivalent and heterotrivalent ligands, which bear low-affinity ligands and a payload (**A**). Normal tissues that express only one of the proteins in the combination display weak affinity towards the multivalent ligand; whereas multivalent expressing cells avidly bind the ligand (**B**). Although heterotrivalent ligands have not been reported, we expect their affinity to be at least as high as that of a bivalent ligand. Furthermore, the putative number of targets increases by several orders of magnitude when using protein combinations compared to single proteins.

Herein, we report the identification and validation of surface protein combinations that could drive selective accumulation and payload delivery to melanoma. In this investigation, we retrieved both melanoma and normal tissue expression profiles for genes that encode surface proteins, derived from a combination of databases, from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) datasets, respectively. Using a threshold-based screening

method, we identified single proteins and protein combinations that are uniquely found in melanoma. We then prioritized these proteins based on selectivity, after which we surveyed the available literature for their reported expression in melanoma, extracellular domains, and known small molecule or peptide ligands. Ten of these putative target proteins were evaluated in established melanoma cell lines using immortalized keratinocytes as a control, eight of which were confirmed to display higher expression in melanoma. These eight target proteins could be used in the development of heteromultivalent scaffolds, immunotherapy, or other applications.

3.3. Experimental section

3.3.1. Materials

Penicillin/streptomycin (100x) and melanoma cell lines A-375, WM-115, and WM-266-4 were obtained from ATCC. Immortalized keratinocyte HaCaT cells were kindly provided by Dr. Lisa DeLouise from University of Rochester Medical Center. Fetal Bovine Serum (FBS), Eagle's Modified Essential Medium (EMEM), and Dulbecco's Modified Essential Medium (DMEM) were procured from Corning, and 2.5% trypsin solution was obtained from Quality Biological. Phosphate buffered saline (PBS) pills, bovine serum albumin, and ethylenediaminetetraacetic acid (EDTA) were purchased from Sigma Aldrich. A mixture of 4% paraformaldehyde in PBS was obtained from Alfa Aesar. All primary antibodies used were raised by the manufacturer in rabbit hosts towards human protein immunogens. Anti-CLN3, -CXCR3, -MFRP, -AMIGO3, -TRPM8, -LRP2, -DLL3, and -MCHR1 were obtained from Life Technologies Corporation. Anti-IL13RA2 was purchased from EMD Millipore. Anti-TRP1 was obtained from Abcam. Goat anti-Rabbit IgG (H+L) cross-adsorbed secondary antibody tagged with DyLight 550 was acquired from Life Technologies Corporation.

3.3.2. Data analysis

Data sorting, formatting, and visualization were conducted using Microsoft® Excel 2016, and Venn diagrams were generated using reference 17. Screening was performed using a Python script developed by the authors that facilitated the selection of expression thresholds (*vide infra*) and the identification of single proteins and protein combinations selectively expressed in one cohort (melanoma) compared to normal tissue cohorts. This script, along with smaller scripts for each step of the screening method, are presented in Appendix A.

3.3.3. Defining the surfaceome

Five databases that report protein subcellular localization were consulted and used to restrict protein screening to those present in the cytosolic membrane, henceforth known as the surfaceome. Briefly, The Human Protein Atlas (THPA),[18-19] the Cell Surface Protein Atlas (CSPA),[20] and the Human Protein Reference Database (HPRD)[21-22] were derived solely from experimental techniques. In contrast, the surfaceome database (SDB) relies on structure-based predictive algorithms to identify membrane-embedded proteins,[23-25] while the Universal Protein Resource (UniProt) utilizes a combination of experimental evidence and predictive algorithms.[26-27] A more detailed description of each database, along with the corresponding search terms utilized and other relevant information, are provided in Table S3.1.

Three different surfaceomes were generated for screening purposes using the five aforementioned databases. The “comprehensive” surfaceome contains the combined entries from the five databases, whereas the “experimental surfaceome” utilizes only those based on experimental reports (HPRD, THPA, and CSPA). In contrast, the “curated” surfaceome contains only proteins present in at least two different databases. The three surfaceomes were compared to

each other and to that used in the work that inspired this study.[28] Further biostatistical analysis was performed for the comprehensive, experimental, and curated surfaceomes independently, after which the putative targets from each screening were compared. The list of proteins that constitute each surfaceome is presented in Appendix B.

3.3.4. Gene expression analysis

Gene expression data for the TCGA melanoma cohort, which includes 474 patients from all four melanoma stages (see Fig. S3.2), and GTEx normal tissues were obtained using the Xena Browser.[29] Expression values, which correspond to $\log_2$-transformed RSEM-normalized counts, were visualized and downloaded without on-the-fly normalization. Genes with no available expression data (see Table S3.2) and GTEx samples where less than 5% of the total genes were detected (i.e., normalized counts >0) were excluded from further screening. Three normal tissue subsets, hereforth "ensembles", were generated by random selection of 10 samples per tissue type in order to prevent tissue bias based on sample number (i.e., GTEx samples are non-homogeneously distributed across tissue types, see Table S3.3). In the case of bladder and fallopian tube tissue samples (with less than 10 samples each), all were selected. In total, each ensemble was comprised of 304 samples across 31 tissue types.

Proteins were considered "expressed" when the gene expression value for a given sample surpassed a particular cutoff, as previously described.[28, 30] Threshold values and minimum cohort coverages for patient and normal tissues were empirically optimized to restrict the number of identified targets as recommended in the references above. Threshold limits were defined based on the expression values for all surface genes in each cohort; the targets with high coverage in

melanoma, and low coverage in at least two of the three normal ensembles, were used in the next step.

All possible combinations comprised of either two or three different targets, henceforth referred to as bivalent and trivalent combinations, were generated and evaluated in melanoma and normal tissues. A bivalent combination was considered "coexpressed" in a given sample if the expression levels of both proteins were higher than the corresponding thresholds. In the case of trivalent combinations, two possible expression patterns were investigated: trivalent coexpression, where all three genes were expressed, and bivalent coexpression, where at least two of the three genes were expressed (*vide infra*). In the case of normal tissue samples, the reported coverage correspond to the average of the three ensembles. Coverage values were calculated using the thresholds for the curated surfaceome, unless specified otherwise.

3.3.5. Literature validation

Single-protein targets were ranked by selectivity (i.e., differential coverage towards melanoma versus normal tissues). In order to be exhaustive, an algorithm was developed to evaluate the likelihood of each identified protein being part of a melanoma-selective bivalent or trivalent combination (see Section 3.7.1), which was comparable to that obtained from selectivity alone (see Fig. S3.1). Starting with the highest-ranked single-protein targets, a literature search was conducted to identify reports of (a) subcellular localization, (b) differential expression in normal versus cancer tissues, (c) involvement in signaling pathways or cancer-relevant roles (e.g., proliferation or cell adhesion), and (d) existence of selective peptides or small-molecule ligands. We paid particular attention to reports on ligand functionalization that did not hinder target binding, in order to guide their potential incorporation into heteromultivalent architectures.

3.3.6. Cell lines and culture conditions

All experiments involving biohazardous materials were conducted under Biosafety Level 2 containment and approved by the Virginia Tech Institutional Biosafety Committee (IBC protocol 14-043). Cell lines were grown under the specific conditions recommended by the supplier. Briefly, cells were incubated in complete medium (DMEM for A-375 and HaCaT, or EMEM for WM-266-4 and WM-115, supplemented with 10% FBS and 1X penicillin/streptomycin) at 37°C in a 5% CO_2 aerobic atmosphere. Cells were subcultured at 80-90% confluency by using 0.25% trypsin/0.53 mM EDTA for 5 min, except for HaCaT which required preincubation with 0.05% EDTA in phosphate buffer saline (PBS) for 15 min, followed by treatment with 0.05% trypsin/0.025% EDTA in PBS for 1-5 min. Cells were counted in an InCyto C-Chip improved Neubauer hemocytometer and seeded into T75 flasks for general culture. Cells were used before passage 12 to decrease phenotypic drift.

3.3.7. Immunocytochemistry

Cells were grown as described above, seeded at 10,000 cells/well in a 96-well plate, and allowed to attach for 24 h. Cells were fixed with 4% paraformaldehyde in PBS for 15 min at room temperature, followed by overnight blocking with 2% bovine serum albumin (BSA) in PBS. Antibodies were diluted in BSA/PBS at a concentration at least equal to that recommended by the supplier (see Table S3.4). Cells were incubated with diluted primary antibody for 1 h, rinsed with BSA/PBS (3x), incubated with fluorophore-tagged secondary antibody for 1 h, rinsed with PBS (3x), and incubated with 2 μg/mL Hoechst 33342 in PBS for 15 min for DNA counterstaining, followed by a PBS rinse. Negative controls were treated under the same conditions, except for

being incubated with BSA/PBS instead of primary antibodies. Cells were imaged using a Biotek Cytation3 Cell Imaging Multi-Mode Reader. Conditions for image acquisition (light intensity, integration time, and camera gain) and processing (contrast and brightness) were adjusted to ensure signal output in stained samples and undetectable fluorescence in negative controls. Imaging and image processing conditions were kept constant across cell lines and samples.

3.4. Results

3.4.1. Defining the protein surfaceome based on available databases

The five analyzed databases displayed high heterogeneity with respect to the proteins defined as membrane-bound, as shown in Fig. 3.1A. We expect the "curated surfaceome," which excludes any proteins listed in only one of the databases, to yield fewer false positives than either of the individual databases or the other surfaceomes. However, this curated list might also exclude proteins with limited experimental evidence (i.e., potential false negatives), or those where predictive algorithms are less accurate (e.g., proteins only shuttled to the membrane when mutated). In turn, the "experimental surfaceome" could potentially include false positives due to the methodological limitations of the methods used to identify subcellular localization, including contamination of the membrane fraction upon isolation or non-specific binding of antibodies. Unlike the curated counterpart, the experimental surfaceome is not affected by structure-driven algorithms that might display limited accuracy in predicting subcellular localization.[23]

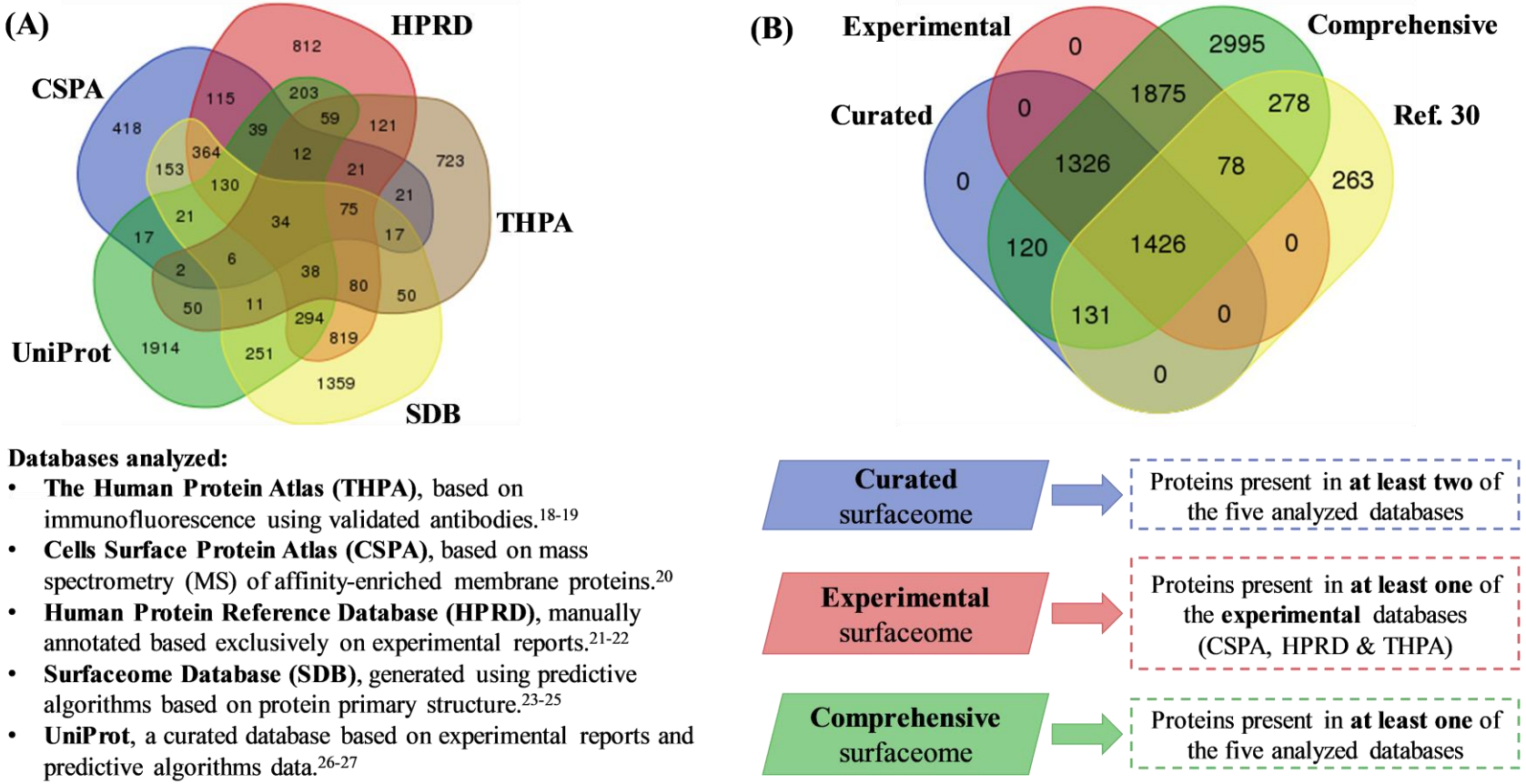
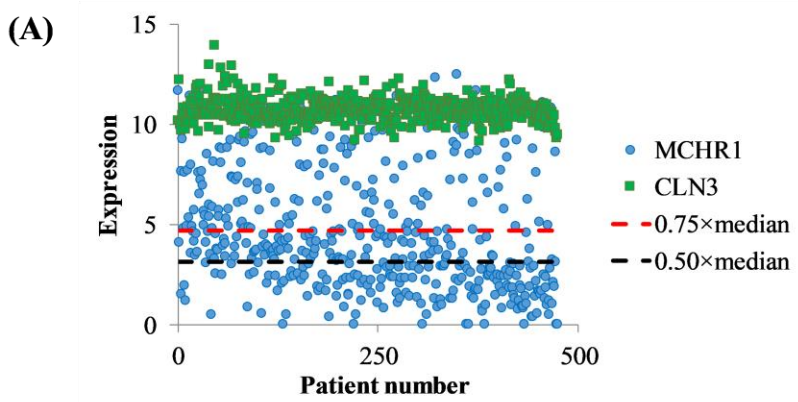


Figure 3.2. Defining the surfaceome. Comparison of the five databases used to derive the three surfaceomes (comprehensive, experimental and curated) used herein (**A**), and their comparison (**B**) with that defined in reference 30.

We compared the curated, experimental and comprehensive surfaceomes (i.e., the combination of all proteins listed in any of the five databases) among themselves and with respect to that used in the nominal work that inspired this study (see Fig., 3.2C).[30] The latter derived a surfaceome using the Cancer Genome Anatomy Project Gene Ontology browser as the main source of annotations on protein subcellular localization, followed by analysis in other databases (e.g., UniProt and Entrez), and instrumental capabilities. We aimed to improve on the definition of the surfaceome as to avoid excluding putative targets in subsequent screening steps, while also consulting databases that were released after the nominal publication. As presented in Fig. 3.2C, most of the genes used in the nominal work, namely 1,913 of 2,176, or 88%, were also present in the comprehensive surfaceome, whereas 6,316 new genes were included. Thus, the comprehensive surfaceome successfully expanded the scope of this screening method.

3.4.2. Identification of single targets based on gene expression levels

As noted, the screening method utilized herein considers proteins to be "expressed" when gene expression surpasses a given threshold, which is empirically optimized to obtain a manageable number of putative targets (see Fig. 3.3A).[30] This nominal work suggests a threshold of 55-75% of the median for cancer cohorts, along with a minimum patient coverage (i.e., the minimum number of patient samples that express the protein). Using the TCGA melanoma cohort, we tested the most stringent threshold limit recommend (75% of the median) along with more restrictive thresholds and coverage values (see Fig. 3.3B). Although the comprehensive, experimental, and curated surfaceomes were screened independently (see Figs. S3.3 & S3.4), only the curated surfaceome will be discussed in the main text for clarity.



(B) ***Melanoma***

Threshold	Patient coverage (%)	Hits
0.75×median	> 0	2,670
	≥ 10	***2,163***
1×median	> 0	2,568
	≥ 10	1,957
2×median	> 0	1,211
	≥ 10	400

(C) ***Normal tissues***

Threshold	Normal tissue coverage (%)	Hits
0.65×median	0	2
	≤ 10	29
	≤ 25	***114***
0.55×median	≤ 10	14
	≤ 25	67
0.45×median	≤ 10	8
	≤ 25	40

(D) ***Differential coverage***

Protein	Melanoma coverage (%)	Normal tissue coverage (%)	Difference (%)*
MFRP	99.8	1.2	98.6
GJD3	83.1	0.6	82.5
SLC45A2	89.4	8.9	80.5
CLN3	100.0	20.2	79.8
TMEFF1	93.5	14.0	79.5
GPR19	94.7	16.2	78.5

* Difference in coverage is calculated using the melanoma cohort and the average coverage from three normal tissue ensembles.

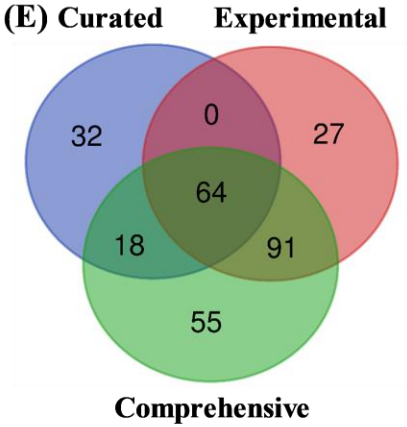


Figure 3.3. Single target screening. Three empirically optimized thresholds (100%, 75% and 50% of the median expression value for the melanoma cohort shown as dashed lines) are shown, along with patient gene expression values for two sample proteins, CLN3 and MCHR1. Only patients where gene expression were above the threshold were counted as "covered" (**A**). Upon testing several thresholds and minimum melanoma patient coverages (**B**), 2,163 proteins (bold italics) were carried to the next step. Thresholds and maximum normal tissue coverage values were optimized using three different ensembles that contain 31 tissue types and 304 samples

each (**C**). As a result, 114 proteins were selected for further study (bold italics), the most selective of which are shown in (**D**). (**E**) compares the single targets obtained in the screenings using the curated, experimental and comprehensive surfaceomes.

For this particular step, we sought to exclude only those genes with weak expression in melanoma. Thus, we selected a threshold of 0.75×median and a minimum patient coverage of 10%, thus generating 2,163 hits, which corresponds to approx. 74% of all genes from the curated surfaceome with available expression data at TCGA and GTEx. Unlike methods that rely on protein overexpression, heteromultivalency drives selective accumulation in cells that coexpress a particular protein combination.[13-14] Therefore, we avoided thresholds at or beyond the median of the cancer array to maintain a useful number of hits. Furthermore, we observed that the following screening step, which involved normal tissues (*vide infra*), readily removed putative targets.

It has been proposed that expression thresholds for normal tissues should be more stringent than those for cancer in order to identify only those genes that display differential expression patterns (i.e., to minimize false positives). In particular, a threshold of 35-55% of the median for normal tissue samples and low coverage values have been recommended.[28] Given that most targets were excluded when we used such a strict threshold (see Fig. 3.3C); we relaxed this value to 65% of the median and a maximum coverage of 25%. The need to widen this threshold might be due to the higher diversity of tissue types and large sample number analyzed herein, which increases the likelihood of finding normal tissue samples with high expression. These screening conditions generated a manageable number of putative targets. Using the curated, experimental, and comprehensive surfaceomes, we identified 114, 182 and 228 single targets, respectively, which corresponds to 3.9%, 4.1%, and 3.8% of all genes with TCGA and GTEx data of each surfaceome). The results for the last two surfaceomes are presented in Figs. S3.3B & S3.4B.

Some of the identified targets displayed striking selectivity towards melanoma, as shown in Fig. 3.3D. For instance, 99.8% of the patients expressed the membrane frizzled-related protein (MFRP), compared to only 1.2% of the normal tissues based on the curated surfaceome screening.

The level of selectivity towards melanoma (i.e., the difference between melanoma and normal tissue coverage) varied greatly across proteins, where only 16 of the 114 identified genes displayed a value greater than 50%. This result is not surprising given that previous studies have identified a rather limited number of melanoma-relevant antigens.[12] Additionally, this selectivity highlights one of the potential advantages of heteromultivalent targeting: the discrimination against normal tissues that express a particular protein at levels comparable to those found in cancer, provided that coexpression with another protein only occurs in the latter.

Furthermore, the screenings using the three derived surfaceomes identified some common targets (see Fig. 3.3E). In particular, 64 of the 228 putative targets obtained using the comprehensive surfaceome were also identified in the corresponding screenings using the curated and the experimental surfaceomes. The curated surfaceome, which is the smallest of the three, led to the identification of 82 targets also found in the comprehensive surfaceome, namely 36% of those obtained using the latter. Thus, the curated dataset might be a valuable alternative for less-computationally intensive screenings.

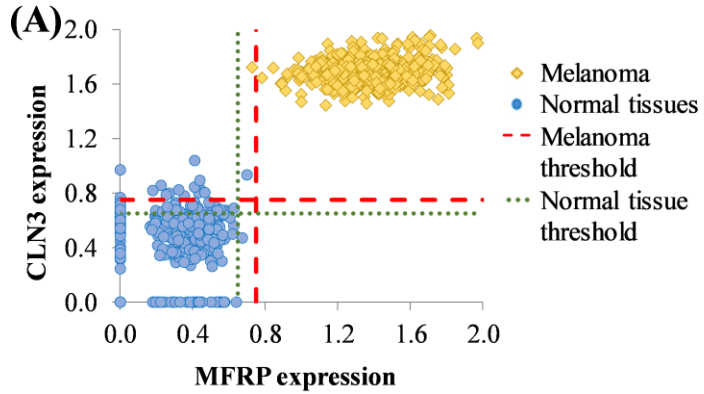
3.4.3. Identification of genes selectively coexpressed in melanoma compared to normal tissues

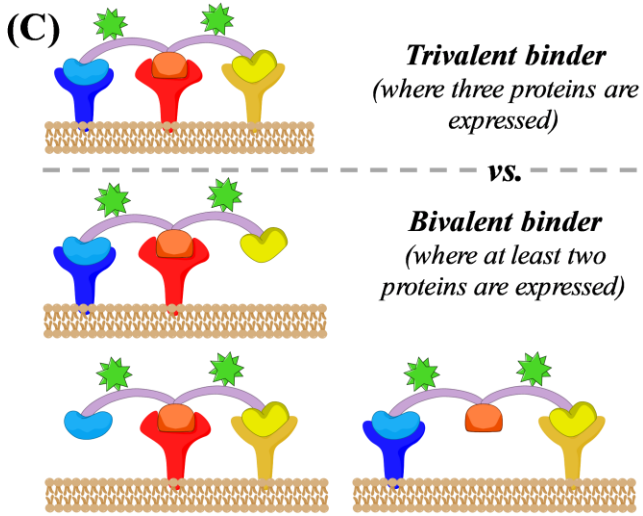
Using the thresholds and single targets identified above, we evaluated the coverage of protein combinations in both cohorts. In this coexpression analysis, patient and normal tissue samples were considered "covered" by a combination when the expression values of more than one gene was found to be above the threshold (see Fig. 3.4A). We identified several bivalent and trivalent protein combinations with excellent coverage in melanoma tissues and minimum presence in normal samples, some of which are shown in Fig. 3.4B&D. For instance, MFRP and CLN3 (Ceroid-Lipofuscinosis Neuronal 3 or battenin) were coexpressed in 98.7% of the

melanoma patients in comparison to only 0.4% in normal tissues, whereas CLN3 alone was expressed in 20.2% of the normal tissues (see Fig. 3.3D). These findings support the hypothesis that selective targeting of melanoma might be accomplished by using protein combinations (e.g., MFRP & CLN3) rather than individual proteins, given that some normal tissues might express one particular target at levels comparable to those found in cancer.

Furthermore, we explored the potential use of trivalent protein combinations to expand the targeting capabilities of heteromultivalent scaffolds. Although trivalent scaffolds have not been reported, we expect that a bivalent binding mode (i.e., the binding of two ligands, see Fig. 3.4C) might be sufficient to drive tissue selectivity, given that heterobivalent ligands effectively discriminate between mono- and dual-expressing cells and tumors.[13-15] Based on the fact that heterotrivalent ligands can undergo three different bivalent binding patterns (see Fig. 3.4C), patient coverage could be expanded through the selection of appropriate bivalent subsets. Conversely, trivalent ligands should be designed such that neither trivalent nor bivalent binding occurs in normal tissues.

Our analysis of patient and normal tissue coverage of trivalent combinations showed striking results. For instance, the combination of CLN3, GPR19, and MFRP showed impressive coverage at both the three-gene level (94.5%) and the two-gene level (virtually complete patient coverage). In contrast, only a minimum number of normal tissue samples displayed trivalent (0.1%) or bivalent (4.8%) protein expression. Several other combinations showed similar levels of selectivity towards melanoma, some of which are presented in Fig. 3.4D. These results support the existence of trivalent protein combinations expressed in melanoma that are absent from normal tissues, which could be used in expression profile-driven targeting.





(B) Bivalent combinations

Combinations	Melanoma coverage (% patients)	Normal tissue coverage (% samples)	Difference (%)
CLN3 & MFRP	99.8	0.4	99.4
GPR19 & MFRP	94.5	0.2	94.3
MFRP & TMEFF1	93.2	0.1	93.1
CLN3 & GPR19	94.7	4.4	90.3
HLA-G & MFRP	90.3	0.4	89.9
MFRP & SLC45A2	89.4	0.0	89.4

(D) Trivalent combinations

Combination	Melanoma coverage (% patients)		Normal tissue coverage (% samples)	
	Three proteins	At least two proteins	Three proteins	At least two proteins
CLN3, GPR19 & MFRP	94.5	100.0	0.1	4.8
CLN3, MFRP & TMEFF1	93.2	100.0	0.1	5.7
CLN3, HLA-G & MFRP	90.3	100.0	0.0	6.0
CLN3, MFRP & SLC45A2	89.4	99.8	0.0	4.1
GPR19, MFRP & TMEFF1	88.4	99.6	0.1	6.7
GPR19, MFRP & SLC45A2	86.5	97.5	0.0	3.3
CLN3, GPR19 & TMEFF1	88.6	99.6	3.1	10.2

Figure 3.4. Evaluation of gene combinations in melanoma and normal tissues using the curated surfaceome. Sample diagram showing expression of MFRP and CLN3 (**A**) in the TCGA melanoma cohort (each yellow diamond represents one patient) or GTEx normal tissue ensemble 1 (each blue circle stands for one tissue sample), and the corresponding thresholds used to define expression. For clarity, all expression values were normalized to the median of the array. We identified several bivalent (**B**) and trivalent (**D**) gene

gene combinations selectively expressed in melanoma compared to normal tissues. In the case of trivalent combinations, we report the number of patients where three or at least two genes are expressed, given that a trivalent ligand could bind bivalently or trivalently (**C**).

3.4.4. Literature validation

We performed a literature survey of the targets identified in the curated, experimental, and comprehensive screenings prior to experimental validation. These putative targets include established melanoma biomarkers (e.g., *MLANA*/MART-1 and TYR)[31] and proteins involved in melanocyte pigmentation (e.g., SLC45A2, SLC24A5, and TYR)[32-33] or cancer-relevant biological roles (e.g., CLN3 in proliferation and apoptosis)[34-35]. Additionally, some targets are overexpressed in melanoma (e.g., SLC24A5)[32] or have been investigated as potential therapeutic targets for this disease (e.g., TYRP1).[36-37] These results suggest that the screening method successfully identified proteins with potential relevance for melanoma targeting, including putative targets with limited or no prior reports of involvement in melanoma, as discussed below.

However, some of the identified proteins were not suitable for heteromultivalency. Given that heteromultivalent ligands rely on protein crosslinking at the cell surface, proteins exclusively expressed in the cytoplasm or nucleus were deemed incompatible with our approach and thus were not pursued further. Similarly, immune biomarkers were excluded, as their presence can be attributed to the infiltrated immune cells in TCGA patient biopsies.[38] Specific examples of targets that were eliminated from further analysis are summarized in Table S3.5.

After an extensive literature search for the parameters described in the method, which including subcellular localization, involvement in cancer, and expression levels, we selected 10 proteins for immunocytochemical validation. These include targets with a) no prior reports of melanoma involvement (MFRP, CLN3 and AMIGO3), b) prior evidence of expression in melanoma but no known synthetic ligands (TYRP1, DLL3 and LRP2), and c) known small molecule or peptide ligands (CXCR3, TRPM8, IL13RA2 and MCHR1). The specific reasons for the selection of each protein target are summarized in Table S3.6. Notably, TYRP1 is a target with

proven clinical relevance in melanoma,[36-37] and thus served as putative target and positive control in this study.

3.4.5. Immunocytochemistry

The results obtained from bioinformatics screening required experimental validation. Although RNAseq is a widely used technique due to its sensitivity and high-throughput character, the correlation between mRNA and protein levels varies across genes.[39-41] Furthermore, the affinity of heteromultivalent ligands depends on the membrane fraction of the protein rather than its total expression in the cell.[14] Even though the transcriptome-driven identification of surface protein combinations has been successfully accomplished in prostate cancer, results must be further validated due to possible discrepancies in protein expression levels on the cell surface.[30, 39]

Thus, we performed immunocytochemistry in established melanoma cell lines WM-266-4, WM-115, and A-375, using immortalized HaCaT keratinocytes as a control. Cell lines were treated with the same concentrations of antibodies and dyes, and imaged under the same conditions to allow for semiquantitative comparison of the expression level of a given protein across cell lines. All proteins were positively identified in the melanoma cell lines and/or the keratinocytes control, whereas some displayed enhanced expression in the former. Selected targets are presented in Fig. 3.5, while the remainder are available in Fig. S3.5. The combined results from gene- and protein-level analysis are summarized in Table 3.1.

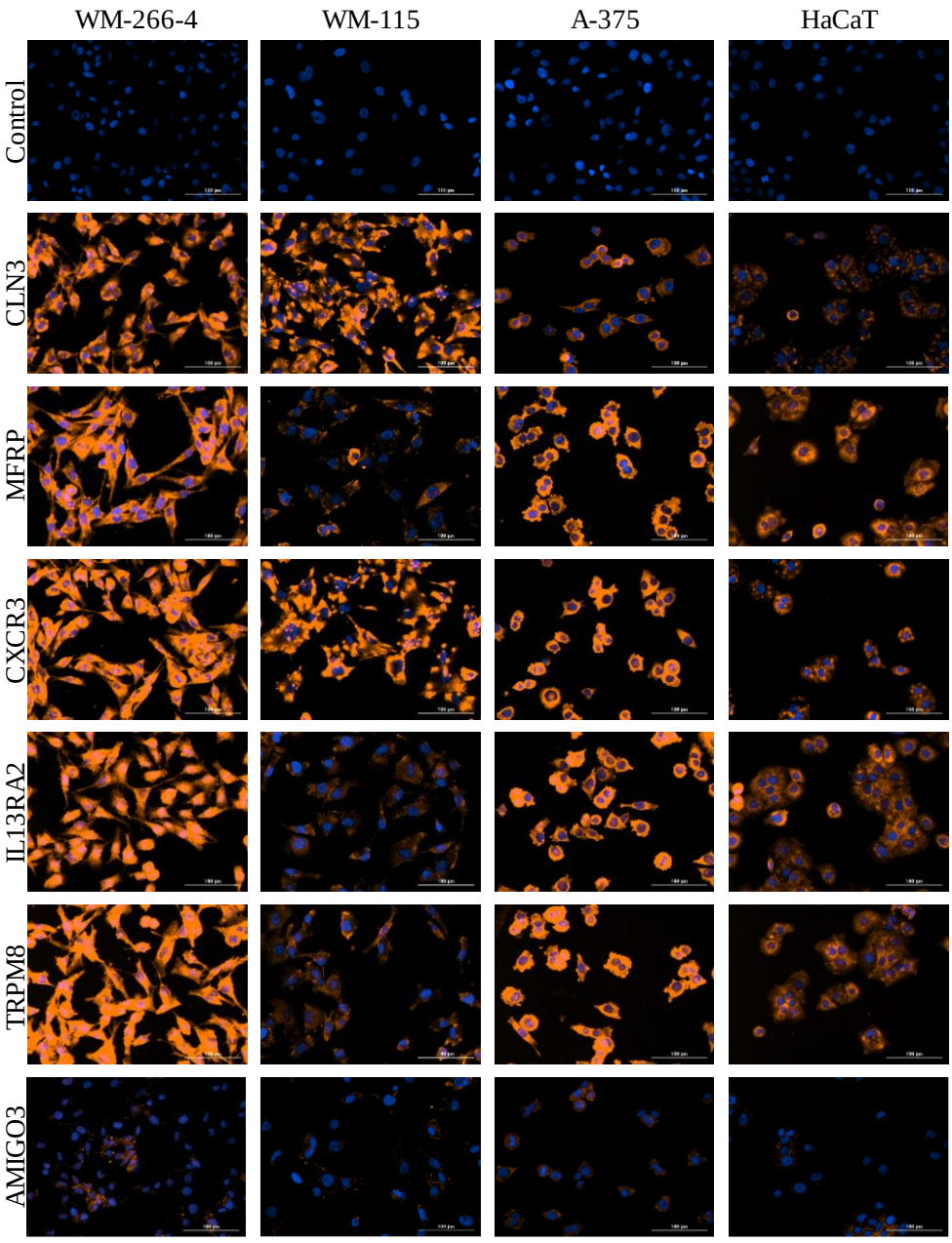


Figure 3.5. Validation of protein expression in melanoma cell lines WM-266-4, WM-115, and A-375, using immortalized keratinocyte HaCaT as a control. All size bars correspond to 100 μm. Cells were fixed but not deliberately permeabilized. Secondary antibodies were tagged with DyLight 550 and fluorescence was collected in the RFP channel (shown in orange). The nucleus was stained with blue fluorescent dye Hoechst 33342, while the controls (first row) were treated with a secondary antibody and nuclear dye only.

Table 3.1. Transcriptome and proteome level coverage for selected combinations.

Combination	Gene level screening*		Protein level‡			
	Melanoma coverage (%)	Normal tissue coverage (%)	WM-266-4	WM-115	A-375	HaCaT
CLN3 & CXCR3	72.6	5.7	H/H	H/H	M/H	L/L
CLN3 & MFRP	99.8	0.2	H/H	H/L	M/H	L/M
CLN3 & TRPM8	55.5	2.9	H/H	H/L	M/H	L/M
CXCR3 & DLL3	64.6	0.9	H/H	H/M	H/M	L/L
CXCR3 & IL13RA2	32.7	3.1	H/H	H/L	H/H	L/M
CXCR3 & TRPM8	39.7	1.2	H/H	H/L	H/H	L/M
DLL3 & IL13RA2	42.2	8.0	H/H	M/L	M/H	L/M
DLL3 & MFRP	90.9	0.4	H/H	M/L	M/H	L/M
DLL3 & TRPM8	50.0	4.5	H/H	M/L	M/H	L/M
IL13RA2 & TRPM8	31.9	8.8	H/H	L/L	H/H	M/M
CLN3, CXCR3 & MFRP	61.2 & 99.2	0.0 & 1.2	H/H/H	H/H/L	M/H/H	L/L/M
CLN3, CXCR3 & TRPM8	26.4 & 79.5	0.0 & 2.9	H/H/H	H/H/L	M/H/H	L/L/M
CLN3, CXCR3 & IL13RA2	25.1 & 78.1	0.2 & 3.8	H/H/H	H/H/L	M/H/H	L/L/M

* Calculated using the aforementioned thresholds from the curated surfaceome. Coverage corresponds to bivalent expression in bivalent combinations, or trivalent followed by bivalent expression in trivalent combinations.

‡ Expression level is represented as “H” for high, “M” for medium” and “L” for low. These values are relative to each protein (e.g., “high” corresponds to the cell line with the highest expression of those tested). Expression levels for each protein are separated by a slash (“/”).

The analyzed protein targets displayed different expression patterns. Specifically, CLN3 and CXCR3 were more intensely expressed in the melanoma cell lines compared to the HaCaT control. In contrast, MFRP, IL13RA2, TRPM8, DLL3, MCHR1, and AMIGO3 were more intensely expressed in the metastatic cell lines WM-266-4 and A-375 compared to HaCaT, whereas expression in WM-115 was observed to be lower or comparable to HaCaT. Although WM-266-4 and WM-115 originated from the same patient,[42] our results agree with prior studies that showed different expression profiles for the two cell lines.[43-44] Additionally, staining patterns for TYRP1 and LRP2 proved to be less conclusive, suggesting that these targets are not as selective towards the melanoma cell lines as those discussed above (see Fig. S3.5). Given that TYRP1 has proven melanoma-relevance,[36-37] other validated targets with higher selectivity might be promising for further studies.

3.5. Discussion

An accurate description of the human surfaceome would prove beneficial in several areas of research. From a fundamental perspective, it could improve our understanding of signaling pathways that involve binding events at the cell surface, along with protein-protein interactions between membrane receptors. Furthermore, such knowledge could drive the further discovery of surface targets involved in diagnosing and treating cancer and other diseases, as well as their subsequent use in a range of therapeutic applications, including receptor-promoted drug uptake agents, regulation of signaling pathways, or selective cell imaging based on specific expression profiles.

In exploring the surfaceomes used in this investigation, we also sought to reconcile conflicting reports from the literature. We observed a high heterogeneity between the five

surfaceomes obtained from the literature (CSPA, THPA, HPRD, SDB, and UniProt), which we attribute to inherent method and sample limitations. For instance, immunofluorescence results can be affected by non-specific binding, fluorophore stability, and sample integrity; in contrast, the subcellular localization of a given protein can vary across tissues (e.g., EN2 is a nuclear protein with no detectable cytosol or membrane expression in melanoma,[19] but is present in the caveolae of rat cerebellum tissues[45]). Moreover, algorithms that predict membrane localization might have limited accuracy in the presence of unknown sorting motifs.[46]

The derived surfaceomes also present distinct advantages and limitations. In the case of the curated surfaceome, targets that were only present in one of the five databases were excluded, which likely led to a decrease in false positives. Given its smaller size compared to the other surfaceomes, namely 3,003 vs. 4,705 and 8,229 proteins for experimental and comprehensive, respectively, it is also the least computationally intensive, and thus it could be beneficial in those applications involving a large amount of data. In contrast, the comprehensive surfaceome is the largest and most exhaustive of the three, thus preventing the exclusion of potentially relevant targets, even if these might only be present on the cell surface of limited cell types.

By utilizing the transcriptome-based screening approach described herein, we were able to identify a number of melanoma-relevant proteins. These include established prognosis or diagnosis biomarkers, therapeutic targets, certain proteins overexpressed in melanoma, and three proteins with limited or non-existent reports in cancer. For instance, AMIGO3, which is a novel target with a widely unknown biological function,[47] and might share a similar role to the homologous protein AMIGO2, which was recently found to be overexpressed in liver metastases and involved in apoptosis and angiogenesis.[48, 49] We identified a higher expression of AMIGO3 in metastatic cell lines WM-266-4 and A-375 compared to WM-115 and HaCaT (see Fig. 3.5). The same expression

pattern was observed for MFRP, a protein expressed in the retina, but with no prior reports in cancer,[50] and MCHR1, a protein with several known ligands and one report of expression in melanoma.[51-54]

A rather interesting finding of our screening was the identification of CXCR3 rather than its more studied homologous, CXCR4. The latter is overexpressed in several types of cancer due to its involvement in cell migration, adhesion, and metastasis, and is known to serve as a prognosis marker and therapeutic target.[55-57] In the case of melanoma, CXCR3 has a wider patient coverage than CXCR4,[58] and a pivotal role in progression and metastasis to lymph nodes.[59-60] Interestingly, the screening showed that CXCR4 was expressed in several normal tissues (unlike CXCR3), which prompted its exclusion. Our immunofluorescence results confirm higher CXCR3 expression in the three melanoma cell lines (WM-266-4, WM-115, and A-375) compared to HaCaT keratinocytes (see Fig. 3.5). These factors, coupled with the existence of several ligands for this protein,[61-64] reinforce its appeal for further studies in heteromultivalency.

Furthermore, we identified several melanoma-specific combinations at the gene and protein levels. Given that 8 of the analyzed proteins showed differential expression with respect to HaCaT controls, the corresponding 28 bivalent and 56 trivalent combinations would virtually display selectivity towards at least one of the melanoma cell lines. In particular, the combination of CLN3 & MFRP evidenced some of the highest coverage results in the transcriptome screening (99.8% in melanoma and 0.2% in normal tissues), as well as higher protein expression in the two metastatic melanoma cell lines compared to the HaCaT control (see Table 3.1). Similarly, CLN3 and CXCR3 were more strongly expressed in the three melanoma cell lines compared to HaCaT, and displayed significant selectivity in the transcriptome level, namely 72.6% vs. 5.7% for melanoma and normal tissues, respectively.

We were also able to identify trivalent combinations of the validated protein targets. For the particular combination of CLN3, CXCR3 & MFRP, we observed wide melanoma patient coverage at the transcriptome level, namely 61.2% and 99.2% for trivalent and bivalent binding, respectively, as well as a similarly high coverage of the melanoma cell lines based on immunofluorescence. WM-266-4 displayed high expression of the three proteins, whereas the other two melanoma cell lines expressed two of the proteins at high intensity (*i.e.*, CLN3 & CXCR3 for WM-115, and CXCR3 & MFRP in A-375); finally, HaCaT only showed one of these proteins at a medium level. These findings support the potential efficacy of using trivalent combinations to improve patient coverage by relying on the coexpression of at least two proteins in the target tissues (*vide supra*). Other combinations with similar expression patterns are presented in Table 3.1.

Although most studies in heteromultivalency utilize proteins as binding anchors to control affinity and selectivity,[13-15] such scaffolds have other potential applications. In particular, these binding events can either suppress or activate signaling pathways; therefore, heterobivalent ligands can regulate the cellular processes associated with the cognate protein combinations.[16, 65-66] For instance, a heterobivalent scaffold bearing a CXCR3 small molecule antagonist[61-63] and an IL13RA2 peptide ligand (e.g., Pep-1)[67] could potentially inhibit progression and invasiveness, which involve the first protein,[60] and lead to IL13RA2-based internalization of the scaffold along with any covalently attached payloads (e.g., drugs or imaging agents).

Importantly, the identified proteins and protein combinations have potential applications beyond heteromultivalency-based targeting. Current immunotherapy strategies that target immune markers expressed in normal tissues could benefit from epitopes selectively expressed in melanoma, including those reported herein. In fact, IL13RA2 has been identified as a putative

antigen for melanoma immunotherapy.[12] Possible approaches include antibody-driven delivery of cytotoxic agents (e.g., antibody-drug conjugates) and immune cell recruitment to the surface of melanoma tissues through bispecific antibodies.[68] Alternatively, drugs intended for these targets could be administered by themselves, or in combination with immunotherapy, to further drive tumor clearance.

Our future studies will focus on the validation of protein coexpression on the cell surface of established melanoma cell lines and patient tissue samples. Preliminary experiments using confocal microscopy to evaluate subcellular distribution of one of the proteins have been inconclusive due to non-deliberate cell permeabilization, resulting in cytoplasmic staining (data not shown). Alternatively, we will explore the use of flow cytometry coupled to epifluorescent microscopy in live cells to avoid fixation. We will also expand the number and types of tissue samples, for both cancer and normal tissues, by using a tissue microarray or a similar setup. Finally, we aim to synthesize targeting scaffolds bearing imaging or therapeutic payloads and combinations of heterologous ligands, ligands and antibodies, or bispecific antibodies towards a melanoma-specific combination.

3.6. Acknowledgments

This work was supported in part by funding from the Department of Chemistry at Virginia Tech (grant number 119744) and VT-ICTAS (grant number 175911). We would like to thank Dr. Carla Finkielstein and Dr. John L. Robertson for their helpful discussions throughout the project and especially during target selection.

3.7. Supporting information

3.7.1. Target ranking based on likelihood of generating selective combinations

After the identification of putative targets and combinations (section 3.3.4), single targets were ranked based on selectivity (i.e., the difference between melanoma and normal tissue coverage). Bivalent combinations were then sorted by selectivity, and trivalent combinations were sorted by trivalent selectivity (i.e., calculated for the three genes) followed by bivalent selectivity (i.e., at least two of the genes in the combination). Combinations with selectivity lower than 10% were excluded because they were no longer considered melanoma-specific.

A weighed ranking based on the likelihood of being able to generate selective combinations was calculated. First, the ranking of the "n" most-selective bivalent combinations that included a particular gene was added together to obtain a total score. Then, genes were ranked based on total score, where lowest values corresponded to best ranking. This process was repeated using the trivalent combinations and averaged with that obtained from bivalent combinations to generate the final weighed rank. Several values of "n" were tested and the resulting rankings were compared to those obtained based on the single target screening alone (section 3.3.4). As shown in Fig. S3.1, high values of "n" (e.g., 50 or 100) significantly skewed the weighed rank for the targets, whereas low "n" values (1-30) revealed ranks comparable to those obtained from the single target ("st") screening.

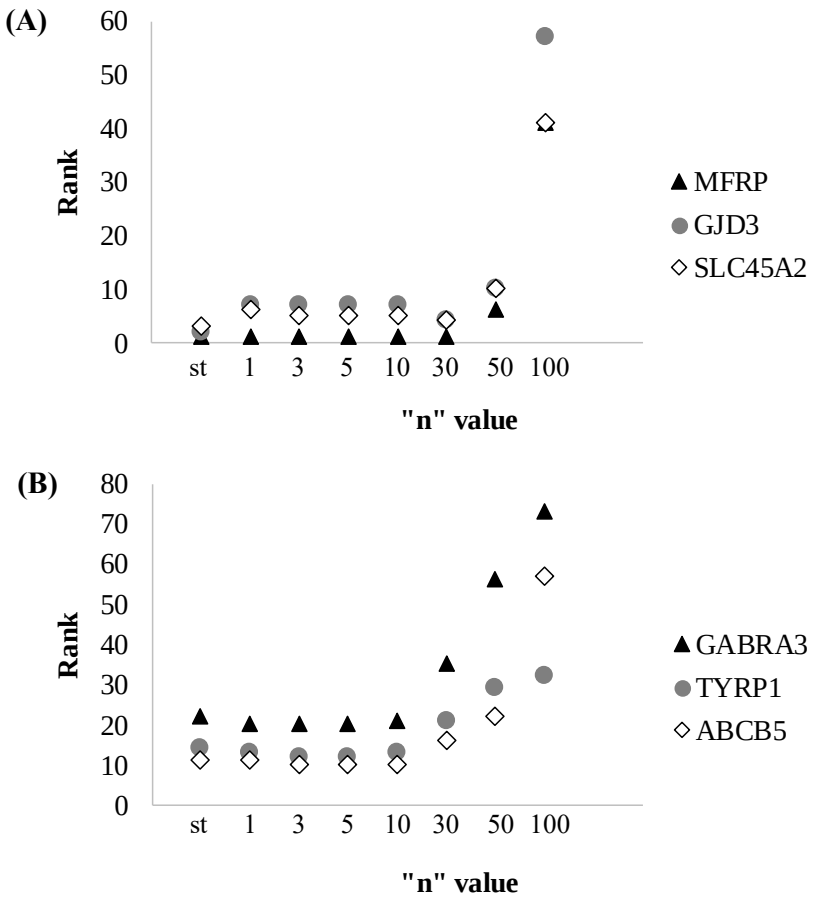


Figure S3.1. Weighed ranked calculated for some of the best proteins based on curated screening using different values of "n", and compared to rank from single target (st) screening. Selected proteins were separated into two graphs (**A** & **B**) for clarity.

3.7.2. Databases used to define the different surfaceomes

Table S3.1. Description of the databases used to derive the human surfaceome

Database	Method used to determine subcellular localization	Terms used for retrieval of surface proteins	Number of surface proteins in database	Version	References
THPA	Immunofluorescence using validated antibodies	Identified on the cell membrane using antibodies with "supported" or "approved" reliability only	1,320	13	18-19
CSPA	Tagging of extracellular proteins followed by affinity enrichment and mass spectrometry	Proteins identified in the cell surface of human tissues and cell lines	1,445	-	20
HPRD	Manual annotation based exclusively on experimental reports	GO:0016327 (apical membrane), GO:0016323 (basolateral membrane), GO:0005903 (brush border), GO:0009986 (cell surface), GO:0005925 (focal adhesion), GO:0005887	3,216	Release 9, last updated on 04-13-10	21-22

		(integral to plasma membrane) and GO:0005886 (plasma membrane)			
SDB	Predictive algorithms based on protein primary structure	Identified as surface protein	3,702	-	23-25
UniProt	Manual annotation based on available literature and predictive algorithms	Proteins located in the cell membrane	3,081	04-10-16	26-27

3.7.3. Stage distribution among patient samples in TCGA

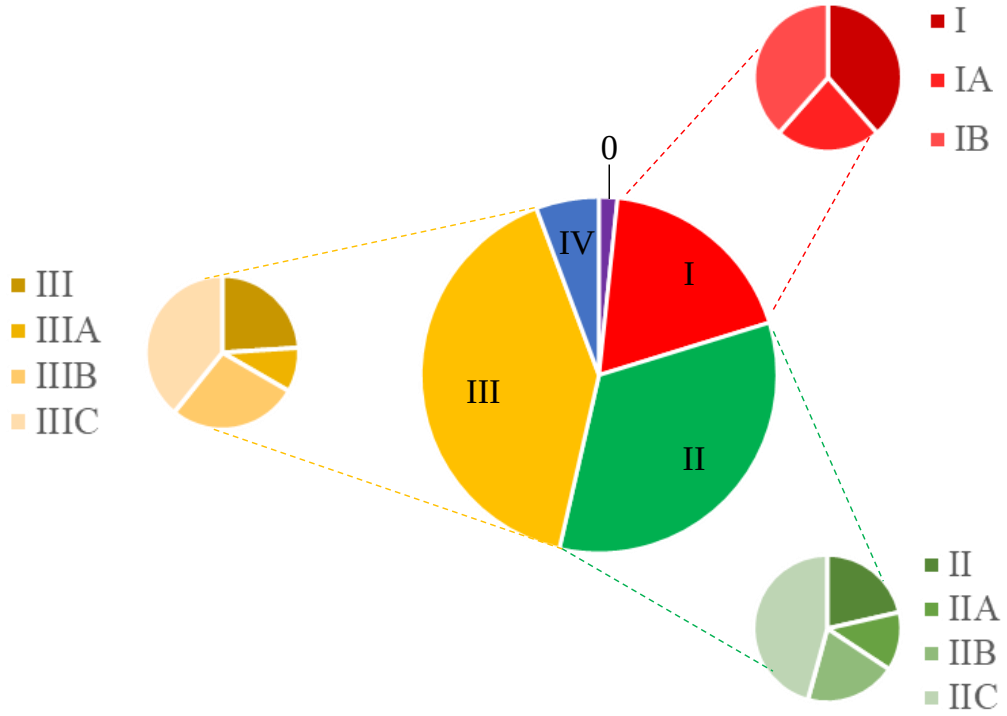


Figure S3.2. Patient staging in the melanoma cohort. Given than some patients were missing the sub-stage classification of A-C, all patients from a particular stage (I-IV) were combined in the central pie chart. Generated using TCGA patient data retrieved from reference 29.

3.7.4. Number of genes analyzed in each surfaceome

Table S3.2. Total genes analyzed per database and surfaceome

Number of genes	Curated surfaceome	Experimental surfaceome	Comprehensive surfaceome
Total	3,003	4,705	8,229
With TCGA data	2,976	4,578	6,187
With GTEx data	2,932	4,572	6,265
With TCGA and GTEx data	2,918	4,483	6,024

3.7.5. Tissue distribution among GTEx samples

Table S3.3. Tissue distribution of GTEx samples

Tissue	Number of samples	Percentage of total samples (%)
Adipose	515	6.58
Adrenal gland	127	1.62
Bladder	9	0.12
Blood	444	5.68
Blood vessel	604	7.72
Bone marrow	70	0.90
Brain	1,136	14.52
Breast	179	2.29
Cervix & uterus	10	0.13
Colon	304	3.89
Esophagus	649	8.30
Fallopian tube	5	0.06
Heart	376	4.81
Kidney	27	0.35
Liver	110	1.41
Lung	287	3.67
Muscle	396	5.06
Nerve	278	3.55
Ovary	88	1.13
Pancreas	165	2.11
Pituitary	107	1.37
Prostate	100	1.28
Salivary gland	55	0.70
Skin	812	10.38
Small intestine	92	1.18
Spleen	99	1.27
Stomach	172	2.20

Testis	165	2.11
Thyroid	278	3.55
Uterus	78	1.00
Vagina	84	1.07

3.7.6. Antibody dilutions

Table S3.4. Antibody dilutions used for immunocytochemistry

Immunogen	Recommended dilution	Used dilution
AMIGO3	1:50 – 1:200	1:200
CLN3	1:25	1:25
CXCR3	1:100 – 1:1,000	1:150
DLL3	1:200 (5 μg/mL from 1 mg/mL stock)	1:200
IL13RA2	1:500	1:500
LRP2	1:1,000 – 1:2,500	1:500
MCHR1	1:200 – 1:333 (3 – 5 μg/mL from 1 mg/mL stock)	1:333
MFRP	No recommended dilution for ICC	1:100*
TRPM8	1:300 – 1:2,000	1:300
TYRP1	1:100 – 1:250	1:50

***** Since there was no recommended dilution for ICC, we performed a titration with antibody and observed intense staining in the samples at 1:50 and 1:100 dilutions.

3.7.7. Sample genes excluded from experimental validation

Table S3.5. Sample genes identified in the screening that were not further investigated and the reason(s) for their elimination.

Gene	Reason(s) for elimination	References
EN2	Nuclear protein with no detectable presence in cytosol or membrane of melanoma samples. Its classification as a membrane protein stems from localization in caveolae in rat cerebellum tissue	18, 45
GAPDHS	Mostly expressed in nucleus and centrosome of testis tissues. Also found on the surface of sperm cells. Immunohistochemical staining of stage III melanoma patients showed cytoplasmic and nuclear staining only	18, 69,70 71
MLANA	Gene encodes MART-1, a widely used melanoma biomarker with cytoplasmic staining	31, 72
TYR	Protein is expressed by a high proportion of melanomas, but its staining is cytoplasmic	31, 72
ICOS	Protein is expressed in T cells, its activation stimulates proimmune activity	73-74

3.7.8. Reasons for the selection of the genes tested in immunochemistry

Table S3.6. Genes selected for immunocytochemical validation in melanoma and reasons for selection.

Gene	Ranking			Literature evidence	Extracellular domains‡	Reasons for selection
	Co	Ex	Cu			
AMIGO3	14	-	-	Protein is involved in inhibition of neurite growth and colocalizes with β-tubulin.[75] Gene expression was detected in several tissues.[47] Another member of the AMIGO family, AMIGO2, showed enhanced expression in liver metastases,[48] whereas AMIGO2 loss of expression promoted apoptosis and inhibited angiogenesis.[49]	76.0% (383 amino acids) of the protein is extracellular. The longest extracellular domain is 383 long and corresponds to the N-terminus.	Wide extracellular domain with no prior reports in cancer. Protein displayed good ranking in the screening using the comprehensive surfaceome. Other protein of the same family has showed relevance in cancer.
CLN3	5	2	4	Anti-apoptotic gene in melanoma,[35] involved in cell death and survival, whereas its inhibition hindered growth and viability.[34] Protein is expressed in several types of cancers, including breast, colon, melanoma, ovarian, neuroblastoma and glioblastoma.[34] CLN3 is	29.0% (127 amino acids) of the protein is luminal. The longest extracellular domain is 69 amino acids long.	Protein is consistently ranked among the best targets in the three surfaceome screenings. Protein has a functional role in cancer as an anti-apoptotic gene and its expression has been validated in

				expressed in lysosomes/endosomes, ER-Golgi apparatus, and to a lesser extent in the plasma membrane.[76] This protein might be trafficked to the cell membrane in its way to early endosomes.[77] CLN3 expression was observed in testis, liver, brain, spleen, and pancreas.[78-79]		several types of cancer.
CXCR3	35	31	15	23 (57%) of 40 primary melanomas express the protein on their cytoplasm and membrane based on immunohistochemical analysis.[58] CXCR3 expression might be involved in melanoma progression and metastasis to lymph nodes.[59-60] One peptide and several small molecule antagonists are reported.[61-64]	30.4% (112 amino acids) of the protein is extracellular. The longest extracellular domain is 53 amino acids long and corresponds to the N-terminus.	CXCR3 is one of the best-ranked proteins with known relevance in melanoma and potent and selective ligands, one of which is a peptide. Furthermore, several other ligands contain amide moieties, which could be readily incorporated into multivalent scaffolds through SPPS.
DLL3	19	14	-	Antibody-drug conjugates for DLL3 showed therapeutic effect in	79.6% (492 amino acids) of the protein	One of the best-ranked proteins with proven therapeutic

				metastatic melanoma, glioblastoma, and lung cancer.[80-82] Protein is expressed in ~50% of melanoma samples.[80] Furthermore, it displayed low-abundance expression on the cell surface of metastatic small cell lung cancer.[82]	is extracellular. The longest extracellular domain is 492 amino acids long and corresponds to the N-terminus.	potential in melanoma. This protein has been found on the surface of cancer cells and is mostly extracellular.
IL13RA2	61	52	-	This protein is selectively expressed in the cell surface of melanoma and, thus, is a potential target for melanoma immunotherapy.[12] Protein knockdown decreased breast cancer metastasis to the lung.[83] Peptides directed to IL13RA2 have been identified and used for therapeutic and imaging scaffolds by themselves or conjugated with nanoparticles.[67, 84-86] The IL13RA2 ligand Pep-1 binds to an extracellular site different from the protein's native ligand and promotes its internalization.[67]	90.3% (343 amino acids) of the protein is extracellular. The longest extracellular domain is 343 amino acids long and corresponds to the N-terminus.	One of the best-ranked proteins that possess known ligands. Selective expression on the cell surface of melanoma patient has been reported, thereby, validating the relevance of this biomarker. Protein contains a wide extracellular portion where peptide ligand Pep-1 binds. Ligand incorporated into constructs (e.g., nanoparticles or longer peptides) did not abolish its activity, which

						suggests that it could be readily coupled to multivalent scaffolds through SPPS.
LRP2	45	40	-	Gene was expressed in 60% of malignant melanomas.[87] Protein is expressed on the plasma membrane and endocytic apparatus of absorptive epithelia (including kidney), endometrium and thyroid.[88] Immunohistochemistry validated the presence of this protein on the plasma membrane of melanoma samples. Further, anti-LRP2 was found to be uptaken into live cells, and LRP2 knockdown lead to decreased proliferation.[87]	95.0% (4,423 amino acids) of the protein is extracellular. The longest extracellular domain is 4,423 amino acids long and corresponds to the N-terminus.	Protein displays an unusually large extracellular domain that could potentially be targeted. Binding-promoted internalization could be exploited to deliver imaging or therapeutic payloads to the cells expressing a particular receptor combination. Known cancer involvement and expression in melanoma samples validates the relevance of this target.
MCHR1	91	72	51	Gene was expressed in brain tissue, cultured melanocytes and 2 of 3	40.0% (169 amino acids) of the protein	MCHR1 is one of the best-ranked proteins with

				melanoma cell lines, but not other skin cells, including keratinocytes, dermal fibroblasts or endothelial cells.[51] Highly avid and selective peptide ligands with antagonistic[52] and agonistic[89] effect have been developed based on the protein's native ligand, melanin concentrating hormone (MCH). Furthermore, small molecule antagonists have been reported.[53-54]	is extracellular. The longest extracellular domain is 113 amino acids long and corresponds to the N-terminus.	reported ligands. The existence of highly avid antagonists and agonists, including both peptides and small molecules with reported structure-activity relationships, might allow facile incorporation into multivalent scaffolds through SPPS.
MFRP	3	1	1	Selective gene expression in adult and fetal brain cells, and undetectable expression level in any other normal tissues and 22 cancer cell lines across several tissue types (gastric, brain, blood, cervix, lung, lymphatic system and melanocytes).[90] Gene is expressed in retina, whereas its protein product is not critical for normal eye function.[50] MFRP is involved in the normal development and	84.5% (489 amino acids) of the protein is extracellular. The longest extracellular domain is 489 amino acids long and corresponds to the C-terminus.	Protein is consistently ranked among the best targets in the three surfaceome screenings. Reported low expression in normal tissues, where it holds no critical role.

				maintenance of photoreceptor outer segments.[91]		
TRPM8	44	42	16	Protein is involved in relevant processes for cancer, including cell proliferation, survival and invasion.[92] TRPM8 is overexpressed in several cancer types, including melanoma.[93] A wide variety of high affinity small molecule antagonists is known.[94-96]	13.0% (144 amino acids) of the protein is extracellular. The longest extracellular domain is 108 amino acids long.	One of the best-ranked proteins with known ligands. Although the extracellular portion of the protein is relatively small, it possesses a decently long extracellular domain of 108 aa. Furthermore, the protein is overexpressed on the cell membrane of several cancers.
TYRP1	31	25	14	Protein is frequently expressed on the cell surface of melanocytes and melanoma.[97] Antibodies for TYRP1 showed therapeutic effect on human melanoma tumors,[37] and moved to clinical trials.[36] TYRP1 expression intensity in melanoma is correlated to clinical outcome parameters.[98]	88.8% (477 amino acids) of the protein is luminal. The longest luminal domain is 477 amino acids long and corresponds to the N-terminus.	Protein is well ranked in all screenings and presents a long extracellular domain. Not only has the protein been identified in melanoma, but it has been the target of clinical trials, which validates the

	relevance of the receptor.

* “Co” = “comprehensive”, “Ex” = “experimental” and “Cu” = “curated”.

‡ Based on topology reported by UniProt.[26-27]

3.7.9. Threshold optimization and single target identified in the experimental and comprehensive surfaceomes

(A) *Melanoma*

Threshold	Patient coverage (%)	Hits
0.75×median	> 0	4,287
	≥ 10	***3,542***
	≥ 25	3,199
1×median	> 0	4,041
	≥ 10	3,038
2×median	> 0	287
	≥ 10	47

(B) *Normal tissues*

Threshold	Normal tissue coverage (%)	Hits
0.65×median	0	1
	≤ 10	43
	≤ 25	***182***
0.55×median	≤ 10	24
	≤ 25	99
0.45×median	≤ 10	14
	≤ 25	56

(C) *Differential coverage*

Protein	Melanoma coverage (%)	Normal tissue coverage (%)	Difference (%)*
MFRP	98.1	0.3	97.8
CLN3	100.0	3.6	96.4
PI4K2B	98.1	3.2	94.9
PIK3R2	100.0	5.6	94.4
TYR	90.9	3.1	87.8
EN2	91.6	6.1	85.5

* Difference in coverage is calculated using the melanoma cohort and the average coverage from three normal tissue ensembles.

Figure S3.3. Threshold optimization and best targets identified through the screening using the experimental surfaceome.

(A) *Melanoma*

Threshold	Patient coverage (%)	Hits
0.75×median	> 0	5,593
	≥ 10	***4,600***
	≥ 25	4,198
1×median	> 0	5,284
	≥ 10	4,031
2×median	> 0	616
	≥ 10	109

(B) *Normal tissues*

Threshold	Normal tissue coverage (%)	Hits
0.65×median	0	3
	≤ 10	65
	≤ 25	***228***
0.55×median	≤ 10	37
	≤ 25	131
0.45×median	≤ 10	17
	≤ 25	79

(C) *Differential coverage*

Protein	Melanoma coverage (%)	Normal tissue coverage (%)	Difference (%)*
PIGY	100.0	0.4	99.6
SFT2D3	100.0	0.8	99.2
MFRP	98.7	0.2	98.5
PI4K2B	98.5	2.2	96.3
CLN3	100.0	5.5	94.5
PIK3R2	100.0	5.6	94.4
TYR	91.1	3.1	88.0
SLC24A5	90.5	3.1	87.4

* Difference in coverage is calculated using the melanoma cohort and the average coverage from three normal tissue ensembles.

Figure S3.4. Threshold optimization and best targets identified through the screening using the comprehensive surfaceome.

3.7.10. Immunocytochemical results of selected protein targets

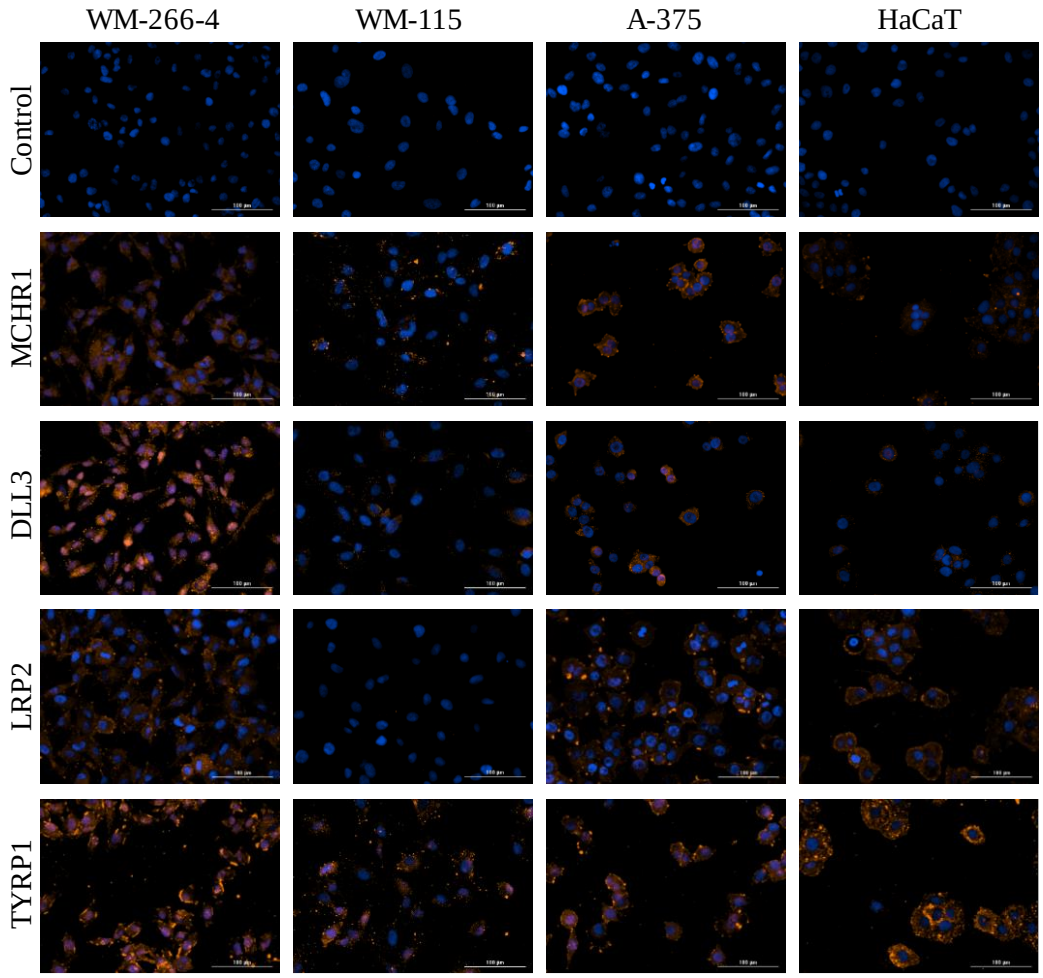


Figure S3.5. Protein target validation in melanoma cell lines WM-266-4, WM-115, and A-375, and immortalized keratinocyte HaCaT as control. All size bars correspond to 100 μm. Secondary antibodies were tagged with DyLight 550 and fluorescence was collected in the RFP (red/orange) channel of the instrument. Nucleus was stained with blue fluorescent dye Hoechst 33342. Controls (first row) were treated with secondary antibody and nuclear dye only.

3.8. References

1. Garbe, C.; Leiter, U., Melanoma epidemiology and trends. *Clin. Dermatol.* **2009,** *27* (1), 3-9.

2. Sacchetto, L.; Zanetti, R.; Comber, H.; Bouchardy, C.; Brewster, D. H.; Broganelli, P.; Chirlaque, M. D.; Coza, D.; Galceran, J.; Gavin, A.; Hackl, M.; Katalinic, A.; Larønningen, S.; Louwman, M. W. J.; Morgan, E.; Robsahm, T. E.; Sanchez, M. J.; Tryggvadóttir, L.; Tumino, R.; Van Eycken, E.; Vernon, S.; Zadnik, V.; Rosso, S., Trends in incidence of thick, thin and *in situ* melanoma in Europe. *Eur. J. Cancer* **2018,** *92*, 108-118.

3. Aitken, J. F.; Youlden, D. R.; Baade, P. D.; Soyer, H. P.; Green, A. C.; Smithers, B. M., Generational shift in melanoma incidence and mortality in Queensland, Australia, 1995–2014. *Int. J. Cancer* **2018,** *142* (8), 1528-1535.

4. American Cancer Society *Cancer facts & figures 2018*; American Cancer Society: Atlanta, GA, 2018; pp 1-76.

5. Curry, B. J.; Farrelly, M.; Mersey, P., Evaluation of S-100β assays for the prediction of recurrence and prognosis in patients with AJCC stage I-III melanoma. *Melanoma Res.* **1999,** *9* (6), 557-568.

6. Balch, C. M.; Gershenwald, J. E.; Soong, S.-j.; Thompson, J. F.; Atkins, M. B.; Byrd, D. R.; Buzaid, A. C.; Cochran, A. J.; Coit, D. G.; Ding, S.; Eggermont, A. M.; Flaherty, K. T.; Gimotty, P. A.; Kirkwood, J. M.; McMasters, K. M.; Mihm, M. C.; Morton, D. L.; Ross, M. I.; Sober, A. J.; Sondak, V. K., Final version of 2009 AJCC melanoma staging and classification. *J. Clin. Oncol.* **2009,** *27* (36), 6199-6206.

7. Vultur, A.; Villanueva, J.; Herlyn, M., Targeting BRAF in advanced melanoma: A first step toward manageable disease. *Clin. Cancer Res.* **2011,** *17* (7), 1658-1663.

8. Davies, H.; Bignell, G. R.; Cox, C.; Stephens, P.; Edkins, S.; Clegg, S.; Teague, J.; Woffendin, H.; Garnett, M. J.; Bottomley, W.; Davis, N.; Dicks, E.; Ewing, R.; Floyd, Y.; Gray, K.; Hall, S.; Hawes, R.; Hughes, J.; Kosmidou, V.; Menzies, A.; Mould, C.; Parker, A.; Stevens, C.; Watt, S.; Hooper, S.; Wilson, R.; Jayatilake, H.; Gusterson, B. A.; Cooper, C.; Shipley, J.; Hargrave, D.; Pritchard-Jones, K.; Maitland, N.; Chenevix-Trench, G.; Riggins, G. J.; Bigner, D. D.; Palmieri, G.; Cossu, A.; Flanagan, A.; Nicholson, A.; Ho, J. W. C.; Leung, S. Y.; Yuen, S. T.; Weber, B. L.; Seigler, H. F.; Darrow, T. L.; Paterson, H.; Marais, R.; Marshall, C. J.; Wooster, R.; Stratton, M. R.; Futreal, P. A., Mutations of the *BRAF* gene in human cancer. *Nature (London)* **2002,** *417* (6892), 949-954.

9. Fedorenko, I. V.; Gibney, G. T.; Sondak, V. K.; Smalley, K. S. M., Beyond BRAF: Where next for melanoma therapy? *Br. J. Cancer* **2015,** *112* (2), 217-226.

10. Topalian, S. L.; Hodi, F. S.; Brahmer, J. R.; Gettinger, S. N.; Smith, D. C.; McDermott, D. F.; Powderly, J. D.; Carvajal, R. D.; Sosman, J. A.; Atkins, M. B.; Leming, P. D.; Spigel, D. R.; Antonia, S. J.; Horn, L.; Drake, C. G.; Pardoll, D. M.; Chen, L.; Sharfman, W. H.; Anders, R. A.; Taube, J. M.; McMiller, T. L.; Xu, H.; Korman, A. J.; Jure-Kunkel, M.; Agrawal, S.; McDonald, D.; Kollia, G. D.; Gupta, A.; Wigginton, J. M.; Sznol, M., Safety, activity, and immune correlates of Anti–PD-1 antibody in cancer. *N. Engl. J. Med.* **2012,** *366* (26), 2443-2454.

11. Robert, C.; Schachter, J.; Long, G. V.; Arance, A.; Grob, J. J.; Mortier, L.; Daud, A.; Carlino, M. S.; McNeil, C.; Lotem, M.; Larkin, J.; Lorigan, P.; Neyns, B.; Blank, C. U.; Hamid, O.; Mateus, C.; Shapira-Frommer, R.; Kosh, M.; Zhou, H.; Ibrahim, N.; Ebbinghaus, S.; Ribas, A., Pembrolizumab versus Ipilimumab in advanced melanoma. *N. Engl. J. Med.* **2015,** *372* (26), 2521-2532.

12. Beard, R. E.; Abate-Daga, D.; Rosati, S. F.; Zheng, Z.; Wunderlich, J. R.; Rosenberg, S. A.; Morgan, R. A., Gene expression profiling using Nanostring digital RNA counting to identify potential target antigens for melanoma immunotherapy. *Clin. Cancer Res.* **2013,** *19* (18), 4941-4950.

13. Josan, J. S.; Vagner, J.; Handl, H. L.; Sankaranarayanan, R.; Gillies, R. J.; Hruby, V. J., Solid-phase synthesis of heterobivalent ligands targeted to melanocortin and cholecystokinin receptors. *Int. J. Pept. Res. Ther.* **2008,** *14* (4), 293-300.

14. Josan, J. S.; Handl, H. L.; Sankaranarayanan, R.; Xu, L.; Lynch, R. M.; Vagner, J.; Mash, E. A.; Hruby, V. J.; Gillies, R. J., Cell-specific targeting by heterobivalent ligands. *Bioconjugate Chem.* **2011,** *22* (7), 1270-1278.

15. Xu, L.; Josan, J. S.; Vagner, J.; Caplan, M. R.; Hruby, V. J.; Mash, E. A.; Lynch, R. M.; Morse, D. L.; Gillies, R. J., Heterobivalent ligands target cell-surface receptor combinations in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** *109* (52), 21295-21300.

16. Hart, N. J.; Chung, W. J.; Weber, C.; Ananthakrishnan, K.; Anderson, M.; Patek, R.; Zhang, Z.; Limesand, S. W.; Vagner, J.; Lynch, R. M., Hetero-bivalent GLP-1/Glibenclamide for targeting pancreatic β-cells. *ChemBioChem* **2014,** *15* (1), 135-145.

17. VIB-Ugent Center for Plant Systems Biology http://bioinformatics.psb.ugent.be/webtools/Venn/ (accessed 04/14/2017).

18. Uhlén, M.; Fagerberg, L.; Hallström, B. M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; Olsson, I.; Edlund, K.; Lundberg, E.; Navani, S.; Szigyarto, C. A.-K.; Odeberg, J.; Djureinovic, D.; Takanen, J. O.; Hober, S.; Alm, T.; Edqvist, P.-H.; Berling, H.; Tegel, H.; Mulder, J.; Rockberg, J.; Nilsson, P.; Schwenk, J. M.; Hamsten, M.;

von Feilitzen, K.; Forsberg, M.; Persson, L.; Johansson, F.; Zwahlen, M.; von Heijne, G.; Nielsen, J.; Pontén, F., Tissue-based map of the human proteome. *Science* **2015,** *347* (6220), 1260419.

19. Human Protein Atlas. www.proteinatlas.org (accessed 04/16/2017).

20. Bausch-Fluck, D.; Hofmann, A.; Bock, T.; Frei, A. P.; Cerciello, F.; Jacobs, A.; Moest, H.; Omasits, U.; Gundry, R. L.; Yoon, C.; Schiess, R.; Schmidt, A.; Mirkowska, P.; Härtlová, A.; Van Eyk, J. E.; Bourquin, J.-P.; Aebersold, R.; Boheler, K. R.; Zandstra, P.; Wollscheid, B., A mass spectrometric-derived cell surface protein atlas. *PLoS ONE* **2015,** *10* (4), e0121314.

21. Prasad, T. S. K.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A.; Balakrishnan, L.; Marimuthu, A.; Banerjee, S.; Somanathan, D. S.; Sebastian, A.; Rani, S.; Ray, S.; Harrys Kishore, C. J.; Kanth, S.; Ahmed, M.; Kashyap, M. K.; Mohmood, R.; Ramachandra, Y. L.; Krishna, V.; Rahiman, B. A.; Mohan, S.; Ranganathan, P.; Ramabadran, S.; Chaerkady, R.; Pandey, A., Human Protein Reference Database—2009 update. *Nucleic Acids Res.* **2009,** *37*, D767-D772.

22. Human Protein Reference Database. http://www.hprd.org/ (accessed 04/16/2016).

23. da Cunha, J. P. C.; Galante, P. A. F.; de Souza, J. E.; de Souza, R. F.; Carvalho, P. M.; Ohara, D. T.; Moura, R. P.; Oba-Shinja, S. M.; Marie, S. K. N.; Silva, W. A.; Perez, R. O.; Stransky, B.; Pieprzyk, M.; Moore, J.; Caballero, O.; Gama-Rodrigues, J.; Habr-Gama, A.; Kuo, W. P.; Simpson, A. J.; Camargo, A. A.; Old, L. J.; de Souza, S. J., Bioinformatics construction of the human cell surfaceome. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (39), 16752-16757.

24. Santana de Souza, J. E.; Favoretto Galante, P. A.; Bueno de Almeida, R. V.; Chagas da Cunha, J. P.; Takatori Ohara, D.; Ohno-Machado, L.; Old, L. J.; de Souza, S. J. Surfaceome DB. http://www.bioinformatics-brazil.org/surfaceome/ (accessed 04/10/2016).

25. Santana de Souza, J. E.; Favoretto Galante, P. A.; Bueno de Almeida, R. V.; Chagas da Cunha, J. P.; Takatori Ohara, D.; Ohno-Machado, L.; Old, L. J.; de Souza, S. J., SurfaceomeDB: A cancer-orientated database for genes encoding cell surface proteins. *Cancer Immun.* **2012,** *12*, 15-20.

26. The Uniprot Consortium, UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* **2017,** *45* (D1), D158-D169.

27. The Uniprot Consortium UniProt. http://www.uniprot.org/ (accessed 04/10/2016).

28. Balagurunathan, Y.; Morse, D. L.; Hostetter, G.; Shanmugam, V.; Stafford, P.; Shack, S.; Pearson, J.; Trissal, M.; Demeure, M. J.; Von Hoff, D. D.; Hruby, V. J.; Gillies, R. J.; Han, H., Gene expression profiling-based identification of cell-surface targets for developing multimeric ligands in pancreatic cancer. *Mol. Cancer Ther.* **2008,** *7* (9), 3071-3080.

29. University of California Santa Cruz Xena Browser. http://xena.ucsc.edu (accessed 03/20/2018).

30. Balagurunathan, Y.; Morse, D.; Stafford, P.; Pearson, J.; Hostetter, G.; Von Hoff, D.; Demeure, M.; Gillies, R.; Han, H., Gene expression profiling based identification of cell surface targets for multimeric ligand development in pancreatic cancer. *Cancer Res.* **2007,** *67* (9 Supplement), 4961-4961.

31. Ordóñez, N. G., Value of melanocytic-associated immunohistochemical markers in the diagnosis of malignant melanoma: A review and update. *Hum. Pathol.* **2014,** *45* (2), 191-205.

32. Fernandez, L. P.; Milne, R. L.; Pita, G.; Avilés, J. A.; Lázaro, P.; Benítez, J.; Ribas, G., *SLC45A2*: A novel malignant melanoma-associated gene. *Hum. Mutat.* **2008,** *29* (9), 1161-1167.

33. Cook, A. L.; Chen, W.; Thurber, A. E.; Smit, D. J.; Smith, A. G.; Bladen, T. G.; Brown, D. L.; Duffy, D. L.; Pastorino, L.; Bianchi-Scarra, G.; Helen Leonard, J.; Stow, J. L.; Sturm, R. A.,

Analysis of cultured human melanocytes based on polymorphisms within the *SLC45A2*/MATP, *SLC24A5*/NCKX5, and *OCA2*/P Loci. *J. Invest. Dermatol.* **2009,** *129* (2), 392-405.

34. Rylova, S. N.; Amalfitano, A.; Persaud-Sawin, D.-A.; Guo, W.-X.; Chang, J.; Jansen, P. J.; Proia, A. D.; Boustany, R.-M., The *CLN3* gene is a novel molecular target for cancer drug discovery. *Cancer Res.* **2002,** *62* (3), 801-808.

35. Zhang, Q.; Wu, J.; Nguyen, A.; Wang, B.-D.; He, P.; Laurent, G. S.; Rennert, O. M.; Su, Y. A., Molecular mechanism underlying differential apoptosis between human melanoma cell lines UACC903 and UACC903(+6) revealed by mitochondria-focused cDNA microarrays. *Apoptosis* **2008,** *13* (8), 993-1004.

36. Khalil, D. N.; Postow, M. A.; Ibrahim, N.; Ludwig, D. L.; Cosaert, J.; Kambhampati, S. R. P.; Tang, S.; Grebennik, D.; Kauh, J. S. W.; Lenz, H.-J.; Flaherty, K. T.; Hodi, F. S.; Lawrence, D. P.; Wolchok, J. D., An open-label, dose–escalation phase I study of anti-TYRP1 monoclonal antibody IMC-20D7S for patients with relapsed or refractory melanoma. *Clin. Cancer Res.* **2016,** *22* (21), 5204-5210.

37. Patel, D.; Balderes, P.; Lahiji, A.; Melchior, M.; Ng, S.; Bassi, R.; Wu, Y.; Griffith, H.; Jimenez, X.; Ludwig, D. L.; Hicklin, D. J.; Kang, X., Generation and characterization of a therapeutic human antibody to melanoma antigen TYRP1. *Hum. Antibodies* **2007,** *16* (3/4), 127-136.

38. Aran, D.; Sirota, M.; Butte, A. J., Systematic pan-cancer analysis of tumour purity. *Nature Commun.* **2015,** *6*, 8971.

39. Kosti, I.; Jain, N.; Aran, D.; Butte, A. J.; Sirota, M., Cross-tissue analysis of gene and protein expression in normal and cancer tissues. *Sci. Rep.* **2016,** *6*, 24799.

40. Edfors, F.; Danielsson, F.; Hallström, B. M.; Käll, L.; Lundberg, E.; Pontén, F.; Forsström, B.; Uhlén, M., Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol. Syst. Biol.* **2016,** *12* (10), 883.

41. Greenbaum, D.; Colangelo, C.; Williams, K.; Gerstein, M., Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* **2003,** *4* (9), 117.

42. Herlyn, M.; Balaban, G.; Bennicelli, J.; Guerry, I. V. D.; Halaban, R.; Herlyn, D.; Elder, D. E.; Maul, G. G.; Steplewski, Z.; Nowell, P. C.; Clark, W. H.; Koprowski, H., Primary melanoma cells of the vertical growth phase: Similarities to metastatic cells. *JNCI, J. Natl. Cancer Inst.* **1985,** *74* (2), 283-289.

43. Al-Ghoul, M.; Brück, T. B.; Lauer-Fields, J. L.; Asirvatham, V. S.; Zapata, C.; Kerr, R. G.; Fields, G. B., Comparative proteomic analysis of matched primary and metastatic melanoma cell lines. *J. Proteome Res.* **2008,** *7* (9), 4107-4118.

44. Qiu, H.; Wang, Y., Quantitative analysis of surface plasma membrane proteins of primary and metastatic melanoma cells. *J. Proteome Res.* **2008,** *7* (5), 1904-1915.

45. Joliot, A.; Trembleau, A.; Raposo, G.; Calvet, S.; Volovitch, M.; Prochiantz, A., Association of Engrailed homeoproteins with vesicles presenting caveolae-like properties. *Development* **1997,** *124* (10), 1865-1875.

46. Guda, C.; Subramaniam, S., TARGET: A new method for predicting protein subcellular localization in eukaryotes. *Bioinformatics* **2005,** *21* (21), 3963-3969.

47. Kuja-Panula, J.; Kiiltomäki, M.; Yamashiro, T.; Rouhiainen, A.; Rauvala, H., AMIGO, a transmembrane protein implicated in axon tract development, defines a novel protein family with leucine-rich repeats. *J. Cell Biol.* **2003,** *160* (6), 963-973.

48. Kanda, Y.; Osaki, M.; Onuma, K.; Sonoda, A.; Kobayashi, M.; Hamada, J.; Nicolson, G. L.; Ochiya, T.; Okada, F., *Amigo2*-upregulation in tumour cells facilitates their attachment to liver endothelial cells resulting in liver metastases. *Sci. Rep.* **2017,** *7*, 43567.

49. Park, H.; Lee, S.; Shrestha, P.; Kim, J.; Park, J. A.; Ko, Y.; Ban, Y. H.; Park, D. Y.; Ha, S. J.; Koh, G. Y.; Hong, V. S.; Mochizuki, N.; Kim, Y. M.; Lee, W.; Kwon, Y. G., AMIGO2, a novel membrane anchor of PDK1, controls cell survival and angiogenesis via Akt activation. *J. Cell Biol.* **2015,** *211* (3), 619-637.

50. Sundin, O. H.; Leppert, G. S.; Silva, E. D.; Yang, J.-M.; Dharmaraj, S.; Maumenee, I. H.; Santos, L. C.; Parsa, C. F.; Traboulsi, E. I.; Broman, K. W.; DiBernardo, C.; Sunness, J. S.; Toy, J.; Weinberg, E. M., Extreme hyperopia is the result of null mutations in *MFRP*, which encodes a Frizzled-related protein. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (27), 9553-9558.

51. Hoogduijn, M. J.; Ancans, J.; Suzuki, I.; Estdale, S.; Thody, A. J., Melanin-concentrating hormone and its receptor are expressed and functional in human skin. *Biochem. Biophys. Res. Commun.* **2002,** *296* (3), 698-701.

52. Bednarek, M. A.; Hreniuk, D. L.; Tan, C.; Palyha, O. C.; MacNeil, D. J.; Van der Ploeg, L. H. Y.; Howard, A. D.; Feighner, S. D., Synthesis and biological evaluation *in vitro* of selective, high affinity peptide antagonists of human melanin-concentrating hormone action at human Melanin-Concentrating Hormone Receptor 1. *Biochemistry* **2002,** *41* (20), 6383-6390.

53. Handlon, A. L.; Zhou, H., Melanin-Concentrating Hormone-1 Receptor antagonists for the treatment of obesity. *J. Med. Chem.* **2006,** *49* (14), 4017-4022.

54. Högberg, T.; Frimurer, T. M.; Sasmal, P. K., Melanin concentrating hormone receptor 1 (MCHR1) antagonists—Still a viable approach for obesity treatment? *Bioorg. Med. Chem. Lett.* **2012,** *22* (19), 6039-6047.

55. Chatterjee, S.; Behnam Azad, B.; Nimmagadda, S., Chapter Two - The intricate role of CXCR4 in cancer. In *Emerging Applications of Molecular Imaging to Oncology. Advances in Cancer Research*, Pomper, M. G.; Fisher, P. B., Eds. Academic Press: London, UK, 2014; Vol. 124, pp 31-82.

56. Domanska, U. M.; Kruizinga, R. C.; Nagengast, W. B.; Timmer-Bosscha, H.; Huls, G.; de Vries, E. G. E.; Walenkamp, A. M. E., A review on CXCR4/CXCL12 axis in oncology: No place to hide. *Eur. J. Cancer* **2013,** *49* (1), 219-230.

57. Teicher, B. A.; Fricker, S. P., CXCL12 (SDF-1)/CXCR4 Pathway in Cancer. *Clin. Cancer Res.* **2010,** *16* (11), 2927-2931.

58. Longo-Imedio, M. I.; Longo, N.; Treviño, I.; Lázaro, P.; Sánchez-Mateos, P., Clinical significance of CXCR3 and CXCR4 expression in primary melanoma. *Int. J. Cancer* **2005,** *117* (5), 861-865.

59. Jenkins, M. H.; Brinckerhoff, C. E.; Mullins, D. W., CXCR3 signaling in $BRAF^{WT}$ melanoma increases IL-8 expression and tumorigenicity. *PLoS ONE* **2015,** *10* (3), e0121140.

60. Kawada, K.; Sonoshita, M.; Sakashita, H.; Takabayashi, A.; Yamaoka, Y.; Manabe, T.; Inaba, K.; Minato, N.; Oshima, M.; Taketo, M. M., Pivotal role of CXCR3 in melanoma cell metastasis to lymph nodes. *Cancer Res.* **2004,** *64* (11), 4010-4017.

61. Crosignani, S.; Missotten, M.; Cleva, C.; Dondi, R.; Ratinaud, Y.; Humbert, Y.; Mandal, A. B.; Bombrun, A.; Power, C.; Chollet, A.; Proudfoot, A., Discovery of a novel series of CXCR3 antagonists. *Bioorg. Med. Chem. Lett.* **2010,** *20* (12), 3614-3617.

62. Chen, X.; Mihalic, J.; Deignan, J.; Gustin, D. J.; Duquette, J.; Du, X.; Chan, J.; Fu, Z.; Johnson, M.; Li, A.-R.; Henne, K.; Sullivan, T.; Lemon, B.; Ma, J.; Miao, S.; Tonn, G.; Collins,

T.; Medina, J. C., Discovery of potent and specific CXCR3 antagonists. *Bioorg. Med. Chem. Lett.* **2012,** *22* (1), 357-362.

63. Andrews, S. P.; Cox, R. J., Small molecule CXCR3 antagonists. *J. Med. Chem.* **2016,** *59* (7), 2894-2917.

64. Ondeyka, J. G.; Herath, K. b.; Jayasuriya, H.; Polishook, J. D.; Bills, G. F.; Dombrowski, A. W.; Mojena, M.; Koch, G.; DiSalvo, J.; DeMartino, J.; Guan, Z.; Nanakorn, W.; Morenberg, C. M.; Balick, M. J.; Stevenson, D. W.; Slattery, M.; Borris, R. P.; Singh, S. B., Discovery of structurally diverse natural product antagonists of chemokine receptor CXCR3. *Mol. Diversity* **2005,** *9* (1), 123-129.

65. Qian, M.; Vasudevan, L.; Huysentruyt, J.; Risseeuw Martijn, D. P.; Stove, C.; Vanderheyden Patrick, M. L.; Van Craenenbroeck, K.; Van Calenbergh, S., Design, synthesis, and biological evaluation of bivalent ligands targeting Dopamine D_2-Like Receptors and the μ-Opioid Receptor. *ChemMedChem* **2018,** *13* (9), 944-956.

66. Soriano, A.; Ventura, R.; Molero, A.; Hoen, R.; Casadó, V.; Cortés, A.; Fanelli, F.; Albericio, F.; Lluís, C.; Franco, R.; Royo, M., Adenosine A_{2A} Receptor-antagonist/Dopamine D_2 Receptor-agonist bivalent ligands as pharmacological tools to detect A_{2A}-D_2 Receptor heteromers. *J. Med. Chem.* **2009,** *52* (18), 5590-5602.

67. Pandya, H.; Gibo, D. M.; Garg, S.; Kridel, S.; Debinski, W., An interleukin 13 receptor α 2–specific peptide homes to human Glioblastoma multiforme xenografts. *Neuro-Oncology* **2012,** *14* (1), 6-18.

68. Chames, P.; van Regenmortel, M.; Weiss, E.; Baty, D., Therapeutic antibodies: Successes, limitations and hopes for the future. *Br. J. Pharmacol.* **2009,** *157* (2), 220-233.

69. Roman, A.; Pall, E.; Moldovan, M.; Rusu, D.; Soritau, O.; Festila, D.; Lupse, M., Cytotoxicity of experimental resin composites on mesenchymal stem cells isolated from two oral sources. *Microsc. Microanal.* **2016,** *22* (5), 1018-1033.

70. Danshina, P. V.; Qu, W.; Temple, B. R.; Rojas, R. J.; Miley, M. J.; Machius, M.; Betts, L.; O'Brien, D. A., Structural analyses to identify selective inhibitors of glyceraldehyde 3-phosphate dehydrogenase-S, a sperm-specific glycolytic enzyme. *MHR: Basic Sci. Reprod. Med.* **2016,** *22* (6), 410-426.

71. Falkenius, J.; Lundeberg, J.; Johansson, H.; Tuominen, R.; Frostvik-Stolt, M.; Hansson, J.; Egyhazi Brage, S., High expression of glycolytic and pigment proteins is associated with worse clinical outcome in stage III melanoma. *Melanoma Res.* **2013,** *23* (6), 452-460.

72. Ohsie, S. J.; Sarantopoulos, G. P.; Cochran, A. J.; Binder, S. W., Immunohistochemical characteristics of melanoma. *J. Cutaneous Pathol.* **2008,** *35* (5), 433-444.

73. Wallin, J. J.; Liang, L.; Bakardjiev, A.; Sha, W. C., Enhancement of $CD8^+$ T cell responses by ICOS/B7h costimulation. *J. Immunol.* **2001,** *167* (1), 132-139.

74. Zhang, Y.; Luo, Y.; Qin, S.-L.; Mu, Y.-F.; Qi, Y.; Yu, M.-H.; Zhong, M., The clinical impact of ICOS signal in colorectal cancer patients. *OncoImmunology* **2016,** *5* (5), e1141857.

75. Ahmed, Z.; Douglas, M. R.; John, G.; Berry, M.; Logan, A., AMIGO3 is an NgR1/p75 co-receptor signalling axon growth inhibition in the acute phase of adult central nervous system injury. *PLoS ONE* **2013,** *8* (4), e61878.

76. Järvelä, I.; Lehtovirta, M.; Tikkanen, R.; Kyttälä, A.; Jalanko, A., Defective intracellular transport of CLN3 is the molecular basis of Batten disease (JNCL). *Hum. Mol. Genet.* **1999,** *8* (6), 1091-1098.

77. Haskell, R. E.; Carr, C. J.; Pearce, D. A.; Bennett, M. J.; Davidson, B. L., Batten disease: Evaluation of *CLN3* mutations on protein localization and function. *Hum. Mol. Genet.* **2000,** *9* (5), 735-744.

78. Narayan, S. B.; Pastor, J. V.; Mitchison, H. M.; Bennett, M. J., CLN3L, a novel protein related to the Batten disease protein, is overexpressed in *Cln3*$^{-/-}$ mice and in Batten disease. *Brain* **2004,** *127* (8), 1748-1754.

79. Margraf, L. R.; Boriack, R. L.; Routheut, A. A. J.; Cuppen, I.; Alhilali, L.; Bennett, C. J.; Bennett, M. J., Tissue expression and subcellular localization of CLN3, the Batten disease protein. *Mol. Genet. Metab.* **1999,** *66* (4), 283-289.

80. Peng, S. L.; Saunders, L.; Bheddah, S.; Williams, S.; Aggarwal, R. R.; Shea, J. E.; Lee, E. Y.; Huang, J.; Zemek, A. J.; Longacre, T. A.; Ball, D. W.; Scaife, C. L.; Nelkin, B.; Anthony, L. B.; Kunz, P. L.; Small, E. J.; Dylla, S., Metastatic melanoma, glioblastoma and high-grade extrapulmonary neuroendocrine carcinomas (NECs) as novel indications for rovalpituzumab tesirine: A delta-like protein 3 (DLL3)-targeted antibody-drug conjugate (ADC). *J. Clin. Oncol.* **2016,** *34* (15 Supplement), 11611.

81. Saunders, L. R.; Bankovich, A. J.; Anderson, W. C.; Aujay, M. A.; Bheddah, S.; Black, K.; Desai, R.; Escarpe, P. A.; Hampl, J.; Laysang, A.; Liu, D.; Lopez-Molina, J.; Milton, M.; Park, A.; Pysz, M. A.; Shao, H.; Slingerland, B.; Torgov, M.; Williams, S. A.; Foord, O.; Howard, P.; Jassem, J.; Badzio, A.; Czapiewski, P.; Harpole, D. H.; Dowlati, A.; Massion, P. P.; Travis, W. D.; Pietanza, M. C.; Poirier, J. T.; Rudin, C. M.; Stull, R. A.; Dylla, S. J., A DLL3-targeted antibody-drug conjugate eradicates high-grade pulmonary neuroendocrine tumor-initiating cells in vivo. *Sci. Transl. Med.* **2015,** *7* (302), 302ra136.

82. Sharma, S. K.; Pourat, J.; Abdel-Atti, D.; Carlin, S. D.; Piersigilli, A.; Bankovich, A. J.; Gardner, E. E.; Hamdy, O.; Isse, K.; Bheddah, S.; Sandoval, J.; Cunanan, K. M.; Johansen, E. B.; Allaj, V.; Sisodiya, V.; Liu, D.; Zeglis, B. M.; Rudin, C. M.; Dylla, S. J.; Poirier, J. T.; Lewis, J. S., Non-invasive interrogation of DLL3 expression in metastatic small cell lung cancer. *Cancer Res.* **2017,** *77* (14), 3931-3940.

83. Minn, A. J.; Gupta, G. P.; Siegel, P. M.; Bos, P. D.; Shu, W.; Giri, D. D.; Viale, A.; Olshen, A. B.; Gerald, W. L.; Massagué, J., Genes that mediate breast cancer metastasis to lung. *Nature (London)* **2005,** *436*, 518-524.

84. Sattiraju, A.; Sai, K. K. S.; Xuan, A.; Pandya, D. N.; Almaguel, F. G.; Wadas, T. J.; Herpai, D. M.; Debinski, W.; Mintz, A., IL13RA2 targeted alpha particle therapy against glioblastomas. *Oncotarget* **2017,** *8* (26), 42997-43007.

85. Sai, K. K. S.; Sattiraju, A.; Almaguel, F. G.; Xuan, A.; Rideout, S.; Krishnaswamy, R. S.; Zhang, J.; Herpai, D. M.; Debinski, W.; Mintz, A., Peptide-based PET imaging of the tumor restricted IL13RA2 biomarker. *Oncotarget* **2017,** *8* (31), 50997-51007.

86. Wang, B.; Lv, L.; Wang, Z.; Zhao, Y.; Wu, L.; Fang, X.; Xu, Q.; Xin, H., Nanoparticles functionalized with Pep-1 as potential glioma targeting delivery system via interleukin 13 receptor α2-mediated endocytosis. *Biomaterials* **2014,** *35* (22), 5897-5907.

87. Andersen, R. K.; Hammer, K.; Hager, H.; Christensen, J. N.; Ludvigsen, M.; Honoré, B.; Thomsen, M.-B. H.; Madsen, M., Melanoma tumors frequently acquire *LRP2*/megalin expression, which modulates melanoma cell proliferation and survival rates. *Pigm. Cell Melanoma Res.* **2015,** *28* (3), 267-280.

88. Christensen, E. I.; Birn, H., Megalin and cubilin: Multifunctional endocytic receptors. *Nat. Rev. Mol. Cell Biol.* **2002,** *3* (4), 258-268.

89. Bednarek, M. A.; Tan, C.; Hreniuk, D. L.; Palyha, O. C.; MacNeil, D. J.; Van der Ploeg, L. H. Y.; Howard, A. D.; Feighner, S. D., Synthesis and biological evaluation *in vitro* of a selective, high potency peptide agonist of human melanin-concentrating hormone action at human Melanin-Concentrating Hormone Receptor 1. *J. Biol. Chem.* **2002,** *277* (16), 13821-13826.

90. Katoh, M., Molecular cloning and characterization of *MFRP*, a novel gene encoding a Membrane-type Frizzled-related Protein. *Biochem. Biophys. Res. Commun.* **2001,** *282* (1), 116-123.

91. Won, J.; Smith, R. S.; Peachey, N. S.; Wu, J.; Hicks, W. L.; Naggert, J. K.; Nishina, P. M., Membrane Frizzled-Related Protein is necessary for the normal development and maintenance of photoreceptor outer segments. *Visual Neurosci.* **2008,** *25* (4), 563-574.

92. Yee, N., Roles of TRPM8 ion channels in cancer: proliferation, survival, and invasion. *Cancers* **2015,** *7* (4), 2134-2146.

93. Yamamura, H.; Ugawa, S.; Ueda, T.; Morita, A.; Shimada, S., TRPM8 activation suppresses cellular viability in human melanoma. *Am. J. Physiol. Cell Physiol.* **2008,** *295* (2), C296-C301.

94. Pérez de Vega, M. J.; Gómez-Monterrey, I.; Ferrer-Montiel, A.; González-Muñiz, R., Transient Receptor Potential Melastatin 8 Channel (TRPM8) modulation: Cool entryway for treating pain and cancer. *J. Med. Chem.* **2016,** *59* (22), 10006-10029.

95. Parks, D. J.; Parsons, W. H.; Colburn, R. W.; Meegalla, S. K.; Ballentine, S. K.; Illig, C. R.; Qin, N.; Liu, Y.; Hutchinson, T. L.; Lubin, M. L.; Stone, D. J.; Baker, J. F.; Schneider, C. R.; Ma, J.; Damiano, B. P.; Flores, C. M.; Player, M. R., Design and optimization of benzimidazole-containing Transient Receptor Potential Melastatin 8 (TRPM8) antagonists. *J. Med. Chem.* **2011,** *54* (1), 233-247.

96. Ohmi, M.; Shishido, Y.; Inoue, T.; Ando, K.; Fujiuchi, A.; Yamada, A.; Watanabe, S.; Kawamura, K., Identification of a novel 2-pyridyl-benzensulfonamide derivative, RQ-00203078, as a selective and orally active TRPM8 antagonist. *Bioorg. Med. Chem. Lett.* **2014,** *24* (23), 5364-5368.

97. Takechi, Y.; Hara, I.; Naftzger, C.; Xu, Y.; Houghton, A. N., A melanosomal membrane protein is a cell surface target for melanoma therapy. *Clin. Cancer Res.* **1996,** *2* (11), 1837-1842.

98. Journe, F.; Boufker, H. I.; Van Kempen, L.; Galibert, M. D.; Wiedig, M.; Salès, F.; Theunis, A.; Nonclercq, D.; Frau, A.; Laurent, G.; Awada, A.; Ghanem, G., TYRP1 mRNA expression in melanoma metastases correlates with clinical outcome. *Br. J. Cancer* **2011,** *105* (11), 1726-1732.

4. EVALUATING RIGOR IN ASSAY & REPORTING IN THE BIOMEDICAL LITERATURE: A CASE STUDY ON A RUN-OF-THE-MILL RESAZURIN ASSAY

José Á. Rodríguez Corrales, Donald H. Clark III, Kristýna Cagašová, Luke R. Vass, Jatinder S. Josan

This chapter has been adapted from a manuscript submitted for publication in the journal *Scientific Reports* and returned due to a conflict with journal scope. The manuscript was reformatted and is ready for submission to *Analytical Biochemistry: Methods in the Biological Sciences* under the "Full-length article" format.

4.1. Abstract

Scientific rigor and reliability are fundamental for the successful extension of *in vitro* and *in vivo* findings of biomedical research into clinical outcomes. Resazurin-based commercial assays, including AlamarBlue®, are widely used to study proliferation and cytotoxicity and have been adapted for high-throughput screening and 3D tissue culture. A comprehensive literature survey on resazurin assays revealed that a glaring percentage of articles failed to report minimal parameters required for replication by other researchers. Furthermore, a vast majority of the articles do not state if assay conditions were validated. We investigated the variability of resazurin reduction assay conditions in seven different cancer cell lines (melanoma, breast, and prostate cancers), and one normal skin cell line. We found that there is no correlation between the type of cell line (i.e., tissue of origin, or cancer *vs.* normal), growth kinetics, or energetic parameters (i.e., ATP and NADH content) and resazurin reduction. We concluded that this assay is variable among cell lines, and thus, care should be taken when extrapolating conditions from studies performed on

other cells. Further, we suggest the minimal parameters required for replicability of this assay and make general comments on the need for higher standards in reporting in the biomedical literature.

4.2. Introduction

Reproducibility in cancer biology research has received substantial attention in the last decade. For instance, a milestone study reported that only 11% of 53 high-profile publications in this area were successfully reproduced due to several factors, including the use of inappropriate preclinical cancer models, problematic endpoints, and non-rigorous testing strategies.[1] The study's authors stressed the need for more rigorous methods and the use of validated reagents in future studies. Similarly, The Reproducibility Project: Cancer Biology, an open-science initiative, has strived to replicate landmark studies in the nominal area using protocols based on peer-review and consultation with authors of the original publications. Aside from poor reproducibility (only 1-2 out of 7 reports were successfully replicated),[2-8] their "clearest" finding was that the methods in the parent publications lacked enough details.[9] The Good Cell Culture Practice Task Force has reinforced the importance of sound methodology in order to allow repetition of the work, while also enabling the target audience to interpret and evaluate any given study.[10] For the specific case of drug evaluation, reproducibility is hampered by the use of improper controls, inappropriate result interpretation, and negligence in reporting experimental conditions, among others.[11-12] Misreporting is one of the easiest issues to resolve and, one that must be undertaken with severity given that its effects on reproducibility are considerable. Specifically, assays would be impossible to reproduce without knowing the conditions used, while any unsuccessful experiments derived from attempts to do so represent a waste of resources and time.

In vitro cytotoxicity evaluation assays are extremely popular in drug screening due to their simplicity and low-cost compared to *in vivo* models. Resazurin, the active component of many commercial assays, including AlamarBlue® and PrestoBlue®, has been used for decades to evaluate bacterial contamination in dairy products,[13] semen quality,[14] and, since the early 1990s, cytotoxicity in somatic cells.[15] Among its many advantages, the assay is reported as non-toxic, simple, inexpensive, homogenous and sensitive.[16] Unlike other methods for the evaluation of viability, the resazurin assay is non-destructive, thereby enabling repeated analysis of the same sets of cells (e.g., same well in a plate), along with further study through other methods (e.g., immunocytochemistry and flow cytometry),[17-18] while keeping at least some biological parameters unaffected.[19] The homogeneous nature of resazurin has permitted its use in high- throughput screening,[20] analysis of cell migration and invasion,[21] and testing of 3D cultures.[22-24]

The resazurin assay relies on its cell-promoted reduction to resorufin, which displays stark differences in its spectroscopic properties compared to the starting material. In particular, the fluorescence of resorufin is substantially stronger than that of resazurin, thus establishing a correlation between fluorescence intensity and the extent of reduction, which is proportional to the number of viable cells.[25] Viability, defined as the proportion of cells in a drug-treated sample (n_s) compared to those in an untreated control (n_c), is calculated as the ratio of fluorescence intensities in a sample (f_s) and an untreated control (f_c) after blank (f_b) subtraction (Eq. 4.1).[26-27]

$$\text{Viability} = \frac{n_s}{n_c} = \frac{f_s - f_b}{f_c - f_b} \qquad \text{(Eq. 4.1)}$$

However, (Eq. 4.1) is only valid if the fluorescence of all samples and controls lie within the linear range of the assay, such that the response factor of the method is constant (Fig. 4.1A). Utilization of a control that is located beyond the limit of linearity leads to an incorrect estimation

of viability, as shown in Fig. 4.1B. Furthermore, both f_s and f_c must be higher than the limit of quantification in order to provide accurate measurements.[28]

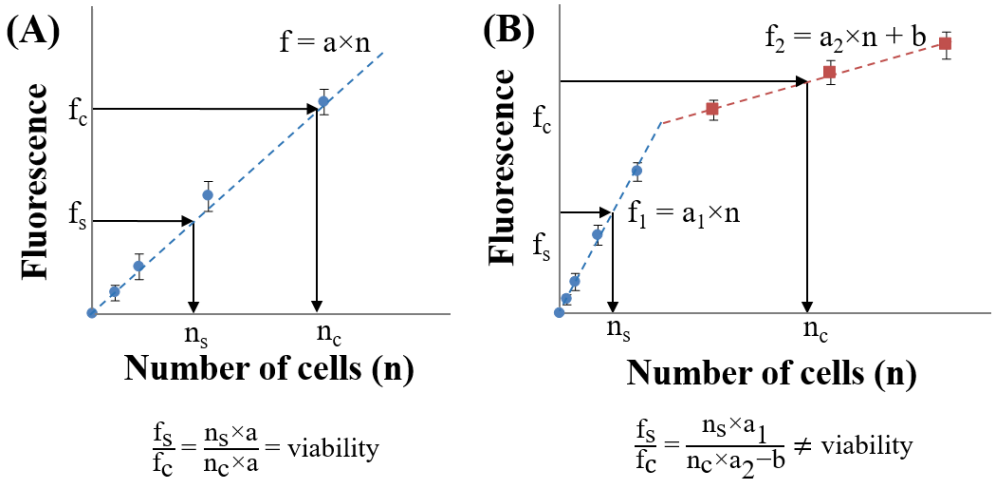


Figure 4.1. Importance of validation of resazurin assay conditions. Sample calibration curve where both sample (s) and control (c) lie within the linear range of the assay, with trendline f (assuming intercept=0), under which conditions Eq. 4.1 is valid (**A**). Sample calibration curve where the control is outside of the linear range. The trendline for the linear range is shown as dashed blue line and represented by equation f_1 (assuming intercept=0), while the plateau is presented as a red dashed line and represented by equation f_2. Under these conditions, Eq. 4.1 is not valid (**B**).

Despite the widespread use of the resazurin assay, specific reduction pathways in cells are poorly understood, at best. Resazurin reduction by sperm cells in semen is enhanced by NADH and decreased by diaphorase inhibitors.[29] Cytosolic, microsomal and mitochondrial cellular fractions catalyze this reaction,[30] which has also been linked to both glycolytic and oxidative metabolism of glucose.[31] Although isolated NADH-oxidoreductase and carnitine dehydrogenase reduce resazurin,[32-33] their role in whole cell models has not been demonstrated. In this context,

determination of the true contributors to resazurin reduction might improve the interpretability and the translation of *in vitro* results into *in vivo* models,[34] especially in studies where mitochondrial activity is impaired or studied.[35]

Herein, we analyze the current reporting standards used for resazurin assays in the scientific literature to determine if they would allow successful replication of experiments. We then studied the relationship between seeding densities and resorufin fluorescence in several cell lines, and analyzed linear ranges, limit of detection, and other parameters required for an effective experimental design. We discuss the correlation between assay conditions and tissue of origin, growth kinetics, and energetic and reductive power, along with evidence pertaining to the role of mitochondrial enzymes in reduction of resazurin by whole cells. Furthermore, we evaluate our results in light of the available literature, and provide suggestions to improve current reporting practices in the scientific community.

4.3. Experimental section

4.3.1. Materials

Resazurin sodium salt, phosphate buffered saline (PBS) pills, and ethylenediaminetetraacetic acid (EDTA) were purchased from Sigma Aldrich. A-375, WM-115, WM-266-4, MCF-7, and BT-474 cell lines, Eagle's Modified Essential Medium (EMEM), Dulbecco's Modified Essential Medium (DMEM), penicillin/streptomycin (100x), and 0.25% trypsin were obtained from American Type Culture Collection (ATCC®). Recombinant insulin was purchased from Life Technologies Corporation. HaCaT, MDA-MB-231, and PC-3 cells were kindly donated by Dr. Lisa DeLouise from University of Rochester Medical Center, Dr. Zhi Sheng from Virginia Tech Carilion Research Institute, and Dr. Donald McDonnell from Duke University.

Fetal Bovine Serum (FBS) was obtained from Corning. Rotenone, reduced nicotinamide adenine dinucleotide (NADH) sodium salt, and adenosine 5'-triphosphate (ATP) sodium salt were purchased from Cayman Chemicals Company. ATP CellTiter-Glo® 2.0, and NAD/NADH-Glo™ assays were obtained from Promega. Guava ViaCount reagent and Check Kit were purchased from EMD Millipore. Statistical and graphical analysis was performed in Microsoft® Excel 2016 and TIBCO® Spotfire® Analyst 7.10.1.

4.3.2. Literature survey

We used the database SciFinder® to identify publications where resazurin was used to quantify the absolute or relative number of cells in a given sample obtained up to 2017. Search inquiries contained either "cytotoxicity", "viability" or "proliferation", and one of three terms describing resazurin, "resazurin", "AlamarBlue" and "Alamar Blue", in order to account for the commercial name of the assay. The results for all searches were combined in SciFinder®, purged for duplicates, and imported into EndNote X8. Automatic removal of duplicates was followed by manual curation of references to exclude assays not performed in human or animal cells.

4.3.3. Preparation of resazurin stock solution

Stock resazurin solution (40 mM) was prepared by dissolving 1 g of resazurin sodium salt in 100 mL of PBS and sterilized by syringe filtration through a 0.22 μm membrane. Working solutions of resazurin were prepared by mixing 180 μL of stock solution with 50 mL of complete medium (final concentration ~14.4 μM).

4.3.4. Cell culture

Experiments involving biohazardous materials, including cell lines, were reviewed and approved by the Institutional Biosafety Committee at Virginia Tech (IBC approval 14-043). All cell lines were grown and handled according to standard protocols. Cells were grown in a 3110 Fisher incubator in a 5% CO_2 aerobic atmosphere at 37 °C, using complete media, comprised of 90% base medium (EMEM for WM-115, WM-2664 and MCF-7; DMEM for A-375, HaCaT, MDA-MB-231 and PC-3; Hybri-Care for BT-474) and 10% fetal bovine serum (FBS), and supplemented with penicillin/streptomycin (1x). MCF-7 medium was further supplemented with insulin (0.01 mg/mL).

Seeding and handling were performed in an A2 Biological Safety Cabinet annually certified by an external vendor. Cells were grown in T75 flasks and subcultured when they reached 80-90% confluency. Detachment was performed using 0.25% trypsin/0.53 mM EDTA for all cell lines, except HaCaT, which required pretreatment with 0.05% EDTA in PBS followed by detachment with 0.05% trypsin/0.025% EDTA. Cells were counted in an InCyto C-Chip improved Neubauer hemocytometer or in a Guava EasyCyte Plus flow cytometer using the ViaCount method, following the manufacturer's instructions in either case. Flow cytometer counts were validated using the Check Kit using the manufacturer's procedure. Cells used in all experiments were of passage 12 or lower to minimize phenotypic drift.

4.3.5. Measurement of resorufin fluorescence

Resorufin fluorescence was analyzed in a Biotek Cytation3 Cell Imaging Multi-Mode Reader, with excitation at 560 nm and emission at 590 nm. Measurements were performed in 100 μL of solution placed in a black polystyrene 96-well plate with clear bottom. Gain was optimized

to obtain high signal to noise ratio while keeping the fluorescence intensity of all samples under the maximum threshold of the instrument.

4.3.6. Determination of limit of detection, limit of quantification, and limit of linearity

Cells were counted using a hemocytometer and seeded on a 96-well plate at a range of densities (15 wells/seeding density) and placed in the incubator. The maximum plating density was adjusted depending on the cell line, with typical values in the range of 100,000-500,000 cells/well, to ensure that at least two densities were located outside of the linear range. After allowing cells to attach for 24 h, media was aspirated and 250 μL of resazurin working solution was added, and the cells were returned to the incubator. After 0.5, 1, 2, 3 or 24 h of incubation, 100 μL of solution (in triplicate) were pipetted from each seeding density, placed in a 96-well plate and measured as described in section 4.3.5. Eight blanks (resazurin solution incubated at 37 °C and 5% CO_2 in the absence of cells) were measured at each time point. Average blank-subtracted fluorescence (f) was plotted versus seeding density at each time point. Limit of linearity was defined as the intercept of the linear range and plateau, with trendline equations $f_1=a_1n$ and $f_2=a_2n+b_2$ respectively.[28] Thus,

$$LL = \frac{b_2}{a_1 - a_2} \qquad \text{(Eq. 4.2)}$$

Furthermore, limits of detection (LD) and quantification (LQ),[28] in terms of cells, were defined as shown below.

$$LD = 3\times\sigma_b/a_1 \qquad \text{(Eq. 4.3)}$$

$$LQ = 10\times\sigma_b/a_1 \qquad \text{(Eq. 4.4)}$$

Where $\bar{f_b}$ and σ_b correspond to the average fluorescence and standard deviation of the blanks at each time point. Finally, linear range was defined as those values between LQ and LL. Any cell

densities lower than LQ for a particular incubation time were removed from the statistical analysis and LD and linear range were reevaluated.

4.3.7. Clustering analysis

In order to observe similarities among assay parameters (LD, LQ, LL) among cell lines, we performed a hierarchical clustering using TIBCO® Spotfire® Analyst. Since parameters obtained at 0.5 h incubation time were not available for all cell lines, these were excluded from the analysis. LD, LQ and LL for each cell line at each time point were directly imported into Spotfire® from an Excel file, Unpivot-transformed and plotted as a heat map. The dendrogram was developed using Unweighted Pair Group Method with Arithmetic Mean (UPGMA) hierarchical clustering for Euclidean distance. BT-474 was excluded from this analysis due to the limited data available.

4.3.8. Kinetics of cell growth

Selected cell lines were seeded in a 96-well plate at a density of 5,000 and 10,000 cells/well and grown under standard conditions (see section 4.3.4). After 24 h, the medium was aspirated, cells were washed twice with PBS and detached by 5 min incubation at 37 °C and 5% CO_2 in 0.25% trypsin/0.53 mM EDTA, in triplicate. 50 µL of the cell suspension was mixed with 100 µL of ViaCount reagent, incubated in the dark for 5 min, and analyzed in a GuavaCyte Plus flow cytometer using the conditions recommended by the manufacturer. This procedure was repeated on separate triplicates of cells two, three, and four days after seeding. Average number of cells/well at each time point (n_t) was normalized to the number of cells recovered on day 1 (n_1), and plotted against time of growth (t). Linear least squares fitting was performed, and population doubling

time (DT) was calculated by evaluating ln(2) multiplied by the reciprocal of the slope, according to the model presented in (Eq. 4.5).[36]

$$\ln\left(\frac{n_t}{n_1}\right) = \frac{\ln(2)\times t}{DT} \qquad \text{(Eq. 4.5)}$$

4.3.9. Cytotoxicity of resazurin

Cells were seeded at 5,000 cells/well in a 96-well plate and incubated under standard conditions. One triplicate of each cell line was incubated in complete medium for 3 days and served as a control. Separate triplicates were incubated in medium for 1 or 2 days, and resazurin working solution in medium for 2 or 1 day respectively. At this point, cells were rinsed with PBS and fresh resazurin solution was added. Fluorescence was measured after a 2-hour incubation using the conditions described in Section 4.3.5; fluorescence intensity was blank-subtracted and normalized to the control.

4.3.10. ATP and NAD(H) cell content

ATP and NADH cell content was assayed to provide insight into cellular metabolism. 5,000 cells/well were seeded in a black 96-well plate and incubated under standard conditions. After 24 h, the medium was aspirated, cells were rinsed twice with PBS, and PBS was added. At the same time, standard solutions of ATP and NADH were prepared from adenosine 5'-triphosphate sodium salt and reduced nicotinamide adenine dinucleotide sodium salt by dissolving an appropriate amount of solid in PBS. A calibration curve ranging from 10 nM to 13 mM ATP was prepared and 100 μL of each standard was loaded into a black 96-well plate. Similarly, a calibration curve was prepared for NADH except the highest standard used was 1 mM and only 50 μL of sample was loaded into the plate.

Luminescence assays were performed using ATP CellTiter-Glo® 2.0 and NAD/NADH-Glo™ kits following manufacturer's instructions. Briefly, 100 μL of CellTiter-Glo® or 50 μL of NAD(H)-Glo detection reagent was added to the corresponding standards, placed in an orbital shaker and incubated as recommended. Luminescence intensity was measured in a Biotek Cytation3 Cell Imaging Multi-Mode Reader, set to a gain of 160 and 1 s integration. The ATP assay was linear up to the highest standard tested, while the NAD(H) assay lost linearity at concentrations higher than 500 nM; thus, only the standards under 500 nM were used for quantification. Cells were treated under the same conditions as the standards, and ATP and NAD(H) content/cell were calculated using the calibration curves. All luminescence measurements of the cell samples were within the linear range of the assays.

4.3.11. Effect of mitochondrial inhibitors in resazurin bioreduction

In order to identify reductases potentially involved in resazurin reduction, co-incubation with mitochondrial inhibitors was performed. Complex I inhibitor rotenone[37-38] and complex IV inhibitor cyanide[39-41] were used. 10,000 cells/well were seeded in a 96-well plate and incubated for 24 h, except for BT-474, which required 48 h for complete attachment to the plate. At this point, media was aspirated and 100 μL of resazurin solution with 0.1-100 μM inhibitor in medium was added. Resazurin solution without inhibitors was used as a control. Blanks were prepared by incubating inhibitor/resazurin or resazurin solutions in the absence of cells, in order to correct for the direct reaction of dye with inhibitors. Fluorescence was measured on the plates immediately after addition of solutions and then every 30 min for 3 h. Cells were returned to the incubator between the measurements. This procedure is broadly based on previous reports involving co-incubation of resazurin with inhibitors.[27, 29, 42]

4.4. Results and discussion

4.4.1. Literature survey

The SciFinder® search engine was used to interrogate the literature for articles where the resazurin assay was used to evaluate the absolute or relative number of cells in a given sample. Upon duplicate removal in SciFinder® and EndNote X7, followed by manual curation, the final library contained 2,157 records, for which the time distribution is shown in Fig. 4.2A. Clearly, use of the resazurin assay to estimate cell number has increased substantially since its first report in 1993,[15] perhaps due to the advantages aforementioned. We then proceeded to determine if the identified publications reported three basic parameters that would be essential for their reproduction by other scientists: cell seeding density, concentration of resazurin (or volume/dilution of commercial solution) and time of incubation with resazurin. Given the vast number of publications, a sample of 252 (11.6% of the total) was selected randomly based on availability, while keeping the publication year distribution very similar to that of the library (see Fig. S4.1).

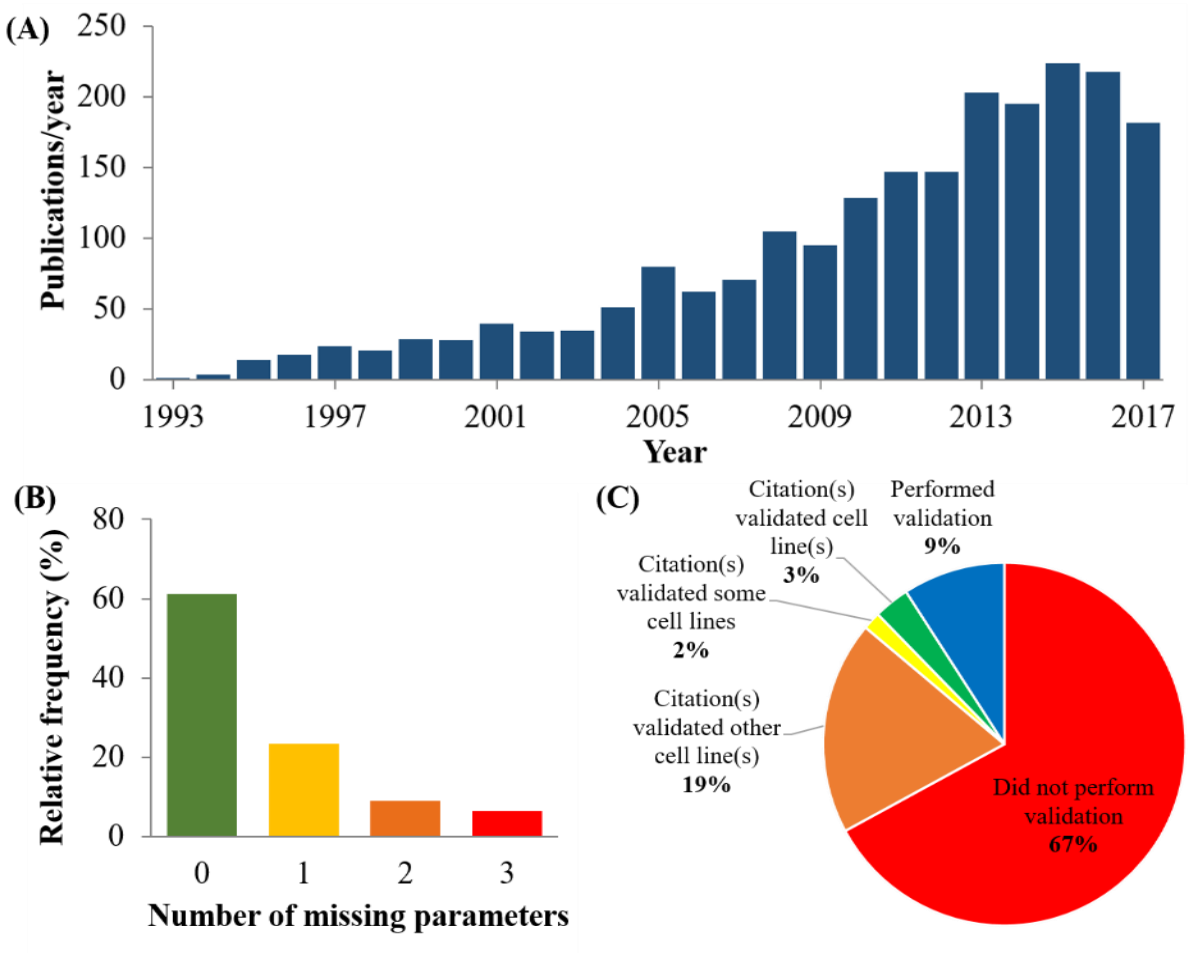


Figure 4.2. Results of the literature survey performed in SciFinder® concerning the resazurin assay in animal and human cells. Number of publications per time range (**A**, n=2157), percentage of publications missing assay parameters (**B**) and validation (**C**) for a subset of the total references (n=252, 11.6%).

Despite the overwhelming use of the resazurin assay, major reporting issues were identified. Nearly 40% of the publications failed to report at least one of the assay parameters; while 15% omitted at least two. Only 61% of the publications specified the three essential parameters in their methods (Fig. 4.2B). An even more crucial factor is assay validation prior to its use in drug screening and assessment of other metabolic end-points. A host of issues such as

cellular phenotypic changes with passage numbers,[43] and cell line contamination with microorganisms such as mycoplasma that can alter metabolic profile,[44] among others, can introduce confounding variables that necessitate assay validation before committing to drug screens. To our surprise, only 14% of publications reported validation of all or at least some of the cell lines used. Most publications rarely acknowledged if any form of method validation was conducted, while in some cases referring to other articles or the manufacturer for the assay conditions, with the latter often including generalized assay conditions that are cell line agnostic. Nearly 65% of publications did not carry out (or at least did not report in the main article or supplementary section) any validation for the assay.

Unexpectedly, 21% of the publications cited articles where the assay was validated using other cell lines. This is particularly troublesome in instances when either the cell line reported in the referred articles was different since, as we will discuss later, the linear range for some cell lines can be different and narrow, which can present problems in obtaining reliable assay results. Although some publications stated that "optimal" assay conditions were used, it was uncommon to find the graphical or algebraic analysis used to determine the linear range of the method[16, 25, 45-46] or the required incubation times.[21, 45, 47-49]

While the results of this survey do not invalidate the findings presented in these publications, they do reinforce the need for improved reporting standards in the scientific literature. Moreover, this laxity in reporting is pervasive in the scientific literature, particularly for cancer biology, and is currently under high scrutiny by several organizations including The Reproducibility Project, Good Cell Culture Practice Task Force, and other open science initiatives (*vide supra*). In this context, our report highlights how even a "simple" drug screening test such as the resazurin assay can introduce substantial variability, and thus, lead to irreproducibility. Such

erroneous results could also lead to further complications when the results are translated into *in vivo* settings.

4.4.2. Analytical parameters for resazurin assay in human cell lines

We selected several commonly used cell lines for studying the analytical parameters of this assay: three melanoma cell lines (WM-115, WM-2664, and A-375), a normal skin line (HaCaT), three breast cancer cell lines (MCF-7, BT-474, and MDA-MB-231), and one prostate cancer cell line (PC-3). In order to assess the variability of the resazurin assay conditions among different cell lines, we determined the linear range of the method under several incubation times. As expected for a cell-promoted reaction, the extent of resorufin production was proportional to the seeding density (Fig. 4.3). For most cell lines, the fluorescence intensity of the samples was higher than the limit of quantification after ~30 min, with the widest linear ranges obtained at 2-3 h of incubation.

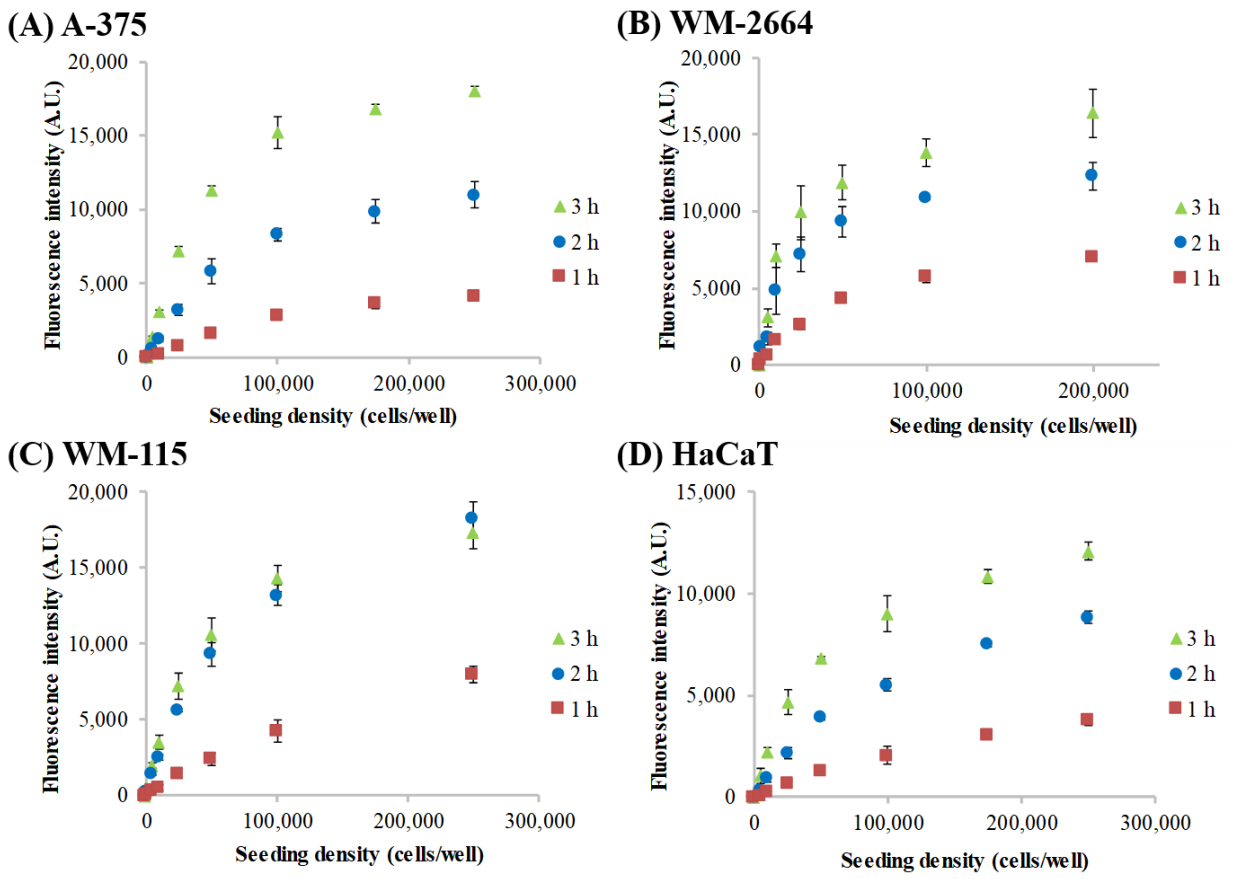


Figure 4.3. Resorufin fluorescence intensity (excitation=560 nm, emission=590 nm) *vs.* cell seeding density for (**A**) A-375, (**B**) WM-266-4, (**C**) WM-115, and (**D**) HaCaT cell lines, measured with Biotek Cytation3 Cell Imaging Multi-Mode Reader after 1 h (red squares), 2 h (blue circles) or 3 h (green triangles) of incubation with 14.4 µM resazurin solution. Error bars correspond to the standard deviation (n=3). All values were blank-corrected prior to plotting. Other incubation times were also measured (see Table 4.1) but not included in the graphs for clarity.

More importantly, the fluorescence intensity increased with incubation time, which confirms that the sensitivity of the assay can be readily enhanced by lengthening the incubation period (see Tables 4.1 and S4.1). For example, the LQ for A-375 cells decreased from ~7,700 cells at 1 h incubation to ~380 cells at 24 h under the conditions studied. However, this advantage comes

with a caveat. The extended incubation time also results in diminishing the limit of linearity from 79,000 at 1 h to 9,900 cells at 24 h, presumably due to the entire consumption of resazurin present in the media with even as few as ~10,000 cells. This finding clearly highlights how variations in assay conditions can significantly alter the linear range under which the assay remains optimal. The results obtained for the studied cell lines are summarized in Tables 4.1 and S4.1. Further, it must be noted that the reported ranges are based on the initial seeding densities, and not the actual number of cells at the end of the assay. Longer incubation times with resazurin can trigger a cytotoxic effect in some cell lines, as will be discussed later.

Table 4.1. Linear range of the resazurin assay for several human cell lines.

Cell line	**Incubation time (h)**				
	0.5	**1**	**2**	**3**	**24**
HaCaT	*	9,600-65,000	4,600-52,000	2,100-33,000	720-8,100
WM-115	17,000-72,000	6,000-74,000	2,600-45,000	1,100-36,000	380-5,600
WM-2664	7,100-27,000	1,500-20,000	860-11,500	450-11,000	44-1,100
A-375	17,000-85,000	7,700-79,000	1,200-61,000	1,400-66,000	380-9,900
MCF-7	*	7,200->200,000	2,600->150,000	2,300-100,000	2,400-61,000
PC-3	16,000-59,000	8,300-59,000	3,600-46,000	680-26,000	360-3,300
MDA-MB-231	7,900-160,000	12,000-100,000	5,300-77,000	1,100-64,000	220-4,800
BT-474	**	**	4,200-110,000	6,600-130,000	**

* Low signal-to-noise ratio hindered evaluation of these values.

** Not evaluated.

The most relevant observation from the optimization performed above is that assay conditions are cell-dependent. Compare with other studied cell lines, the melanoma cell line WM-

266-4 consistently displayed the lowest LQ and LL at all incubation times, which we attribute to fast metabolic processes that catalyze the reduction of resazurin and generate a quick depletion of the reagent from the media. At the other extreme are the breast cancer cell lines, MCF-7 and BT-474, which displayed significantly slower resazurin reduction kinetics and wider linear ranges, even higher than 200,000 cells/well at some incubation times. Thus, for the purpose of this discussion, we classified MCF-7 and BT-474 as "slow" resazurin reducing cells.

Clustering analysis of the cells lines using the resazurin assay parameters confirmed the strikingly different behaviors. Fig. 4.4 shows that MCF-7 and WM-266-4 possess the highest dissimilarity compared to the other cell lines, which agrees with the qualitative analysis presented above. The remaining cells lines in our cohort behaved similarly to each other. WM-115, A-375 HaCaT, PC-3, and MDA-MB-231 displayed LLs of 45,000-75,000 cells/well at 2 h of incubation. Since LLs for these cells lines is consistently between those of "fast"-reducing WM-266-4 and "slow"-reducing MCF-7 and BT-474, we denoted them as "medium"-reducers. It is noteworthy that this category is comprised of both cancer and normal cells of the same lineage (e.g., melanoma WM-115, A-375, and immortalized keratinocyte HaCaT cell lines), and cells derived from different tissues, including prostate (PC-3) and breast (MDA-MB-231). Furthermore, clustering analysis showed the highest similarity between PC-3 and HaCaT cells, which stem from different tissue types; whereas the former is metastatic, the latter is a healthy immortalized cell line. Our results agree with former reports on the resazurin assay that have found cell-dependent linear limits.[16, 26, 50]

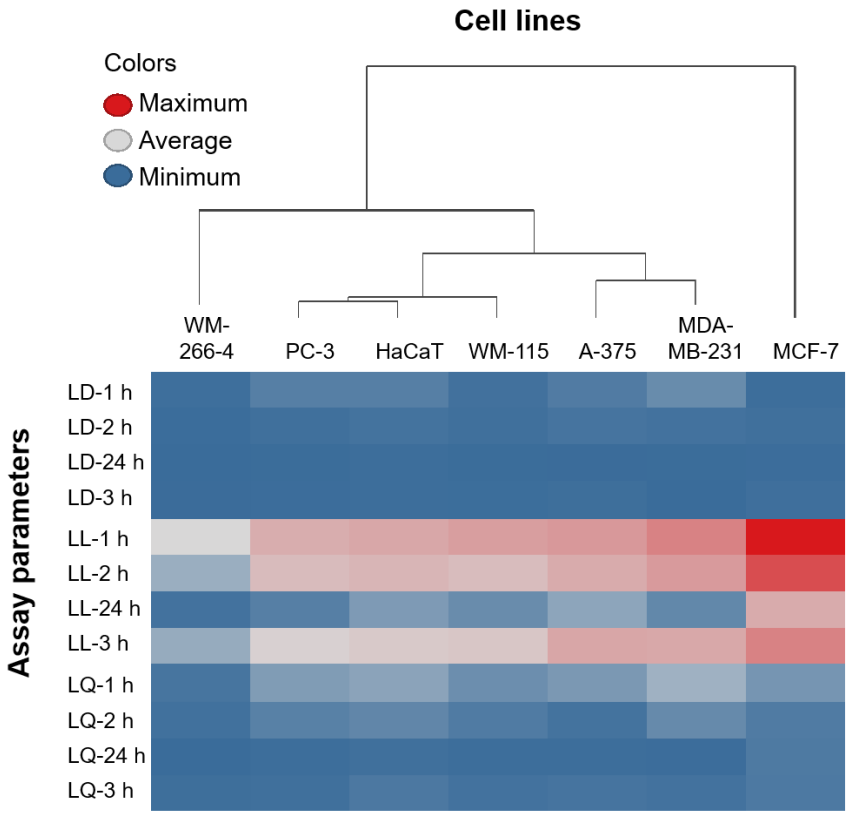


Figure 4.4. Unsupervised clustering of cell lines based on resazurin assay parameters. Clustering is performed based on Euclidean distance using limits of detection (LD), quantification (LQ) and linearity (LL) at 1, 2, 3, and 24 h.

It is noteworthy that seeding densities, although extensively used in the literature, do not necessarily correspond to the number of cells at the time of the resazurin assay for two main factors. First, some of the cells that are seeded might not attach to the plate, and second, lag time and kinetics of growth might differ between cell lines, leading to differences in the actual number of cells at the time of the assay. Similarly, resazurin assays performed over long periods (e.g., 24 h) are inherently influenced by cell growth and division. Thus, we caution the scientific community to question the seeding densities reported by others or provided in vendor protocols. Rather, it is essential to conduct the validation first-hand in order to determine the optimal conditions for the

assay. Further, we urge the readers to report all such conditions tested in the supplementary information, thereby enabling others to reproduce the assay or extrapolate to their own experimental design.

4.4.3. Long-term cytotoxicity studies with resazurin

One of the advantages reported for the use of resazurin in drug evaluation is its low inherent toxicity, which avoids overestimation of toxicity due to resazurin itself when conducting screens for anti-proliferative agents. We studied this behavior in a subset of cell lines and found that 2 h exposure to resazurin does not generate a detectable cytotoxic effect in WM-266-4 or A-375 cells (Fig. S4.2). In contrast, longer-term exposure to resazurin (1 or 2 days) remarkably affected cell health. Four cell lines, A-375, HaCaT, WM-115, and WM-266-4, were grown in complete media for a total of 3 days, and media was replenished every day. A subset of the samples and supplemented with resazurin for either 0 (control), 24, or 48 h. At the end of the incubation time, cells were rinsed, the resazurin assay was performed, and resorufin fluorescence was measured after 2 h (Fig. 4.5). Most of the cell lines pre-exposed to resazurin generated less fluorescence than the controls, which we attribute to the toxicity of either resazurin or resorufin, or the depletion of reducing agents (e.g., NADH) present in the cells. Interestingly, the cytotoxic effects of long-term exposure to resazurin are cell dependent, since A-375 pre-exposure for up to 2 days did not change its ability to catalyze the formation of resorufin, the reason(s) for which are beyond the goals of this study.

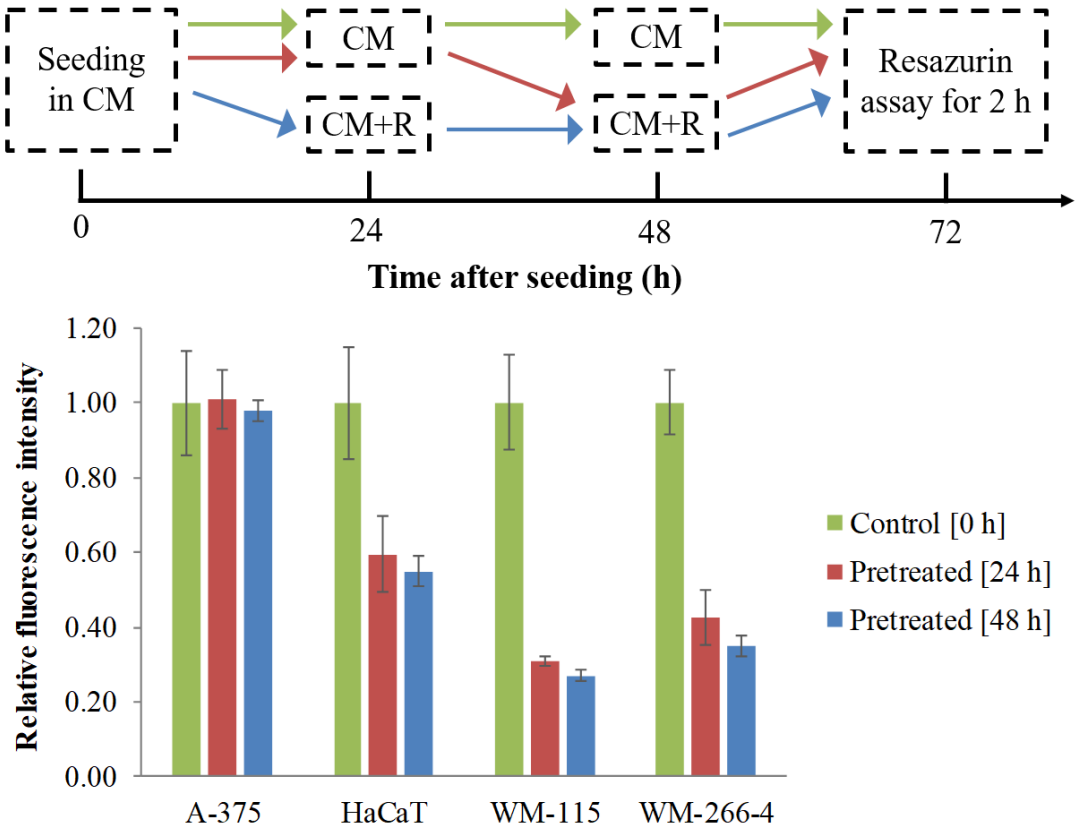


Figure 4.5. Long-term cytotoxicity of resazurin. All samples were tested 72 after seeding. Controls cells were grown in complete medium (CM) for the entire duration of the experiment, whereas media in the other samples was supplemented with resazurin (CM+R) 24 h or 48 h prior to the end of the 72 h growth period. At this point, fresh resazurin solution was supplied to samples and controls and fluorescence was measured after 2 h. Values for each cell line were normalized to the controls that were grown in the absence of resazurin. Error bars correspond to standard deviation (n=3).

Our results agree with previous reports that have shown resazurin toxicity in A2780-cp-20 ovarian carcinoma[51] under similar conditions, and in MCF-7 breast cancer and 3T3-L1 mouse fibroblasts as early as 4-8 h exposure.[52] In contrast, a 24 h treatment of cytotoxic T lymphocytes did not decrease their viability.[18] Our results suggest that any long-term cytotoxicity studies

performed on resazurin-sensitive cells such as WM-266-4, WM-115 and HaCaT, where these cells are exposed to resazurin for days, would lead to inaccurate results and likely an overestimation of the potency of the tested drug.[53] This finding further supports our argument regarding the need for resazurin optimization and validation for each cell line. If extended incubation times are required (e.g. when cell number is low), then the inherent cytotoxic effects of resazurin must be considered.

Furthermore, cytotoxicity studies require special considerations due to cell growth during the testing period. The correlation between cell number and fluorescence discussed above was determined 24 h after seeding, and thus, data findings are specific to the number of cells at that particular time point. In contrast, cytotoxicity assays require extra time allocated for incubation with the drug (e.g. hours to days) after seeding/attachment and prior to testing with resazurin. In this context, we suggest that assay conditions should be refined based on the cell line (accounting for linear range and kinetics of growth) and drug toxicity, such that the number of cells in all samples and controls at the time of resazurin assay falls within the determined linear ranges. We recently proposed a protocol for fine-tuning the assay for this purpose.[54]

4.4.4. Correlation between analytical parameters and cell metabolism

As discussed above, the resazurin assay parameters were cell line-dependent. Since reduction is a cell-catalyzed process, we sought to explore if the difference between fast-, medium-, and slow-reducing cells could be correlated to metabolic state. In principle, we hypothesized that faster dividing cells could have a faster metabolism, thereby enabling them to reduce resazurin more quickly. This hypothesis was proven wrong, however, since we observed that the fast-reducing cell line WM-266-4 displayed a doubling time (DT) of 20.3±2.5 h, which was in between

the DTs of two medium-reducing cell lines, A-375 (14.7±1.9 h) and WM-115 (36±10 h) (Fig. 4.6A). Thus, doubling time did not correlate to resazurin metabolic reduction kinetics.

NADH, the concentration of which relates to the reductive power of the cell,[55] serves as the electron source in resazurin reduction in isolated enzymes and semen.[29, 32] Thus, we explored if NAD(H) content could be correlated to reduction kinetics. Although fast-reducing WM-266-4 showed the highest NAD(H) content per cell, this value is within error of medium-reducing PC-3 cells. Further, NAD(H) level in the slow-reducing MCF-7 cells was found to be in between those of HaCaT and MDA-MB-231, which are medium-reducers. Inspection of the energetic content, when measured as ATP content, also did not show any correlation with resazurin reduction kinetics. These results for NAD(H) and ATP quantification are presented in Fig. 4.6.

The fact that neither kinetics of growth, reductive power, or energetic power, measured through doubling time, and NADH and ATP content, did not correlate with the kinetics of resazurin reduction highlights the dangers with *a priori* assumptions with respect to extrapolated assay parameters, thus reinforcing the importance of conducting assay optimization independently of cell type. Indeed, throughout our assays we found that no one factor could be used to predict resazurin assay conditions for a particular cell line.

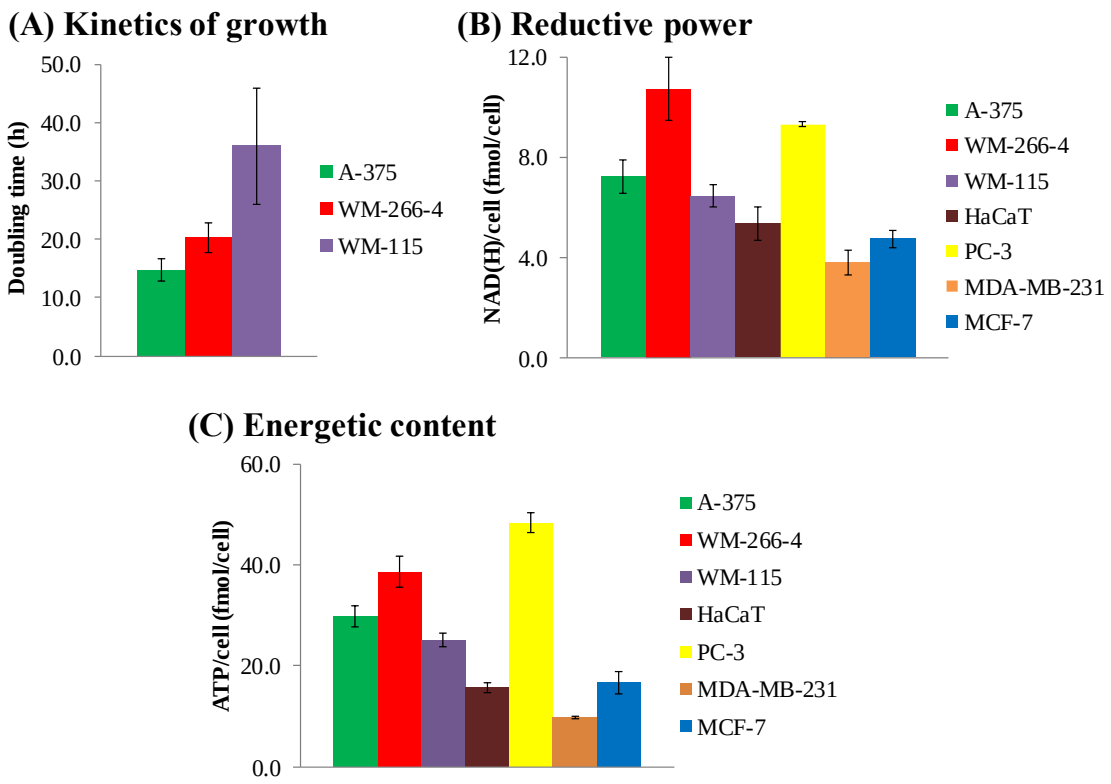


Figure 4.6. (**A**) Kinetics of growth (measured as doubling time), (**B**) Reductive power as NAD(H) content and (**C**) Energetic or ATP cellular content determined for several cell lines. Error bars correspond to the standard deviation (n=3).

4.4.5. Resazurin reduction metabolic pathways might differ between cell lines

Although isolated NADH-oxidoreductase, also known as mitochondrial complex I,[32] and carnitine dehydrogenase[33] reduce resazurin *in vitro*, the actual enzymes responsible for this process in whole cells have not been conclusively identified. Thus, in order to obtain insight into the metabolic pathways that lead to resazurin reduction, we investigated the role of mitochondrial complexes due to their relevance in cellular respiration and previously theorized role of mitochondrial reductases in the formation of resorufin.[25] We hypothesized that if resazurin reduction occurred at one of the complexes, the addition of its corresponding inhibitor would

hinder the formation of resorufin in a dose-dependent manner. Accordingly, we incubated WM-266-4, A-375, WM-115, MCF-7, and BT-474 with resazurin in the presence of either rotenone or KCN, which are inhibitors for complexes I and IV, respectively. After 2 h co-incubation, fluorescence was measured and normalized with respect to controls treated with resazurin only.

Interestingly, observed results differed greatly among cell lines (Fig. 4.7). In particular, KCN inhibited resazurin reduction in WM-266-4 (fast), BT-474 (slow) and WM-115 (medium), but not in MCF-7 (slow) or A-375 (med). In contrast, treatment with rotenone only showed a clear dose-response pattern in BT-474, whereas the other cell lines did not show any response. Intriguingly, the highest concentration of compound (100 μM) consistently enhanced the fluorescence in all cell lines up to levels at or above those of controls incubated with resazurin alone. Given that rotenone is known to react with resazurin in the absence of cells to produce enhanced fluorescence,[35] all measurements were corrected using blanks with resazurin and the inhibitors at the corresponding concentrations prior to normalization. Thus, the increase in fluorescence could stem from a biological response (e.g., a feedback loop) that is unknown and beyond the scope of this study.

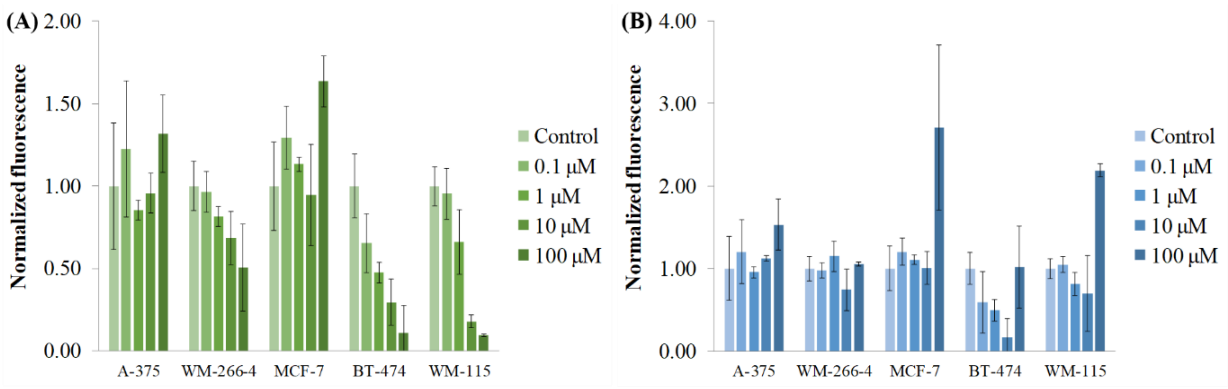


Figure 4.7. Dose-response curves for the treatment of different cell lines with mitochondrial inhibitors, (**A**) KCN, or (**B**) rotenone. All fluorescence values were blank-corrected using solutions of resazurin and inhibitor at each concentration and then normalized using controls with cells that were not exposed to the inhibitors. Error bars correspond to the standard deviation (n=3).

Overall, the results discussed herein suggest that complex IV participates in resazurin reduction, but that it is not the predominant reducing enzyme in all cell lines. This is consistent with the results of prior studies that have shown the involvement of complex IV in resazurin reduction by corneal endothelial cells,[56] whereas complexes I and II appear to be involved in neural tissue,[57] and inhibitors of complexes I and IV did not affect metabolism of resazurin in PC12 cells.[42] Thus, it seems apparent that resazurin reduction might involve several enzymes, some which could have a subcellular localization other than mitochondria. This finding is in line with a study where resazurin reduction was promoted by cytosolic, mitochondrial and microsomal enzymes.[30] Thus, caution must be exercised in using this assay as a health endpoint of a specific metabolic process or organelle such as mitochondria.[35]

4.5. Conclusions

Our literature survey revealed the pervasive underreporting of assay parameters for as basic as a cell proliferation and cytotoxicity assay, thus reinforcing the earlier calls for improving rigor in journal reporting to facilitate reproducibility.[1] This is not only true for more complex and system specific assays, but also for the run-of-the-mill workhorse assays such as cell proliferation, and viability/cytotoxicity measurements. Indeed, analysis of a substantial sample of the available literature showed that, despite the widespread and growing use of resazurin-based assays to estimate absolute or relative cell numbers, important parameters needed for experimental replication (e.g., incubation time, or resazurin concentration) were not reported. Furthermore, most authors failed to acknowledge if the said parameters or the conditions under which the experiments were performed were validated in their laboratories, or adopted from the literature.

Extrapolation from the literature without in-house validation can lead to erroneous results. The present study used LL, LQ, and LD to show how some cell lines have very narrow linear ranges, which can change under different experimental conditions even for the same cell line. In the present age when page limit is no longer a concern due to digital storage of supplementary information, it is hard to rationalize the omission of critical data on assay parameters. Further, we argue that careless "recycling" of previously published methods tends to introduce lapses in reporting when, in fact, a detailed methodology is critical for other researchers who wish to replicate a given study. Hence, each study must be carefully vetted. We cannot emphasize this point strongly enough given this host laboratory's experience in reproducing this assay from a published protocol.

The most relevant observation from these optimization studies is that the requisite assay conditions are heavily cell-dependent, with some cells being very rapid reducers of resazurin,

while others are slow. Accordingly, the limits of linearity (LL) vary between cell lines, with fast-reducers showing a lower LL, i.e., the linear range under which the experiment could be performed is narrow. Further, the limits also vary within the same cell line when long-term incubation with resazurin was used for increased sensitivity, which in turn decreased the limit of quantification but also decreased the linear range (Fig. 4.3 and Table 4.1). This result should caution the researcher to not indiscriminately fix the seeding density for every resazurin experiment, but rather to account for the varied assay conditions (e.g., exposure to drugs, inclusion of agonists and inhibitors in the same screen, and incubation time), even when using the same cell line. Further, such variations in conditions and the validation of linear range under different conditions must be reported in detail in the supplementary information and cannot be generally assumed to encompass a generic resazurin assay protocol.

It has been reported that resazurin has low inherent toxicity towards cells. Here we find that while this assertion may hold true for short-term exposure of few hours, any experimental conditions longer than a few hours introduces confounding variables from the inherent toxicity of resazurin. From our tested cohort, while A-375 remained unaffected at all incubation times (up to 2 days), the non-metastatic WM-115 saw more than 70% reduction in cell number within 24 hours (Fig. 4.5). The normal skin line HaCaT and the metastatic cell line WM-2664 were also affected but less than WM-115. It is interesting to note that both WM-2664 and WM-115 cells were isolated and immortalized from the same patient,[58] but different growth kinetics and limits of linearity.

We studied the kinetics of resazurin reduction by studying the doubling time, NADH, and ATP content. No clear trends were observed between the reduction rate and reductive power and energy. It has been suggested that the mitochondrial respiratory chain enzymes are responsible for the resazurin reduction.[46] Thus, we looked at the cellular basis of this process by investigating the

effect of mitochondrial respiration inhibition. Our results suggest that the degree of involvement of various reductases is variable between different cell lines, with WM-115 and BT-474 showing the highest sensitivity to Complex I inhibition; in contrast, BT-474 was only sensitive to Complex IV inhibition. After correcting for rotenone-mediated direct resazurin reduction,[35] our results indicate that the mitochondrial reductase activity does not solely explain resazurin reduction. This ambiguity is likely due to the involvement of several enzymes, some of which could have higher or lower activity across cell lines. Further, the unperturbed reduction activity, despite mitochondrial inhibition, suggests that some cell lines reduce resazurin predominantly through the cytosolic fraction of reductases.

In conclusion, our results agree with several prior studies that showed the possibility of estimating cell viability using the extent of cell-catalyzed resazurin reduction. We expect that the methods described herein for determining linear range, kinetics of growth, and inherent resazurin toxicity, will help researchers in the design of experiments where the fluorescence measured after the assay is accurately correlated with viability and thus cytotoxicity. However, our overall findings reinforce the high variability of reduction kinetics across different cell lines, which we were unable to correlate with tissue of origin, normal versus metastatic character, or other cellular metabolic parameters, including kinetics of growth, and reductive or energetic power. This outcome supports a need for regular assay validation for any new set of experimental conditions employed. Caution must also be exercised when interpreting the resazurin reduction in the context of mitochondrial health. Failure to conduct such experiments can lead to erroneous results. Most importantly, negligence to implement rigor in validation and reporting assay conditions can lead to unintended assumptions on the part of the repeater (whether in the same or different labs), or worse still, significant wastage of time and funding.

4.6. Acknowledgments

Authors would like to thankfully acknowledge funding from the Department of Chemistry at Virginia Tech (grant number 119744) and VT-ICTAS (grant number 175911).

4.7. Supplementary information

4.7.1. Time distribution of analyzed literature

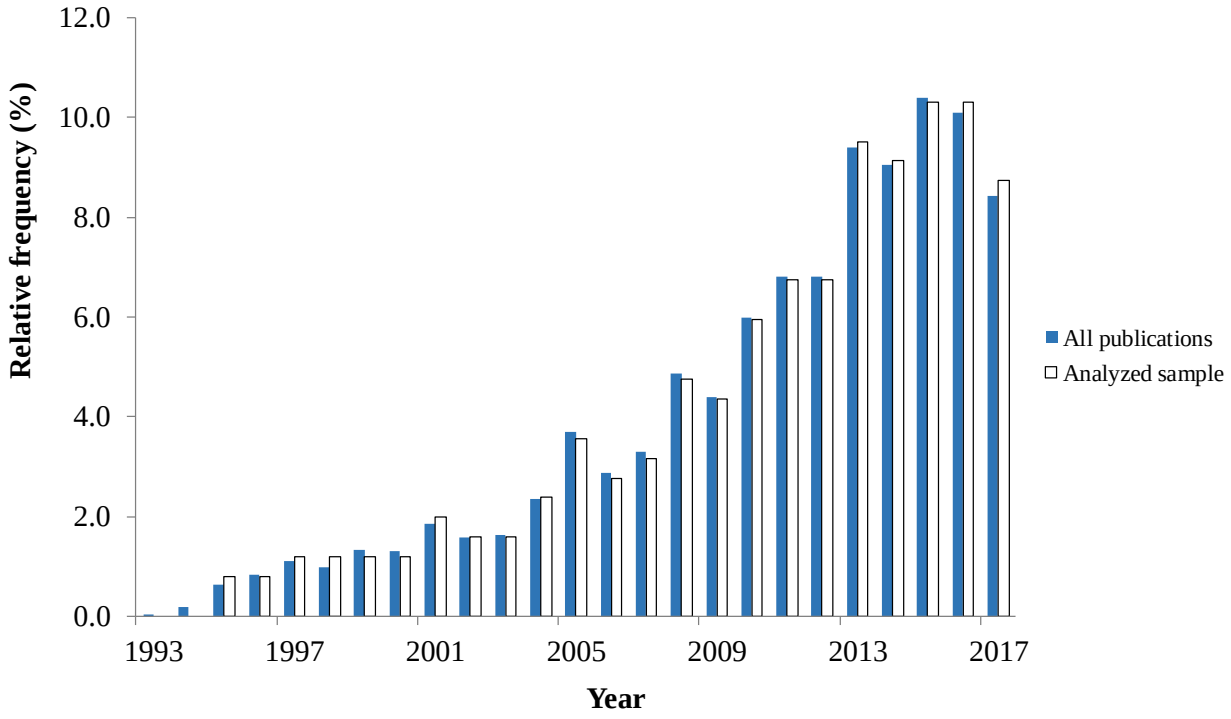


Figure S4.1. Time distribution of all publications identified (blue) using Scifinder® with respect to the use of Alamar Blue for animal and human cell lines. Distribution of publications analyzed for reporting parameters and validation of assay conditions are shown in light blue.

4.7.2. Assay validation and analytical parameters for selected cell lines

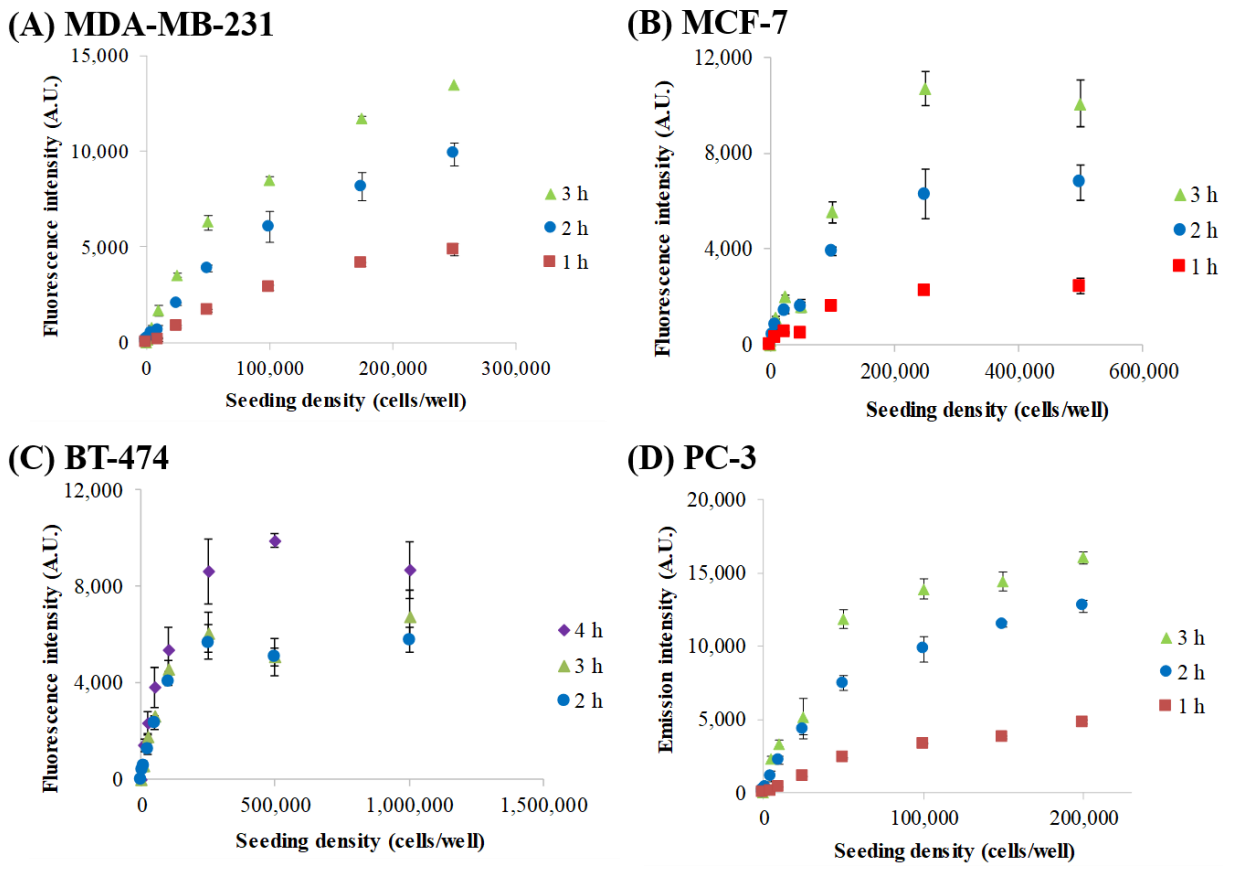


Figure. S4.2. Resorufin fluorescence intensity (excitation=560 nm, emission=590 nm) *vs.* cell seeding density for (**A**) MDA-MB-231, (**B**) MCF-7, (**C**) BT-474, and (**D**) PC-3 cell lines, measured with Biotek Cytation3 Cell Imaging Multi-Mode Reader after 1 h (red squares), 2 h (blue circles) or 3 h (green triangles) of incubation with 14.4 μM resazurin solution. Error bars correspond to standard deviation (n=3). All values were blank-corrected prior to plotting. Other incubation times were also measured (see Table 4.1) but not included in the graphs for clarity.

Table S4.1. Limit of detection (in cells) for the resazurin assay in studied cell lines.

Cell line	Incubation time (h)				
	0.5	1	2	3	24
HaCaT	*	3,400	1,200	350	330
WM-115	3,900	900	690	280	100
WM-266-4	1,700	480	180	82	6
A-375	5,500	2,700	1,400	460	81
MCF-7	*	420	700	610	260
PC-3	4,900	3,400	750	200	110
MDA-MB-231	3,200	5,500	1,100	48	98
BT-474	**	**	1,300	2,000	**

* Value cannot be determined based on gathered data

** Not determined

4.7.3. Short-term exposure of cells to resazurin

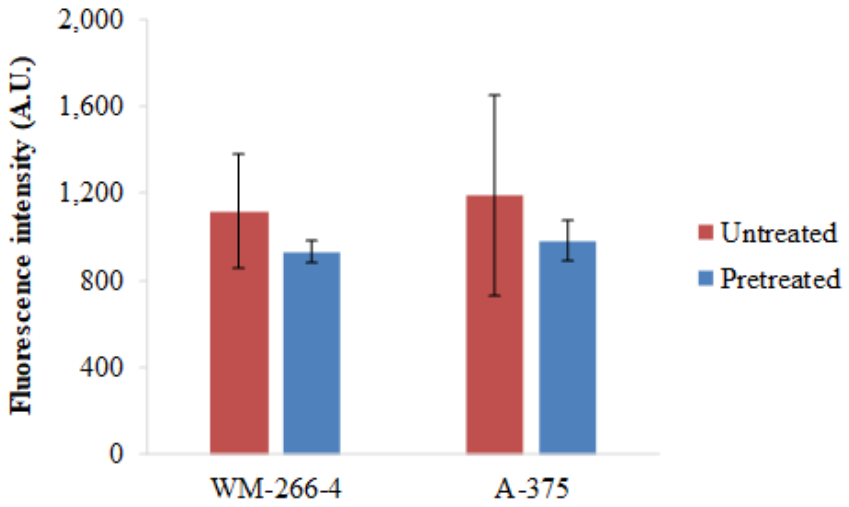


Figure S4.3. Cells treated with resazurin for 2 h, grown in media for 24 h and exposed to resazurin for 2 h. Controls were not pretreated with resazurin. Error bars correspond to standard deviation (n=3).

4.8. References

1. Begley, C. G.; Ellis, L. M., Drug development: Raise standards for preclinical cancer research. *Nature (London)* **2012,** *483* (7391), 531-533.

2. Aird, F.; Kandela, I.; Mantis, C.; Reproducibility Project: Cancer Biology, Replication Study: BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *eLife* **2017,** *6*, e21253.

3. Mantis, C.; Kandela, I.; Aird, F.; Reproducibility Project: Cancer Biology, Replication Study: Coadministration of a tumor-penetrating peptide enhances the efficacy of cancer drugs. *eLife* **2017,** *6*, e17584.

4. Kandela, I.; Aird, F.; Reproducibility Project: Cancer Biology, Replication Study: Discovery and preclinical validation of drug indications using compendia of public gene expression data. *eLife* **2017,** *6*, e17044.

5. Shan, X.; Fung, J. J.; Kosaka, A.; Danet-Desnoyers, G.; Reproducibility Project: Cancer Biology, Replication Study: Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *eLife* **2017,** *6*, e25306.

6. Horrigan, S. K.; Courville, P.; Sampey, D.; Zhou, F.; Cai, S.; Reproducibility Project: Cancer Biology, Replication Study: Melanoma genome sequencing reveals frequent *PREX2* mutations. *eLife* **2017,** *6*, e21634.

7. Horrigan, S. K.; Reproducibility Project: Cancer Biology, Replication Study: The CD47-signal regulatory protein alpha (SIRPa) interaction is a therapeutic target for human solid tumors. *eLife* **2017,** *6*, e18173.

8. Showalter, M. R.; Hatakeyama, J.; Cajka, T.; VanderVorst, K.; Carraway, K. L., III; Fiehn, O.; Reproducibility Project: Cancer Biology, Replication Study: The common feature of leukemia-

associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *eLife* **2017,** *6*, e26030.

9. Baker, M.; Dolgin, E., Cancer reproducibility project yields muddy results. *Nature (London)* **2017,** *541* (7637), 269-270.

10. Coecke, S.; Balls, M.; Bowe, G.; Davis, J.; Gstraunthaler, G.; Hartung, T.; Hay, R.; Merten, O.-W.; Price, A.; Schechtman, L.; Stacey, G.; Stokes, W., Guidance on good cell culture practice. A report of the second ECVAM Task Force on good cell culture practice. *ARLA, Altern. Lab. Anim.* **2005,** *33*, 261-287.

11. Baker, M., 1,500 scientists lift the lid on reproducibility. *Nature (London)* **2016,** *533* (7604), 452-454.

12. Prinz, F.; Schlange, T.; Asadullah, K., Believe it or not: How much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **2011,** *10* (9), 712.

13. Baker, W.; Davis, J. G.; Leeds, W. G.; Oxley, P.; Short, W. F.; Twigg, R. S.; Watson, D. W., The synthesis and standardization of sodium resazurate for testing the hygienic quality of milk. *Biochem. J.* **1942,** *36* (1-2), i-ii.

14. Erb, R. E.; Ehlers, M. H., Resazurin reducing time as an indicator of bovine semen fertilizing capacity. *J. Dairy Sci.* **1950,** *33* (12), 853-864.

15. Pagé, B.; Pagé, M.; Noël, C., A new fluorometric assay for cytotoxicity measurements *in vitro*. *Int. J. Oncol.* **1993,** *3* (3), 473-476.

16. Nakayama, G. R.; Caton, M. C.; Nova, M. P.; Parandoosh, Z., Assessment of the Alamar Blue assay for cellular growth and viability in vitro. *J. Immunol. Methods* **1997,** *204* (2), 205-208.

17. Ahmed, S. A.; Gogal, R. M.; Walsh, J. E., A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: An alternative to [3H]thymidine incorporation assay. *J. Immunol. Methods* **1994,** *170* (2), 211-224.

18. Nociari, M. M.; Shalev, A.; Benias, P.; Russo, C., A novel one-step, highly sensitive fluorometric assay to evaluate cell-mediated cytotoxicity. *J. Immunol. Methods* **1998,** *213* (2), 157-167.

19. Zhi-Jun, Y.; Sriranganathan, N.; Vaught, T.; Arastu, S. K.; Ansar Ahmed, S., A dye-based lymphocyte proliferation assay that permits multiple immunological analyses: mRNA, cytogenetic, apoptosis, and immunophenotyping studies. *J. Immunol. Methods* **1997,** *210* (1), 25-39.

20. Evans, S. M.; Casartelli, A.; Herreros, E.; Minnick, D. T.; Day, C.; George, E.; Westmoreland, C., Development of a high throughput in vitro toxicity screen predictive of high acute in vivo toxic potential. *Toxicol. In Vitro* **2001,** *15* (4–5), 579-584.

21. Al-Nasiry, S.; Geusens, N.; Hanssens, M.; Luyten, C.; Pijnenborg, R., The use of Alamar Blue assay for quantitative analysis of viability, migration and invasion of choriocarcinoma cells. *Hum. Reprod.* **2007,** *22* (5), 1304-1309.

22. Tung, Y. C.; Hsiao, A. Y.; Allen, S. G.; Torisawa, Y.-s.; Ho, M.; Takayama, S., High-throughput 3D spheroid culture and drug testing using a 384 hanging drop array. *Analyst* **2011,** *136* (3), 473-478.

23. Bonnier, F.; Keating, M. E.; Wróbel, T. P.; Majzner, K.; Baranska, M.; Garcia-Munoz, A.; Blanco, A.; Byrne, H. J., Cell viability assessment using the Alamar blue assay: A comparison of 2D and 3D cell culture models. *Toxicol. In Vitro* **2015,** *29* (1), 124-131.

24. Ampuja, M.; Jokimäki, R.; Juuti-Uusitalo, K.; Rodriguez-Martinez, A.; Alarmo, E. L.; Kallioniemi, A., BMP4 inhibits the proliferation of breast cancer cells and induces an MMP-dependent migratory phenotype in MDA-MB-231 cells in 3D environment. *BMC Cancer* **2013,** *13* (1), 429.

25. Fields, R. D.; Lancaster, M. V., Dual-attribute continuous monitoring of cell proliferation/cytotoxicity. *Am. Biotechnol. Lab.* **1993,** *11* (4), 48-50.

26. Zachari, M. A.; Chondrou, P. S.; Pouliliou, S. E.; Mitrakas, A. G.; Abatzoglou, I.; Zois, C. E.; Koukourakis, M. I., Evaluation of the AlamarBlue assay for adherent cell irradiation experiments. *Dose-Response* **2014,** *12* (2), 246-258.

27. Mitrakas, A. G., Effect of mitochondrial metabolism-interfering agents on cancer cell mitochondrial function and radio/chemosensitivity. *Anti-cancer drugs* **2014,** *25* (10), 1182-1191.

28. Miller, J. N.; Miller, J. C., Calibration methods in instrumental analysis: Regression and correlation. In *Statistics and Chemometrics for Analytical Chemistry*, 6th ed.; Pearson Education: Harlow, England, 2010; pp 125-128.

29. Zalata, A. A.; Lammertijn, N.; Christophe, A.; Comhaire, F. H., The correlates and alleged biochemical background of the resazurin reduction test in semen. *Int. J. Androl.* **1998,** *21* (5), 289-294.

30. Gonzalez, R. J.; Tarloff, J. B., Evaluation of hepatic subcellular fractions for Alamar blue and MTT reductase activity. *Toxicol. In Vitro* **2001,** *15* (3), 257-259.

31. Abe, T.; Takahashi, S.; Fukuuchi, Y., Reduction of Alamar Blue, a novel redox indicator, is dependent on both the glycolytic and oxidative metabolism of glucose in rat cultured neurons. *Neurosci. Lett.* **2002,** *326* (3), 179-182.

32. Barnes, S.; Spenney, J. G., Stoichiometry of the NADH-oxidoreductase reaction for dehydrogenase determinations. *Clin. Chim. Acta* **1980,** *107* (3), 149-154.

33. Matsumoto, K.; Yamada, Y.; Takahashi, M.; Todoroki, T.; Mizoguchi, K.; Misaki, H.; Yuki, H., Fluorometric determination of carnitine in serum with immobilized carnitine dehydrogenase and diaphorase. *Clin. Chem.* **1990,** *36* (12), 2072-2076.

34. McKim, J. M., Jr., Building a tiered approach to *in vitro* predictive toxicity screening: A focus on assays with *in vivo* relevance. *Comb. Chem. High Throughput Screening* **2010,** *13* (2), 188-206.

35. Abu-Amero, K. K.; Bosley, T. M., Detection of mitochondrial respiratory dysfunction in circulating lymphocytes using resazurin. *Arch. Pathol. Lab. Med.* **2005,** *129* (10), 1295-1298.

36. American Type Culture Collection, *ATCC® animal cell culture guide*. American Type Culture Collection: Manasas, VA, 2014; p 4.

37. Chance, B.; Hollunger, G., Inhibition of electron and energy transfer in mitochondria: I. Effects of amytal, thiophental, rotenone, progesterone, and methylene glycol. *J. Biol. Chem.* **1963,** *238* (1), 418-431.

38. Barrientos, A.; Moraes, C. T., Titrating the effects of mitochondrial Complex I impairment in the cell physiology. *J. Biol. Chem.* **1999,** *274* (23), 16188-16197.

39. Yoshikawa, S.; Caughey, W. S., Infrared evidence of cyanide binding to iron and copper sites in bovine heart cytochrome *c* oxidase. Implications regarding oxygen reduction. *J. Biol. Chem.* **1990,** *265* (14), 7945-7958.

40. Hargreaves, I. P.; Duncan, A. J.; Wu, L.; Agrawal, A.; Land, J. M.; Heales, S. J. R., Inhibition of mitochondrial complex IV leads to secondary loss complex II–III activity: Implications for the

pathogenesis and treatment of mitochondrial encephalomyopathies. *Mitochondrion* **2007,** *7* (4), 284-287.

41. Van Buuren, K. J. H.; Zuurendonk, P. F.; Van Gelder, B. F.; Muijsers, A. O., Biochemical and biophysical studies on cytochrome aa_3. V. Binding of cyanide to cytochrome aa_3. *Biochim. Biophys. Acta, Bioenerg.* **1972,** *256* (2), 243-257.

42. Hawtin, S. R.; Dobbins, A. C.; Tailor, V. J.; Shearman, M. S., β-amyloid inhibition of MTT reduction is not mimicked by inhibitors of mitochondrial respiration. *Biochem. Soc. Trans.* **1995,** *23* (1), 56S-56S.

43. Torsvik, A.; Stieber, D.; Enger, P. Ø.; Golebiewska, A.; Molven, A.; Svendsen, A.; Westermark, B.; Niclou, S. P.; Olsen, T. K.; Chekenya Enger, M.; Bjerkvig, R., U-251 revisited: Genetic drift and phenotypic consequences of long-term cultures of glioblastoma cells. *Cancer Med.* **2014,** *3* (4), 812-824.

44. Geraghty, R. J.; Capes-Davis, A.; Davis, J. M.; Downward, J.; Freshney, R. I.; Knezevic, I.; Lovell-Badge, R.; Masters, J. R. W.; Meredith, J.; Stacey, G. N.; Thraves, P.; Vias, M., Guidelines for the use of cell lines in biomedical research. *Br. J. Cancer* **2014,** *111* (6), 1021-1046.

45. Uhlig, S.; Jestoi, M.; Knutsen, A. K.; Heier, B. T., Multiple regression analysis as a tool for the identification of relations between semi-quantitative LC-MS data and cytotoxicity of extracts of the fungus *Fusarium avenaceum* (syn. *F. arthrosporioides*). *Toxicon* **2006,** *48* (5), 567-579.

46. de Fries, R.; Mitsuhashi, M., Quantification of mitogen induced human lymphocyte proliferation: Comparison of alamarBlue™ assay to ^{3}H-thymidine incorporation assay. *J. Clin. Lab. Anal.* **1995,** *9* (2), 89-95.

47. Liu, W.; Xu, J.; Wu, S.; Liu, Y.; Yu, X.; Chen, J.; Tang, X.; Wang, Z.; Zhu, X.; Li, X., Selective anti-proliferation of HER2-positive breast cancer cells by anthocyanins identified by high-throughput screening. *PLoS ONE* **2013,** *8* (12), e81586.

48. Uzunoglu, S.; Karaca, B.; Atmaca, H.; Kisim, A.; Sezgin, C.; Karabulut, B.; Uslu, R., Comparison of XTT and Alamar blue assays in the assessment of the viability of various human cancer cell lines by AT-101 (−/− gossypol). *Toxicol. Mech. Methods* **2010,** *20* (8), 482-486.

49. Gong, X.; Fang, Q.; Li, X.; Han, X.; Wu, Y.; Yang, S.; Shen, B. Q., A high-throughput end-point assay for viable mammalian cell estimation. *Cytotechnology* **2005,** *49* (1), 51-58.

50. Quent, V. M. C.; Loessner, D.; Friis, T.; Reichert, J. C.; Hutmacher, D. W., Discrepancies between metabolic activity and DNA content as tool to assess cell proliferation in cancer research. *J. Cell Mol. Med.* **2010,** *14* (4), 1003-1013.

51. Squatrito, R. C.; Connor, J. P.; Buller, R. E., Comparison of a novel redox dye cell growth assay to the ATP bioluminescence assay. *Gynecol. Oncol.* **1995,** *58* (1), 101-105.

52. Pace, R. T.; Burg, K. J. L., Toxic effects of resazurin on cell cultures. *Cytotechnology* **2015,** *67* (1), 13-17.

53. Solit, D. B.; Garraway, L. A.; Pratilas, C. A.; Sawai, A.; Getz, G.; Basso, A.; Ye, Q.; Lobo, J. M.; She, Y.; Osman, I.; Golub, T. R.; Sebolt-Leopold, J.; Sellers, W. R.; Rosen, N., BRAF mutation predicts sensitivity to MEK inhibition. *Nature (London)* **2006,** *439* (7074), 358-362.

54. Rodríguez-Corrales, J. Á.; Josan, J. S., Resazurin live cell assay: Setup and fine-tuning for reliable cytotoxicity results. In *Proteomics for Drug Discovery. Methods in Molecular Biology*, Lazar, I. M.; Kontoyianni, M.; Lazar, A. C., Eds. Humana Press: New York, NY, 2017; Vol. 1647, pp 207-219.

55. Teodoro, J. S.; Rolo, A. P.; Palmeira, C. M., The NAD ratio redox paradox: Why does too much reductive power cause oxidative stress? *Toxicol. Mech. Methods* **2013,** *23* (5), 297-302.

56. Larson, E. M.; Doughman, D. J.; Gregerson, D. S.; Obritsch, W. F., A new, simple, nonradioactive, nontoxic in vitro assay to monitor corneal endothelial cell viability. *Invest. Ophthalmol. Vis. Sci.* **1997,** *38* (10), 1929-1933.

57. Springer, J. E.; Azbill, R. D.; Carlson, S. L., A rapid and sensitive assay for measuring mitochondrial metabolic activity in isolated neural tissue. *Brain Res. Protoc.* **1998,** *2* (4), 259-263.

58. Herlyn, M.; Balaban, G.; Bennicelli, J.; Guerry, I. V. D.; Halaban, R.; Herlyn, D.; Elder, D. E.; Maul, G. G.; Steplewski, Z.; Nowell, P. C.; Clark, W. H.; Koprowski, H., Primary melanoma cells of the vertical growth phase: Similarities to metastatic cells. *JNCI, J. Natl. Cancer Inst.* **1985,** *74* (2), 283-289.

5. RESAZURIN LIVE CELL ASSAY: SETUP AND FINE-TUNING FOR RELIABLE CYTOTOXICITY RESULTS

José Á. Rodríguez Corrales, Jatinder S. Josan

This chapter has been adapted from a published manuscript: Resazurin Live Cell Assay: Setup and Fine-Tuning for Reliable Cytotoxicity Results. In *Proteomics for Drug Discovery. Methods in Molecular Biology*, Lazar, I. M.; Kontoyianni, M.; Lazar, A. C., Eds. Humana Press: New York, NY, 2017; Vol. 1647, pp 207-219. Reprinted and adapted by permission from Springer Nature, Copyright (2017).

5.1. Abstract

In vitro cytotoxicity tests allow for fast and inexpensive screening of drug efficacy prior to *in vivo* studies. The resazurin assay (commercialized as Alamar Blue®) has been extensively utilized for this purpose in 2D and 3D cell cultures, and high-throughput screening. However, improper or lack of assay validation can generate unreliable results and limit reproducibility. Herein, we report a detailed protocol for the optimization of the resazurin assay to determine relevant analytical (limits of detection, quantification and linear range) and biological (growth kinetics) parameters, and, thus, provide accurate cytotoxicity results. Fine-tuning of the resazurin assay will allow accurate and fast quantification of cytotoxicity for drug discovery. Unlike more complicated methods (e.g. mass spectrometry), this assay utilizes fluorescence spectroscopy and, thus, provides a less costly alternative to observe changes in the reductase proteome of the cells.

5.2. Introduction

Drug development is a costly (>1 billion dollars per drug) and lengthy (several years) process that utilizes several "filters" to select a lead molecule that can be carried to clinical trials and, finally, be commercialized. Drug screening is performed early in this process to determine cytotoxicity of candidates and identify leads, which are later optimized. Tiered approaches to drug evaluation involving multiple parameters (e.g., cytotoxicity, changes in the proteome) have been proposed to improve our ability to predict the *in vivo* activity of drugs.[2] The resazurin assay, commercially available as Alamar Blue®, has been extensively used to determine toxicity in human and animal *in vitro* models since the early 1990s.[3-4] Resazurin reduction by metabolically active, or "live", cells produces the highly fluorescent resorufin, thus allowing spectroscopic evaluation of cell viability under certain conditions. This reduction process is promoted by NADH reductase and carnitine dehydrogenase *in vitro*.[5-6] However, it has been shown that resazurin reduction is not carried out by only one specific enzyme, as incubation with different subcellular fractions (cytosolic, microsomal, and mitochondrial) led in all cases to the formation of resorufin.[7] This observation indicates that the results of this assay correlate to the reductase activity of the proteome as a whole, and not to any one enzyme in particular.

In this context, resazurin can be used to screen for drugs that target reductases directly or indirectly (e.g., enzymes involved in the same signaling pathway), or promote cell death or growth. Furthermore, reductase activity varies among different cell lines, which influences the assay conditions. Improper or lack of assay validation conditions can generate, therefore, unreliable results and limit reproducibility. Several scientists and organizations have raised concerns due to the irreproducibility of published results that could arise from improper assay validation, among other factors.[8-9]

Herein, we present a method for the optimization of the resazurin assay for cell viability and cytotoxicity measurements, where both analytical and biological parameters are studied and validated. Although the amount of information gathered with this technique is rather limited (resazurin reduction is affected by changes in the reductase proteome and not only one enzyme), it provides a facile, fast and accurate assessment of drug cytotoxicity in drug discovery, as long as the assay is properly validated. Thus, this inexpensive assay allows identification of hit molecules in the earlier stages of a drug discovery screening. These hits can be further studied through more selective yet cost-intensive proteomics techniques, including SDS-PAGE and MALDI-TOF, to determine their specific biological targets.

5.3. Materials

Prepare all solutions in phosphate buffered saline (PBS) or complete medium, i.e., medium supplemented with fetal bovine serum (FBS) and antibiotics, as specified. Use chemical reagents of reagent grade or higher. Sterilize all solutions by filtering through a sterile 0.22 μm membrane or an equivalent method, or, procure sterile solutions from dedicated manufacturers. Prepare and store all solutions at 4 °C, except if noted otherwise, and dispose following authorized procedures. Use only sterile equipment and supplies (pipettes, pipette tips, culture flasks, vials, tubes, etc.) for handling the cells. Cell work must be performed under aseptic conditions, in a biosafety cabinet, following all appropriate safety regulations.

5.3.1. Reagent solutions

1. Mammalian cells.

2. Complete cell culture medium: Follow the instructions provided by the supplier of the cells to prepare a cell culture medium of adequate composition. For example, supplement the base medium (e.g., EMEM or equivalent) with 10% FBS and penicillin/streptomycin (1X). Sterile filter and store in the refrigerator at 4 ºC.
3. PBS solution (1X): 0.01 M phosphate buffer, 0.0027 M KCl and 0.137 M NaCl, pH 7.4 at 25 °C. Dissolve a PBS tablet (e.g., Sigma-Aldrich P4417-100g) in 200 mL (or volume suggested by manufacturer) of autoclaved deionized water (see **Note 1**). Mix and filter under sterile conditions.
4. Trypsin solution: 0.25% Trypsin with 0.53 mM EDTA in PBS (1X). Use any high quality commercial product.
5. Penicillin/streptomycin solution (100X): 10,000 units/mL penicillin and 10 mg/mL streptomycin. Use any high quality commercial product.
6. Resazurin stock solution: 10 mg/mL. Add 1 g of resazurin into 100 mL of PBS and mix until the reagent has dissolved completely. Sterile filter and store in the refrigerator protected from light at 4 ºC (see **Note 2**).
7. Resazurin working solution: 15 µM. Add 180 µL of cold resazurin stock solution into 50 mL of complete medium at 37 °C. Prepare fresh immediately before use (see **Note 3**).

5.3.2. Equipment

1. Cell culture supplies: sterile cell culture flasks, pipettes, vials, tubes, etc.
2. Conventional cell culture lab equipment: biosafetycabinet, incubator, microscope, water bath, 4 ºC refrigerator, -20 ºC and -80 ºC freezers, centrifuges, and hemocytometer or flow cytometer.

5.4. Methods

5.4.1 Validation of assay conditions

1. Grow cells in a T75 flask under the conditions recommended by the supplier. Typical conditions may include EMEM:FBS 9:1 with 1% penicillin/streptomycin, with incubation performed at 37 °C in a 5% CO_2 atmosphere (see **Note 4**).
2. Harvest the cells when confluency reaches ~70-90%. Typical detachment conditions include a PBS wash, followed by incubation with 0.25% trypsin/EDTA for 5-10 min at 37 °C. After trypsinization, the trypsin must be inactivated by the addition of complete medium. Next, the cells must be dispersed by gently aspirating and dispensing the cell suspension with a pipette.
3. Transfer the cell suspension to a 15 mL Falcon tube and centrifuge according to supplier's recommendation, e.g., 5 min at 125 × g (see **Note 5**).
4. Gently aspirate the supernatant without disturbing the cell pellet. Resuspend the pellet in complete medium by careful, repeated pipetting.
5. Determine the cell concentration using a hemocytometer, a flow cytometer, or any other available method (see **Note 6**).
6. Prepare a stock solution of 2,000,000 cells/mL (stock A). If the concentration is higher than 2,000,000 cells/mL, dilute with complete medium. If the concentration is lower, centrifuge, remove the supernatant and resuspend in an appropriate volume.
7. Prepare stocks of 250,000 cells/mL (stock B) and 25,000 cells/mL (stock C) by dilution in cell medium.
8. Seed two 96-well plates with the volumes provided in Table 5.1, i.e., seed three wells per cell plating density, including blanks, and per incubation time. If five incubation times are

tested, the total number of wells per plating density is 3 × 5 = 15, and the total number of wells needed is 3 × 12 × 5 = 180 wells (see **Notes 7** through **9**, and Fig. 5.1).

Table 5.1. Volume of stock solutions to be placed per well.

Columns	Plating density (cells/well)	Stock A (µL)	Stock B (µL)	Stock C (µL)	Complete medium (µL)
1	300,000	150	-	-	-
2	200,000	100	-	-	50
3	150,000	75	-	-	75
4	100,000	50	-	-	100
5	50,000	25	-	-	125
6	25,000	-	100	-	50
7	10,000	-	40	-	110
8	5,000	-	20	-	130
9	1,000	-	-	40	110
10	500	-	-	20	130
11	100	-	-	4	146
12 (blanks)	0	-	-	-	150

Note: row H of one of the plates should be used for blanks and, thus, filled with medium only.

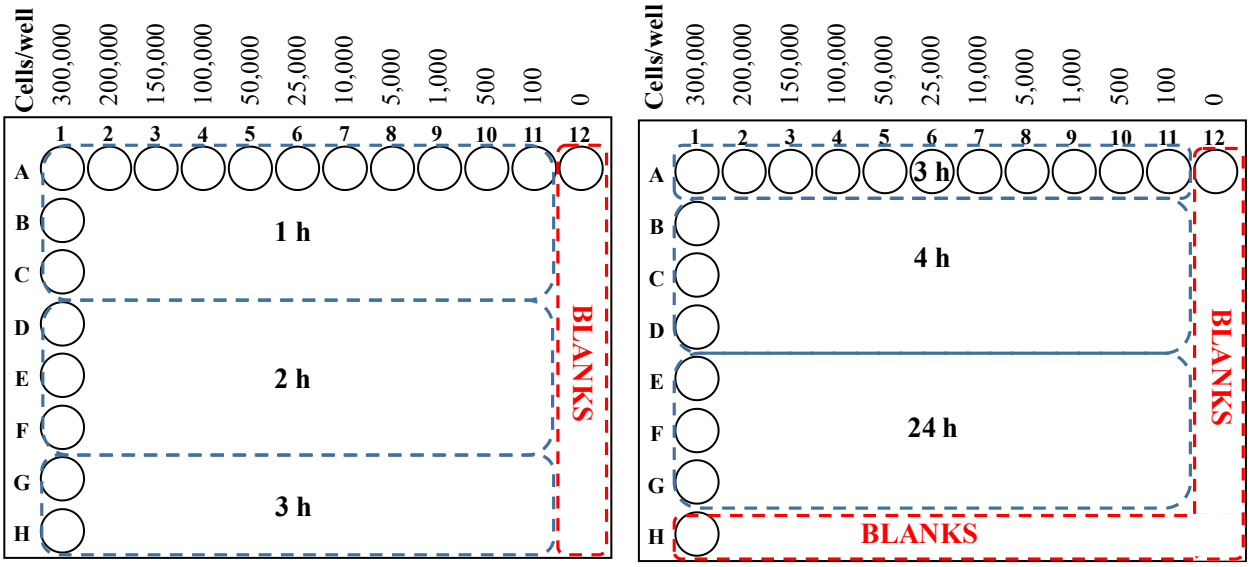


Figure 5.1. Sample plating setup. The recommended setup allows triplicate testing of 12 plating densities at 5 resazurin incubation times per triplicate (15 rows/density). **Notes 6** and **7** provide important considerations if the setup is to be modified.

9. Place 96-well plates ("culture plates") in the incubator for 24 h to allow for cell attachment (see **Note 10**).
10. Aspirate the medium from all wells. Add 250 µL of resazurin working solution per well using a multichannel pipette. Return the plate to the incubator (see **Note 11**).
11. After 1 h, remove 100 µL from the wells located in wells A1 through C11 and pipette into a separate 96-well plate with black walls and clear bottom ("reader plate"). Also, pipette 100 µL from eight blank wells (column 12 or row-H of one of the plates, see Fig. 5.1). Return the culture plate to incubator (see **Notes 12** and **13**).
12. Place the black reader plate in a fluorescence plate reader and measure the fluorescence intensity using excitation at 560 nm and detection at 590 nm (see **Note 14**).

13. Repeat steps 11-12 using a different triplicate of wells after 2 h, 3 h, 4 h, and 24 h incubation with resazurin.
14. For each time point, determine the average fluorescence of the corresponding blanks ($\bar{f}_b$) and the standard deviation (σ_b).
15. Subtract $\bar{f}_b$ from all samples to obtain the blank-corrected fluorescence intensity. Exclude any blank-corrected values lower than $10\times\sigma_b$ from further analysis. These values are considered to be under the limit of quantification.
16. Plot blank-corrected fluorescence intensity *vs*. number of plated cells (Fig. 5.2). Identify the linear portion and plateau in the graph (see **Note 15**) and determine the limit of linearity (LOL) using an appropriate method (see **Note 16**).

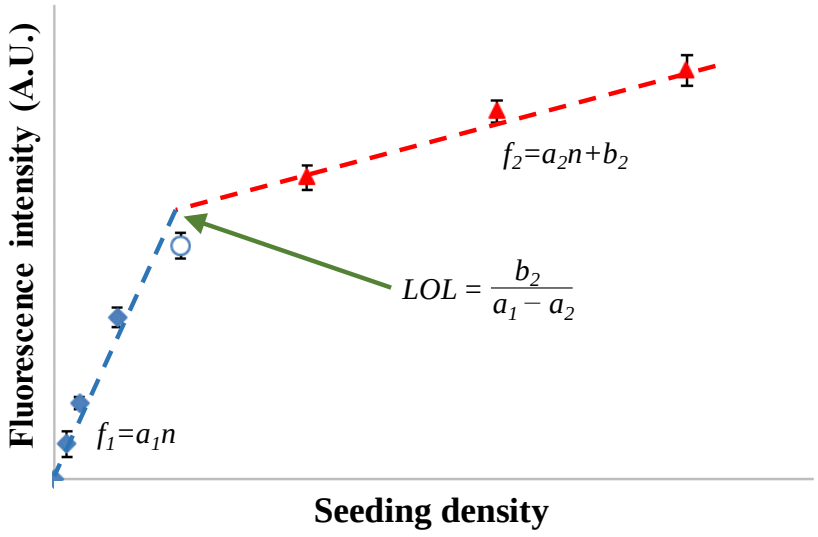


Figure 5.2. Sample fluorescence *vs.* plating density graph, displaying the linear (blue diamonds) and plateau (red triangles) regions. Best-fit linear equations are provided in the graph, along with the LOL.

17. Calculate limit of detection (LOD) and quantification (LOQ), using equations 5.1 and 5.2, where a_1 stands for the slope of the best-fit equation for the linear range.

$$LOD = 3\times\sigma_b/a_1 \quad \text{(Eq. 5.1)}$$

$$LOQ = 10\times\sigma_b/a_1 \quad \text{(Eq. 5.2)}$$

18. Under most circumstances, choose the "optimum" incubation time with resazurin ($t_{resazurin}$) that produces the widest linear range (LOL/LOQ) (*see* **Note 17**).

5.4.2 Evaluation of growth kinetics

1. Perform steps 1-5 from section 5.4.1**.**
2. Seed 1,000-5,000 cells/well in 12 wells (i.e., 3 wells per 4 time-points) of a 96-well plate, that will be used to assess growth kinetics (see **Note 18**).
3. Place 96-well plate in the incubator for 24 h to allow cell attachment (see **Note 10**).
4. Carefully aspirate medium from all wells.
5. Rinse three wells with 100 µL of PBS (1X), and aspirate.
6. Add 100 µL of trypsin 0.25% with 0.53 mM EDTA to the wells prepared in step 5 (see **Note 19**). Pipette 100 µL of complete medium into all other wells.
7. Return the 96-well plate to the incubator for 10 min.
8. Gently mix the wells with trypsin with a pipette to improve cell detachment from the plate.
9. Check cells under the microscope to confirm detachment from the plate. If incomplete, repeat step 8.
10. Determine the number of cells/well in the three wells treated with trypsin, using a hemocytometer, a flow cytometer, or any other available method (see **Note 6**).
11. Return the 96-well plate to the incubator.

12. Repeat steps 4-11, 2, 3 and 4 days after seeding, each day trypsinizing the following three wells.
13. Determine the average number of cells/well for each time point.
14. Normalize all values by using the number of cells/well in day 1 (24 h after seeding) (see **Note 20**).
15. Plot natural logarithm of the normalized number of cells/well *vs.* time and determine the equation for the line of best fit.
16. Determine the population doubling time (τ) using Eq. 5.3, where a_k is the slope of the line of best fit as determined in step 15.[10]

$$\tau = \ln(2)/a_k \qquad \text{(Eq. 5.3)}$$

5.4.3 Cytotoxicity experiment design

In order to illustrate the reasoning behind the assay validation and the experimental design, the general procedure is summarized as a flow chart in Fig. 5.3.

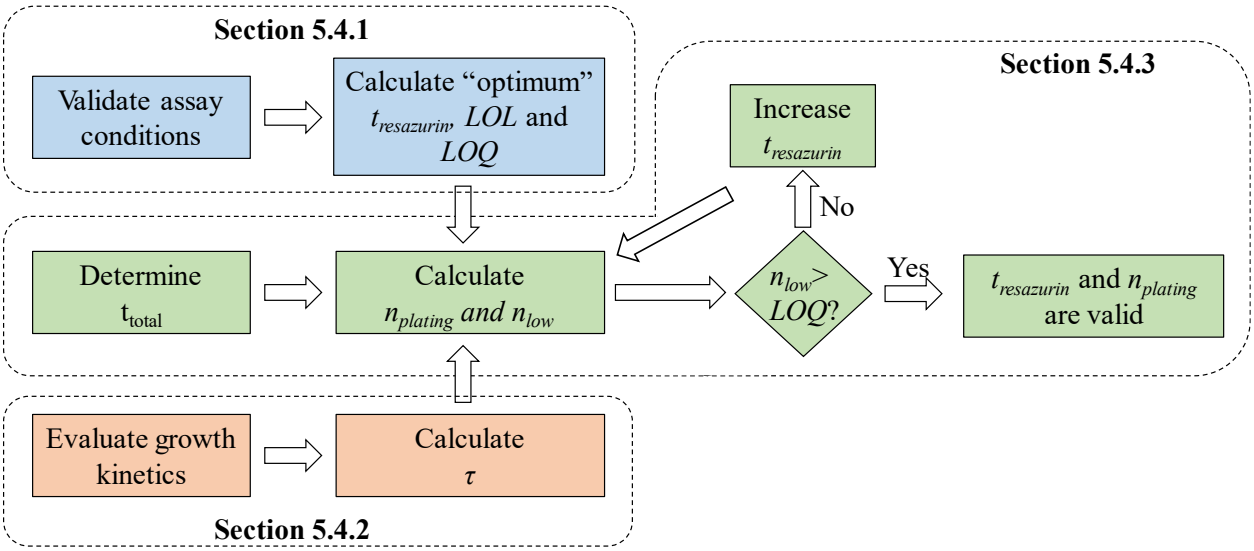


Figure 5.3. Flow chart for effective cytotoxicity experiment design.

1. Based on the goals of the experiment, calculate the total number of days that cells will be grown (t_{total}). Take into account the time required for attachment to the plate (t_{attach}), time of incubation with cytotoxic compound (t_{test}) and any other relevant incubations (e.g., toxicity evaluation 24 h after removing cytotoxic compound). Do not add time of incubation of resazurin.
2. Based on the validated resazurin assay conditions, namely incubation time ($t_{resazurin}$) at a given concentration of resazurin (see Section 5.4.1, step 18), determine the seeding density ($n_{plating}$) required to ensure that the number of cells (n_{final}) at t_{total} remains in the linear range. n_{final} can be estimated as 90% of the LOL, and $n_{plating}$ can be solved using Eq. 5.4 (see **Notes 21** and **22**).

$$n_{plating} = \frac{0.9 \times LOL}{2^{(t_{total}/\tau)}} \qquad \text{(Eq. 5.4)}$$

3. Assuming ~95% cytotoxicity in the sample treated with the highest concentration of drug, calculate the estimated "lowest" number of live cells after treatment with cytotoxic drug (n_{low}).

$$n_{low}=0.045\times LOL \quad \text{(Eq. 5.5)}$$

4. If n_{low} > LOQ at $t_{resazurin}$, all parameters have been correctly calculated and the resazurin assay conditions should be valid.
5. If n_{low} < LOQ, $t_{resazurin}$ should be extended to increase the sensitivity of the assay and steps 1-4 of this section should be repeated.

5.4.4. Cytotoxicity experiment

1. Perform steps 1-5 from section 5.4.1.
2. Pipette enough solution to seed $n_{plating}$ cells per well, as determined in step 2 from section 5.4.3. Seed at least two or three wells for each concentration of drug to be tested, and for the controls (growth medium without drug, and growth medium with any solvents used).
3. Incubate cells at 37 °C and 5% CO_2 for 24 h (see **Note 10**).
4. Prepare stock solutions of drug to be tested in complete medium. If dimethyl sulfoxide (DMSO) or other cytotoxic solvents are used to dissolve the drug, controls should be prepared using the same concentration of solvent.
5. Replace medium from the wells with solutions of drug.
6. Return plate(s) to the incubator for desired period of time.
7. Aspirate drug solution and add 250 μL of resazurin working solution to all wells using a multichannel pipettor. Prepare blanks by pipetting 250 μL of resazurin working solution into at least three empty wells.

8. Incubate with resazurin for the time determined in section 5.4.3.
9. Pipette 100 μL out of each well and into a 96-well plate with black walls and clear bottom.
10. Place the black 96-well plate in a plate reader and measure fluorescence intensity using excitation at 560 nm with detection at 590 nm (see **Note 14**).
11. Calculate the average fluorescence of the blanks ($\bar{f}_b$).
12. Subtract $\bar{f}_b$ from all values to obtain the blank-corrected fluorescence intensity of samples ($\bar{f}_s$) and untreated controls ($\bar{f}_c$) (see **Note 23**).
13. Calculate % viability using Eq. 5.6.

$$\% \text{ viability} = \frac{\bar{\bar{f}}_s}{\bar{\bar{f}}_c} \times 100 \qquad \text{(Eq. 5.6)}$$

14. Determine IC_{50} using an appropriate statistical program. A sample % viability *vs*. drug concentration plot is shown in Fig. 5.4.

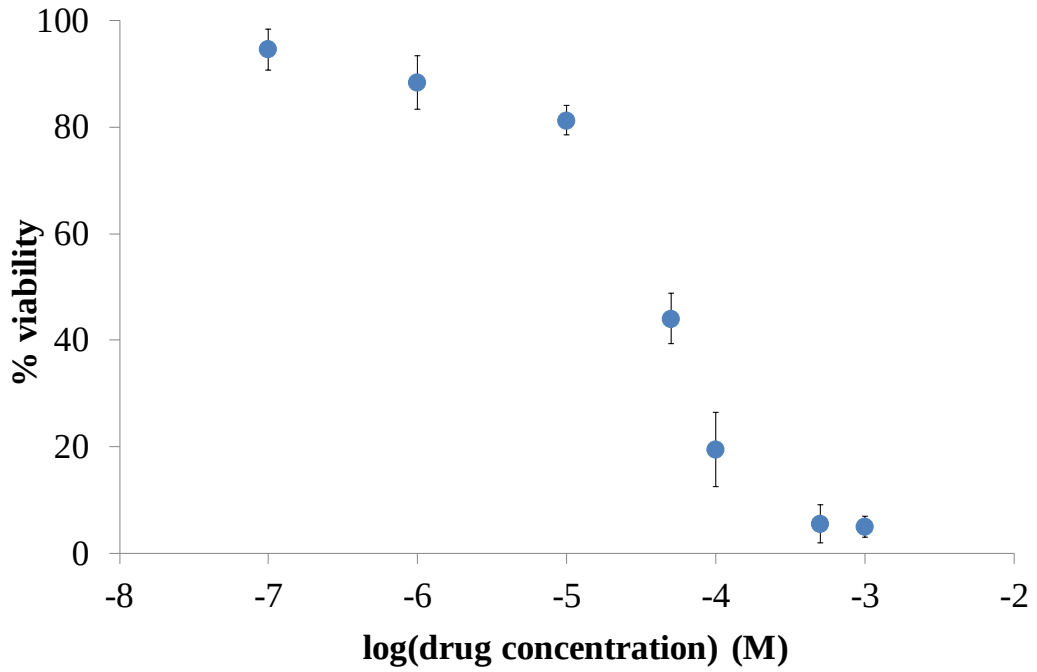


Figure 5.4. Sample cytotoxicity graph showing cell viability *vs.* log of drug concentration. Error bars correspond to standard deviation of samples at each concentration of drug.

15. In order to report average IC_{50}, at least two or three different cytotoxicity experiments should be performed. Uncertainty should correspond to the standard deviation of the three experiments.

5.5. Notes

1. Non-autoclaved deionized water can also be used since the solution will be sterile filtered. However, we recommend use of autoclaved water as an extra cautionary measure.
2. Resazurin stock solution can be stored in 50 mL centrifuge tubes wrapped in aluminum foil to prevent photodegradation of resazurin. Solution stability can be checked periodically by diluting the stock and calculating the concentration based on wavelength of maximum absorption (usually 600 nm).

3. Resazurin stock solution should be kept cold at all times to avoid its thermal degradation. Thus, we recommend warming up the complete medium and then adding the cold resazurin stock solution. Given the small volume used of the latter, the resulting solution will remain warm. For more accurate results, working resazurin solutions should be used immediately to avoid photo- or thermal degradation of resazurin. Never leave working resazurin solutions in a warm water bath, or its equivalent.
4. The growth conditions presented in section 5.4.1 are representative of procedures recommended by ATCC for several cancer cells lines. Any growth protocols recommended by the supplier of a particular cell line supersede those presented here.
5. Centrifugation allows further removal of trypsin from the cell suspension. Although trypsin should be inactivated by the components of the complete medium, we have observed that cell morphology can be slightly affected if not removed.
6. If possible, perform a viability assay in the cell suspension, such as trypan blue or propidium iodide exclusion for bright field or fluorescence microscopy, respectively. Non-viable cells might skew the results of the experiment.
7. Values presented in Table 5.1 are based on our experience with several cancer cell lines in 96 well plates. The higher plating densities should be used to ensure that a plateau in resazurin reduction is reached, while the lower densities are especially necessary for long-term (e.g., 24 h) incubations. Plating densities should be changed depending on the specific cell line to be used and the size of the well plate (e.g., 6-, 24- or 384- well plates).
8. The recommended plating allows for triplicate fluorescence measurements of each plating density at each of five different incubation times. Plating can be modified to include less

samples per time point (e.g., duplicate instead of triplicate) or less times points (e.g., two *vs.* five).

9. If the recommended plating setup is used, only 15 wells per seeding density (5 time points × 3) are needed. Thus, one row of the two 96-well plates would remain unused. We recommend using this row for blanks (see Fig. 5.1) to perform a more accurate evaluation of the limits of detection and quantification.
10. Some cell lines might require longer time to attach to the plate. If visualization of cells under the microscope reveals a significant proportion of floaters, plates should be returned to the incubator and allowed extra time. Once an appropriate time has been determined, it should be used in ALL experiments, including evaluation of cytotoxicity and kinetics of growth.
11. Alternatively, a lower volume of resazurin working solution can be used at the expense of a shorter linear range for the method, since resazurin will be metabolized faster by the cells. Furthermore, a more concentrated resazurin working solution could be added directly to the cells without need of aspirating the old medium. Although this might reduce the time of analysis, results might be skewed by differences in the volume of medium throughout the wells due to evaporation.
12. Alternatively, fluorescence can be measured directly in the plates containing cells. In our experience, this leads to lower fluorescence intensities perhaps attributed to self-absorption. Moreover, pipetting resazurin solution into black 96-well plates instead of growing cells directly on them might prove more cost effective since “reader plates” can be reused after proper disinfection. Such plates are significantly more expensive than pipette tips or regular 96-well plates (“culture plates”). Periodic inspection for cracks or scratches in the black plates should be performed to ensure acceptable spectroscopic quality.

13. The same blank wells can be used at multiple incubation times (e.g., 1 and 2 h) since no cells are present to reduce resazurin.
14. Excitation and emission wavelengths should be optimized for the instrument. Resazurin working solution should be used as a blank, while a fully reduced sample can be prepared by adding sodium borohydride to resazurin working solution.[4] Alternatively, one of the samples with high plating densities can be used for this purpose. Measure fluorescence at different excitation and emission wavelengths ranging from 540 to 580 nm and 570 to 600 nm, respectively. Wavelengths should be selected as those that yield the highest sample/blank emission ratio. Instrument gain should be adjusted so that no samples surpass the detector's threshold, while keeping the highest possible sample/blank emission ratio.
15. Determination of the linear and plateau portions of the graph is critical for latter steps. It is recommended to start with the linear portion by performing graphical examination based on the lowest plating densities. In some cases, the plateau might be accurately described using a linear equation with a slope of or close to zero.
16. Assuming linear fits are applied to the linear and plateau regions, best-fit equations would be expressed as $f_1=a_1n$ and $f_2=a_2n+b_2$, respectively, where f stands for blank-corrected fluorescence, n is the number of cells or plating density, a is the slope, and b is the intercept on the y-axis. A sample graph displaying both regions and equations is presented in Fig. 5.2. Thus, the limit of linearity (expressed in number of cells) would be calculated as:

$$LOL=\frac{b_2}{a_1-a_2} \quad \text{(Eq. 5.7)}$$

17. Although the resazurin assay is moderately non-toxic for most cell lines, inherent cytotoxicity should be tested. In our experience, short exposure (2-3 h) has minimal toxicity, whereas long contact (days) leads to decreased viability.
18. The exact number of cells to be plated depends on their growth kinetics. Seeding density should be low enough to avoid reaching 100% confluency at any time point.
19. Detachment conditions should be validated for each cell line. In our experience, treatment with 0.25% trypsin with 0.53 mM EDTA for 10 min works well for most cell lines. For highly adherent cell lines (e.g., HaCaT), PBS can be substituted with 0.05% EDTA for rinses, and EDTA concentration in trypsin solution can be increased to 0.025%.
20. Although the number of cells seeded in the well is known, it is best to actually measure it 24 h after seeding. This allows correction for incomplete attachment (e.g., cells might die after counting and viability evaluation but before attaching to plate) and incomplete detachment from well plate upon trypsination.
21. Eq. 5.4 assumes 100% attachment efficiency, which might not be true for all cell lines. Although the equation could be corrected to account for this error, we recommend to use as is to ensure that cells do not exceed the linear range of the resazurin assay.
22. Eq. 5.4 was derived from a basic exponential growth function (Eq. 5.8), where the number of cells (n_i) at a given time (t_i) after plating depends on the doubling time (τ) and initial number of cells (n_0).

$$n_i = n_0 \times 2^{\left(\frac{t_i}{\tau}\right)} \qquad \text{(Eq. 5.8)}$$

Given the assumptions described in step 2 of section 5.4.2, n_i at t_{total} corresponds to 90% of the LOL, and n_0 is the plating density.

23. If a “solvent control” was prepared with DMSO or any other potential cytotoxic solvent, the drug samples with the same solvent concentration should be compared to the “solvent control” instead.

5.6. References

1. Rodríguez-Corrales, J. Á.; Josan, J. S., Resazurin live cell assay: Setup and fine-tuning for reliable cytotoxicity results. In *Proteomics for Drug Discovery. Methods in Molecular Biology*, Lazar, I. M.; Kontoyianni, M.; Lazar, A. C., Eds. Humana Press: New York, NY, 2017; Vol. 1647, pp 207-219.

2. McKim, J. M., Jr., Building a tiered approach to *in vitro* predictive toxicity screening: A focus on assays with *in vivo* relevance. *Comb. Chem. High Throughput Screening* **2010,** *13* (2), 188-206.

3. Fields, R. D.; Lancaster, M. V., Dual-attribute continuous monitoring of cell proliferation/cytotoxicity. *Am. Biotechnol. Lab.* **1993,** *11* (4), 48-50.

4. Pagé, B.; Pagé, M.; Noël, C., A new fluorometric assay for cytotoxicity measurements *in vitro*. *Int. J. Oncol.* **1993,** *3* (3), 473-476.

5. Barnes, S.; Spenney, J. G., Stoichiometry of the NADH-oxidoreductase reaction for dehydrogenase determinations. *Clin. Chim. Acta* **1980,** *107* (3), 149-154.

6. Matsumoto, K.; Yamada, Y.; Takahashi, M.; Todoroki, T.; Mizoguchi, K.; Misaki, H.; Yuki, H., Fluorometric determination of carnitine in serum with immobilized carnitine dehydrogenase and diaphorase. *Clin. Chem.* **1990,** *36* (12), 2072-2076.

7. Gonzalez, R. J.; Tarloff, J. B., Evaluation of hepatic subcellular fractions for Alamar blue and MTT reductase activity. *Toxicol. In Vitro* **2001,** *15* (3), 257-259.

8. Begley, C. G.; Ellis, L. M., Drug development: Raise standards for preclinical cancer research. *Nature (London)* **2012,** *483* (7391), 531-533.

9. Prinz, F.; Schlange, T.; Asadullah, K., Believe it or not: How much can we rely on published data on potential drug targets? *Nat. Rev. Drug Discov.* **2011,** *10* (9), 712.

10. American Type Culture Collection, *ATCC® animal cell culture guide*. American Type Culture Collection: Manasas, VA, 2014; p 4.

6. CONCLUSIONS

Even though cancer treatment has experienced significant developments in the last decades (e.g., the introduction of alkylating/platinating agents) and especially over the past several years (e.g., immunotherapy), there is a need for novel drugs that overcome the limitations of current treatments. A major ongoing concern is off-target toxicity, i.e., the harmful effects exerted by the drugs in tissues other than the tumor, which is particularly challenging to address given the similarities between the putative targets (e.g., DNA or proteins) in cancer and non-cancer tissues, including those from which the tumor originated. The studies presented in this dissertation explored two potential approaches to restricting the activity of putative drugs: photodynamic therapy and heteromultivalency. The former approach seeks to limit off-target toxicity through the development of drugs that are inactive until photoactivated at the site of the tumor, while the latter relies on the selective accumulation of heteromultivalent ligands in the cancer cells based on the expression of a combination of surface proteins not found in any other tissues. Finally, we showed the conditions under which the resazurin assay, a widely-used reagent with a variety of commercial names, provides an accurate estimation of cytotoxicity or proliferation, which is imperative in drug evaluation for cancer and other diseases.

Our studies on a potential photodynamic agent, $[(phen)_2Ru(dpp)Rh(phen)Cl_2]^{3+}$ or **Ru(II)-Rh(III)**, revealed a strong correlation between the non-covalent interactions established with DNA (i.e., a putative cellular target) and its covalent modification. **Ru(II)-Rh(III)** is a cationic bimetallic complex that binds strongly to DNA with a binding constant on the order of 10^4-10^5 $M^{-1}cm^{-1}$ depending on the ionic strength of the medium. This non-covalent interaction appears to involve a combination of electrostatic and groove binding, based on the fact that the amount of DNA-bound **Ru(II)-Rh(III)** decreased in the presence of known minor groove binder Hoechst

33342. Further, this metal complex is able to promote multifunctional covalent photomodification of DNA (i.e., oxygen-independent photocleavage and photobinding) that inhibits its amplification by the Polymerase Chain Reaction (PCR), whereas the complex remains widely inactive in samples protected from light. Even though the non-covalent interactions formed between **Ru(II)-Rh(III)** and DNA are not sufficient to inhibit PCR amplification in the absence of photoactivation, they play a major role as an association step prior to photolysis. We showed that samples photolyzed at high ionic strength, which decreased non-covalently bound fraction of **Ru(II)-Rh(III)**, displayed a significantly lower extent of covalent photomodification and, in turn, abolished its activity in the PCR amplification assay. We propose that the design of future polymetallic systems that target DNA should consider the contribution of non-covalent interactions that may substantially enhance their activity, especially under physiologically-relevant conditions (e.g., ionic strength and pH).

In the second approach, focused on developing heteromultivalent compounds, our analysis of protein and gene expression in patient samples and cell lines allowed the identification of several protein combinations selectively expressed in melanoma. Starting with the definition of the human protein surfaceome using five available databases, including experimental and predictive assignments of subcellular localization, we analyzed gene expression profiles of 474 melanoma patients and 31 normal tissue types using a threshold-based screening approach with Python scripts written by the authors. Upon further inspection into the coexpression patterns of protein targets followed by an analysis of the available literature, we tested ten proteins in three established melanoma cell lines, using a normal keratinocyte cell line as a control, to find that at least eight of these proteins were enriched in melanoma cells. Some of the proteins with the highest selectivity include CLN3 (Ceroid-Lipofuscinosis Neuronal 3 or battenin), MFRP (Membrane Frizzled-

Related Protein), CXCR3 (C-X-C Motif Chemokine Receptor 3) and IL-13RA2 (Interleukin 13 Receptor Subunit Alpha 2), which have shown varying degrees of involvement in cancer and other diseases. While these proteins could be used in traditional chemotherapeutic approaches or as melanoma antigens for immunotherapy, combinations of these targets could be harnessed to drive heteromultivalent strategies. Furthermore, the screening method used, along with the corresponding scripts, could be used for diseases beyond cancer, especially due to the availability and relevance of surface-bound proteins in targeting and signaling. Future studies should validate protein coexpression on the cell surface of melanoma cell lines and patient biopsies and explore the design and synthesis of heteromultivalent ligands for validated combinations.

Finally, our studies on the resazurin assay revealed the existence of common inadequate reporting and optimization of analysis conditions in the scientific literature. A literature survey using the SciFinder® database showed a substantial and growing use of resazurin-based assays for the evaluation of cytotoxicity and proliferation, with a significant proportion of these articles failing to report minimum parameters required for experimental replication. Moreover, most authors failed to acknowledge whether these parameters had been validated in their own laboratories, since in some cases the assay conditions had been selected based on publications that utilized different cell lines. This practice is troublesome given that the resazurin assay is based on the metabolism of the title chromophore, the rate of which is cell line-dependent and cannot be extrapolated from tissue of origin (e.g., breast *vs.* pancreas, or healthy *vs.* cancer), kinetics of growth (i.e., doubling time) or energetic parameters (e.g., NADH or ATP content). Furthermore, our enzymatic inhibition assays suggest that resazurin reduction might not be promoted by specific reductases inside the cell or attributed exclusively to cellular respiration, which should be accounted for when interpreting the results of this assay. We hope that the recommendations made

with respect to reporting standards along with the protocol for experimental validation of the resazurin assay will help the scientific community improve the validity and reproducibility of future studies.

It is not surprising that cancer remains a deeply feared disease in the general population in spite of the massive amount of funding and effort invested in developing effective new detection and treatment methods. Cancer is, in fact, a complex array of related diseases that originate from different tissues via diverse mechanisms and that display radically different behaviors and drug sensitivities (e.g., cisplatin is effective in testicular cancer but ineffective in melanoma). Moreover, intra- and intertumor heterogeneity, i.e., genetic and phenotypic differences among cells in a given tumor or across patients with the same type of tumor, restrict drug efficacy and contribute to the development of resistance. Unlike other diseases that arise from the malfunction of one specific protein, cancer clinicians must fight an ever-evolving target. At times, the treatment of cancer resembles the mythical battle of Hercules with the Lernaean Hydra, where each decapitation of this monster led to the growth of two new heads. In this Greek story, Hercules overcame this singular challenge by recruiting his nephew Iolaus, who cauterized the open stumps thus precluding the appearance of new heads. In turn, cancer researchers must work together to harness our individual strengths and approach this ever-evolving monster with interdisciplinary strategies. The studies described in this dissertation represent our contributions to the field at the interface of chemistry and biology, in hopes that they might improve our understanding of drug activation and accumulation (e.g., PDT and heteromultivalency, respectively) and its eventual evaluation in preclinical models (e.g., resazurin assays). In doing so we remain true to the final goal of scientific pursuit as it impacts all of humanity and, as stated by Begley & Ellis (2012), “we in the field must remain focused on the purpose of cancer research: to improve the lives of patients.”

7. APPENDICES

7.1. Appendix A. Python scripts using in the bioinformatics analysis

7.1.1. Script 1: Single protein identification and bivalent and trivalent combination screening

```
import statistics
import copy

print("General guidelines:")
print("1. Patient and normal sample lists must have one patient/sample per line, starting with protein names and then followed by expression values, all separated by commas. If samples are downloaded from Xena Browser, it is highly recommended that patient/samples with less than 5% of transcripts measured are deleted prior to using this script.")
print("2. Output file(s) will be overwritten every time the script is run. Save copies and/or change file names as needed.")
print("3. File names should include full directory location unless they are placed in the same folder as the script. Do not forget to write the file's extension")
print("4. You will be asked to propose thresholds and minimum patient coverages for cancer patients and normal tissue samples. The script will return the number of protein targets that fit those parameters and ask if the latter need changed, in which case you will need to provide new output file names.")
print("5. Proteins must be presented in the same order in all three normal ensembles. Proteins in the cancer patients can have a different order than the normal samples.")
cancer_file=open(input("Enter file with cancer patient expression:"),'r')
```

```
#norm_file_1=open(input("Enter first file with normal tissue expression:"),'r')
#norm_file_2=open(input("Enter second file with normal tissue expression:"),'r')
#norm_file_3=open(input("Enter third file with normal tissue expression:"),'r')
output_file=open(input("Enter a name for the file with protein names and coverage in cancer and normal tissues:"), "w")

#these five lines are left here for debugging, in case the script needs to be modified. This saves time by avoiding the need for user input
#cancer_file=open("mel.csv",'r')
norm_file_1=open("GTEx_1.csv",'r')
norm_file_2=open("GTEx_2.csv",'r')
norm_file_3=open("GTEx_3.csv",'r')
#output_file=open("single_target_cov.txt",'w')

all_cancer=[] #list of lists of floaters, will hold lists of expression values per patient without protein names
prot_cancer=[] #list of strings, will contain protein names
expr_cancer=[] #list of floaters, will contain all expression values
for sample_str in cancer_file: #analyze cancer file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each lines and splits the line string into a list
    prot_cancer.append(sample_list[0]) #store protein names in prot_cancer
```

```
    sample_i=[] #starts/restarts a list for temporary storage of expression values for a particular protein across patients
    for expr in sample_list[1:]:
        sample_i.append(float(expr)) #temporarily stores expression values per patient in sample_i
        expr_cancer.append(float(expr)) #stores expression values for array's median calculation
    all_cancer.append(sample_i) #when expression for all patients has been stored in the temporary list "sample_i", export expression values for the protein to list "all_cancer"

all_norm_1=[]
prot_norm=[]
expr_norm_1=[]
for sample_str in norm_file_1: #analyze normal tissue file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each line and splits the line string into a list
    prot_norm.append(sample_list[0]) #store protein names in prot_norm
    sample_i=[]
    for expr in sample_list[1:]:
        sample_i.append(float(expr))
        expr_norm_1.append(float(expr))
    all_norm_1.append(sample_i) #store expression values of each protein through all normal tissue samples

all_norm_2=[]
```

```
expr_norm_2=[]
for sample_str in norm_file_2: #analyze normal tissue file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each line and splits the line string into a list
    sample_i=[]
    for expr in sample_list[1:]:
        sample_i.append(float(expr))
        expr_norm_2.append(float(expr))
    all_norm_2.append(sample_i) #store expression values of each protein through all normal tissue samples

all_norm_3=[]
expr_norm_3=[]
for sample_str in norm_file_3: #analyze normal tissue file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each line and splits the line string into a list
    sample_i=[]
    for expr in sample_list[1:]:
        sample_i.append(float(expr))
        expr_norm_3.append(float(expr))
    all_norm_3.append(sample_i) #store expression values of each protein through all normal tissue samples
```

```
cancer_file.close() #closes files that will not be needed any more
norm_file_1.close()
norm_file_2.close()
norm_file_3.close()

compliance_str="N"
while compliance_str=="N":
    thres_canc=float(input('Input percentage threshold for cancer patients (e.g., "0.75" stands for "75% of the median of the cancer array"):'))
    cov_canc=100*float(input('Input minimum acceptable cancer patient coverage (e.g., "0.10" stands for protein expression in "at least 10% of patients"):'))
    thres_norm=float(input('Input percentage threshold for normal samples (e.g., "0.65" stands for "65% of the median of the normal tissue array"):'))
    cov_norm=100*float(input('Input maximum acceptable normal tissue coverage (e.g., "0.25" stands for protein expression in "no more than 25% of samples"):'))
    n_thres=0
    n_options=[1,2,3]
    #thres_canc=0.75 #these five lines are left here for debugging, in case the script needs to be modified. This saves time by avoiding the need for user input
    #cov_canc=10
    #thres_norm=0.65
    #cov_norm=25
    #n_thres=2
```

```
    while n_thres==0:
        n_thres_temp=int(input("In how many normal tissue ensembles must the protein be present in to be selected (1, 2 or 3 only)?"))
        if n_thres_temp in n_options:
            n_thres=n_thres_temp
        else:
            print("Must be 1, 2 or 3 only")
    prot_num=0
    prot_list=[]
    thres_canc_rsem=thres_canc*statistics.median(expr_cancer) #calculate the expression threshold for the cancer patients by multiplying the user-defined percentage and the median of the cancer array
    thres_norm_rsem_1=thres_norm*statistics.median(expr_norm_1) #calculate the expression threshold for the normal samples by multiplying the user-defined percentage and the median of the normal tissue array
    thres_norm_rsem_2=thres_norm*statistics.median(expr_norm_2)
    thres_norm_rsem_3=thres_norm*statistics.median(expr_norm_3)
    print("Data hereafter was obtained using a cancer threshold of {}x(median of the cancer array), which corresponds to {:.4f}, and a minimum cancer patient coverage of {}%, as well as a normal tissue threshold of {}x(median of each normal array), and a maximum normal tissue sample coverage of {}%".format(thres_canc,thres_canc_rsem,cov_canc,thres_norm,cov_norm), file=output_file)
    print("Protein, cancer coverage (%), normal tissue coverage (%)", file=output_file)
```

```
    patient_num=len(all_cancer[0]) #calculates total number of patients
    sample_num=len(all_norm_1[0]) #calculates total number of normal tissue samples
    canc_index=0
    prot_count=0
    all_cancer_trimmed=[] #starts a list for cancer samples after screening. Only protein targets that meet criteria for cancer AND  normal tissues are exported into this AND OTHER TRIMMED lists
    all_norm_trimmed_1=[] #starts a list for normal samples from ensemble 1 after screening. Expression values on trimmed normal lists are placed in the same order as in cancer trimmed list (protein 1 in cancer is also protein 1 in normal and protein 1 in prot_list to further simplify coexpression analysis later on)
    all_norm_trimmed_2=[]
    all_norm_trimmed_3=[]
    for prot_i in all_cancer:
        patient_count=0
        for patient_i in prot_i:
            if patient_i>=thres_canc_rsem:
                patient_count+=1 #counts patients where the protein is "expressed" (i.e., at or above the threshold)
        cov_canc_prot=patient_count/patient_num*100 #calculates the percentage of cancer patients where the protein is "expressed"
        if cov_canc_prot>=cov_canc and prot_cancer[canc_index] in prot_norm: #checks if protein satisfies minimum patient coverage and if there is data in normal tissues
            prot_norm_index=prot_norm.index(prot_cancer[canc_index])
```

```
        sample_count=0
        cov_norm_list=[]
        cov_norm_count=0
        for sample_i in all_norm_1[prot_norm_index]:
            if sample_i>=thres_norm_rsem_1:
                sample_count+=1 #counts normal tissue samples in first ensemble where the protein is "expressed" (i.e., at or above the threshold)
        cov_norm_list.append(sample_count/sample_num*100) #calculates normal tissue coverage
        sample_count=0
        for sample_i in all_norm_2[prot_norm_index]:
            if sample_i>=thres_norm_rsem_2:
                sample_count+=1 #counts normal tissue samples in second ensemble where the protein is "expressed" (i.e., at or above the threshold)
        cov_norm_list.append(sample_count/sample_num*100) #calculates normal tissue coverage
        sample_count=0
        for sample_i in all_norm_3[prot_norm_index]:
            if sample_i>=thres_norm_rsem_3:
                sample_count+=1 #counts normal tissue samples in third ensemble where the protein is "expressed" (i.e., at or above the threshold)
        cov_norm_list.append(sample_count/sample_num*100) #calculates normal tissue coverage
```

```
        for cov_i in cov_norm_list:
            if cov_i<=cov_norm:
                cov_norm_count+=1 #checks if protein satisfies maximum normal tissue coverage in
each ensemble
        if cov_norm_count>=n_thres: #checks if protein is under threshold in as many ensembles
as defined above
            prot_count+=1 #counts number of protein hits
            prot_list.append(prot_cancer[canc_index])
            all_cancer_trimmed.append(all_cancer[canc_index])
            all_norm_trimmed_1.append(all_norm_1[prot_norm_index])
            all_norm_trimmed_2.append(all_norm_2[prot_norm_index])
            all_norm_trimmed_3.append(all_norm_3[prot_norm_index])

print("{},{:.2f},{:.2f}".format(prot_cancer[canc_index],cov_canc_prot,statistics.mean(cov_norm
_list)),file=output_file) #saves protein name, cancer coverage and normal tissue coverage in
corresponding file
        canc_index+=1
    output_file.close()

    while compliance_str=="N":
        compliance_temp=input('{} proteins identified. Are the threshold and patient coverage
acceptable? (Answer "Y" or "N"):'.format(prot_count))
        compliance_temp=compliance_temp.upper()
```

```
        if compliance_temp=='Y':
            compliance_str=compliance_temp
        elif compliance_temp=="N":
            output_file=open(input("Enter a new name for the file with protein names and coverage in cancer and normal tissues:"), "w")
            print("Please set new thresholds and minimum coverages.")
            break
        else:
            print('Please answer "Y" or "N" only.')


prot_sorted=sorted(prot_list)
all_comb_biv=[] #generates a list for combinations
all_comb_tri=[]


for prot_a in prot_sorted: #goes through the protein list and picks the first protein in the combination
    for prot_b in prot_sorted[prot_sorted.index(prot_a)+1:]: #to get the second protein in the combination, it moves one position over in the protein list
        all_comb_biv.append([prot_a,prot_b])
        for prot_c in prot_sorted[prot_sorted.index(prot_b)+1:]:#to get the third protein in the combination, it moves one more position over in the protein list
                all_comb_tri.append([prot_a,prot_b,prot_c]) #appends the three proteins as a list
```

```
output_biv=open(input("Enter name for the file with bivalent combination coverage across cancer and normal samples:"), "w")
output_tri=open(input("Enter name for the file with trivalent combination coverage across cancer and normal samples:"), "w")
print("Data hereafter is organized as follows: protein a, protein b, % bivalent coverage of cancer patients, % bivalent coverage of normal tissues (average of three ensembles)",file=output_biv)
print("Data hereafter is organized as follows: protein a, protein b, protein c, % cancer coverage as a trivalent binder (% of patients where all three proteins are expressed), % cancer coverage of bivalent subsets (% of patients where at least two of the proteins are expressed), % normal coverage as a trivalent binder (average of three ensembles), % normal coverage of bivalent subsets (average of three ensembles)",file=output_tri)

binary_cancer=[] #starts a list binary conversion of expression values in cancer samples that met the screening criteria
for expr_list in all_cancer_trimmed:
    sample_i=[]
    for exp_i in expr_list: #converts expression values to a binary system where 1=expressed (equal or higher than threshold) and 0=not expressed (under threshold)
        if exp_i>=thres_canc_rsem:
            sample_i.append(1)
        else:
            sample_i.append(0)
```

```
    binary_cancer.append(sample_i) #appends all expression values for each protein to binary_cancer list

binary_norm_1=[] #starts a list binary conversion of expression values in normal ensemble 1 samples that meet the screening criteria
for expr_list in all_norm_trimmed_1:
    sample_i=[]
    for exp_i in expr_list: #converts expression values to a binary system where 1=expressed (equal or higher than threshold) and 0=not expressed (under threshold)
        if exp_i>=thres_norm_rsem_1:
            sample_i.append(1)
        else:
            sample_i.append(0)
    binary_norm_1.append(sample_i) #appends all expression values for each protein to binary_norm_2 list

binary_norm_2=[] #starts a list binary conversion of expression values in normal ensemble 2 samples that meet the screening criteria
for expr_list in all_norm_trimmed_2:
    sample_i=[]
    for exp_i in expr_list: #converts expression values to a binary system where 1=expressed (equal or higher than threshold) and 0=not expressed (under threshold)
        if exp_i>=thres_norm_rsem_2:
```

```
            sample_i.append(1)
        else:
            sample_i.append(0)
    binary_norm_2.append(sample_i) #appends all expression values for each protein to binary_norm_3 list

binary_norm_3=[] #starts a list binary conversion of expression values in normal ensemble 3 samples that meet the screening criteria
for expr_list in all_norm_trimmed_3:
    sample_i=[]
    for exp_i in expr_list: #converts expression values to a binary system where 1=expressed (equal or higher than threshold) and 0=not expressed (under threshold)
        if exp_i>=thres_norm_rsem_3:
            sample_i.append(1)
        else:
            sample_i.append(0)
    binary_norm_3.append(sample_i) #appends all expression values for each protein to binary_norm_1 list

for comb_i in all_comb_biv: #checks one bivalent combination at a time
    prot_a_index=prot_list.index(comb_i[0])
    prot_b_index=prot_list.index(comb_i[1])
    patient_i=0
```

```
    sample_i=0
    cov_biv_canc=0
    cov_biv_norm_1=0
    cov_biv_norm_2=0
    cov_biv_norm_3=0
    while patient_i<patient_num: #check one patient at a time in the binarized cancer ensemble
        if binary_cancer[prot_a_index][patient_i]+binary_cancer[prot_b_index][patient_i]==2:
            cov_biv_canc+=1
        patient_i+=1
    while sample_i<sample_num: #check one patient at a time in each of the binarized normal tissue ensembles
        if binary_norm_1[prot_a_index][sample_i]+binary_norm_1[prot_b_index][sample_i]==2:
            cov_biv_norm_1+=1
        if binary_norm_2[prot_a_index][sample_i]+binary_norm_2[prot_b_index][sample_i]==2:
                cov_biv_norm_2+=1
        if binary_norm_3[prot_a_index][sample_i]+binary_norm_3[prot_b_index][sample_i]==2:
                cov_biv_norm_3+=1
        sample_i+=1
    cov_biv_norm_avg=(cov_biv_norm_1+cov_biv_norm_2+cov_biv_norm_3)/3 #calculates the average bivalent coverage of normal tissues (in number of samples, not percentage)

print("{},{},{:.2f},{:.2f}".format(comb_i[0],comb_i[1],cov_biv_canc/patient_num*100,cov_biv_norm_avg/sample_num*100),file=output_biv)
```

```
for comb_i in all_comb_tri: #checks one combination at a time
    prot_a_index=prot_list.index(comb_i[0])
    prot_b_index=prot_list.index(comb_i[1])
    prot_c_index=prot_list.index(comb_i[2])
    patient_i=0
    sample_i=0
    cov_biv_canc=0
    cov_biv_norm_1=0
    cov_biv_norm_2=0
    cov_biv_norm_3=0
    cov_tri_canc=0
    cov_tri_norm_1=0
    cov_tri_norm_2=0
    cov_tri_norm_3=0
    while patient_i<patient_num: #check one sample/patient at a time
        if
binary_cancer[prot_a_index][patient_i]+binary_cancer[prot_b_index][patient_i]+binary_cancer[
prot_c_index][patient_i]==3:
            cov_biv_canc+=1
            cov_tri_canc+=1
```

```
        elif
binary_cancer[prot_a_index][patient_i]+binary_cancer[prot_b_index][patient_i]+binary_cancer[
prot_c_index][patient_i]==2:
            cov_biv_canc+=1
        patient_i+=1
    while sample_i<sample_num: #check one patient at a time in each of the binarized normal tissue
ensembles
        if
binary_norm_1[prot_a_index][sample_i]+binary_norm_1[prot_b_index][sample_i]+binary_nor
m_1[prot_c_index][sample_i]==3:
            cov_tri_norm_1+=1
            cov_biv_norm_1+=1
        elif
binary_norm_1[prot_a_index][sample_i]+binary_norm_1[prot_b_index][sample_i]+binary_nor
m_1[prot_c_index][sample_i]==2:
            cov_biv_norm_1+=1
        if
binary_norm_2[prot_a_index][sample_i]+binary_norm_2[prot_b_index][sample_i]+binary_nor
m_2[prot_c_index][sample_i]==3:
            cov_tri_norm_2+=1
            cov_biv_norm_2+=1
```

```
        elif
binary_norm_2[prot_a_index][sample_i]+binary_norm_2[prot_b_index][sample_i]+binary_nor
m_2[prot_c_index][sample_i]==2:
            cov_biv_norm_2+=1
        if
binary_norm_3[prot_a_index][sample_i]+binary_norm_3[prot_b_index][sample_i]+binary_nor
m_3[prot_c_index][sample_i]==3:
            cov_tri_norm_3+=1
            cov_biv_norm_3+=1
        elif
binary_norm_3[prot_a_index][sample_i]+binary_norm_3[prot_b_index][sample_i]+binary_nor
m_3[prot_c_index][sample_i]==2:
            cov_biv_norm_3+=1
        sample_i+=1
    cov_biv_norm_avg=(cov_biv_norm_1+cov_biv_norm_2+cov_biv_norm_3)/3 #calculates the
average bivalent coverage of normal tissues (in number of samples, not percentage)
    cov_tri_norm_avg=(cov_tri_norm_1+cov_tri_norm_2+cov_tri_norm_3)/3

print("{},{},{},{:.2f},{:.2f},{:.2f},{:.2f}".format(comb_i[0],comb_i[1],comb_i[2],cov_tri_canc/
patient_num*100,cov_biv_canc/patient_num*100,cov_tri_norm_avg/sample_num*100,cov_biv
_norm_avg/sample_num*100),file=output_tri)

output_biv.close()
```

```
output_tri.close()

print("Done. {} bivalent combinations and {} trivalent combinations were analyzed across {} cancer patients and three ensembles of {} normal tissue samples each.".format(len(all_comb_biv),len(all_comb_tri),patient_num,sample_num))
```

7.1.2. Script 2. Single target identification

```
import statistics


print("General guidelines:")

print("1. Patient and normal sample lists must have one patient/sample per line, starting with protein names and then followed by expression values, all separated by commas. If samples are downloaded from Xena Browser, it is highly recommended that patient/samples with less than 5% of transcripts measured are deleted prior to using this script.")

print("2. Output file will be overwritten every time the script is run. Save copies and/or change file names as needed.")

print("3. File names should include full directory location unless they are placed in the same folder as the script. Do not forget to write the file's extension")

print("4. You will be asked to propose thresholds and minimum patient coverages for cancer patients and normal tissue samples. The script will return the number of protein targets that fit those parameters and ask if the latter need changed, in which case you will need to provide new output file names.")

cancer_file=open(input("Enter file with cancer patient expression:"),'r')

norm_file=open(input("Enter file with normal tissue expression:"),'r')

output_file=open(input("Enter a name for the file with protein names and coverage in cancer and normal tissues:"), "w")
```

```
output_prot=open(input("Enter a name for the file with protein names only:"),"w")


#these four lines are left here for debugging, in case the script needs to be modified. This saves time by avoiding the need for user input
#cancer_file=open("mel.csv",'r')
#norm_file=open("GTEx_1.csv",'r')
#output_file=open("test.txt",'w')
#output_prot=open("test_prot.txt",'w')


all_cancer=[] #list of lists of floaters, will hold lists of expression values per patient without protein names
prot_cancer=[] #list of strings, will contain protein names
expr_cancer=[] #list of floaters, will contain all expression values
for sample_str in cancer_file: #analyze cancer file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each lines and splits the line string into a list
    prot_cancer.append(sample_list[0]) #store protein names in prot_cancer
    sample_i=[] #starts/restarts a list for temporary storage of expression values for a particular protein across patients
```

```
    for expr in sample_list[1:]:
        sample_i.append(float(expr)) #temporarily stores expression values per patient in sample_i
        expr_cancer.append(float(expr)) #stores expression values for array's median calculation
    all_cancer.append(sample_i) #when expression for all patients has been stored in the temporary list "sample_i", export expression values for the protein to list "all_cancer"


all_norm=[]
prot_norm=[]
expr_norm=[]
for sample_str in norm_file: #analyze normal tissue file
    sample_list=sample_str.strip().split(',') #removes '\n' at the end of each line and splits the line string into a list
    prot_norm.append(sample_list[0]) #store protein names in prot_norm
    sample_i=[]
    for expr in sample_list[1:]:
        sample_i.append(float(expr))
        expr_norm.append(float(expr))
    all_norm.append(sample_i) #store expression values of each protein through all normal tissue samples
```

```
compliance_str="N"

while compliance_str=="N":

    thres_canc=float(input('Input percentage threshold for cancer patients (e.g., "0.75" stands for "75% of the median of the cancer array"):'))

    cov_canc=100*float(input('Input minimum acceptable cancer patient coverage (e.g., "0.10" stands for protein expression in "at least 10% of patients"):'))

    thres_norm=float(input('Input percentage threshold for normal samples (e.g., "0.65" stands for "65% of the median of the normal tissue array"):'))

    cov_norm=100*float(input('Input maximum acceptable normal tissue coverage (e.g., "0.25" stands for protein expression in "no more than 25% of samples"):'))

    #thres_canc=0.75 #these four lines are left here for debugging, in case the script needs to be modified. This saves time by avoiding the need for user input

    #cov_canc=10

    #thres_norm=0.65

    #cov_norm=25

    prot_num=0

    thres_canc_rsem=thres_canc*statistics.median(expr_cancer) #calculate the expression threshold for the cancer patients by multiplying the user-defined percentage and the median of the cancer array
```

thres_norm_rsem=thres_norm*statistics.median(expr_norm) #calculate the expression threshold for the normal samples by multiplying the user-defined percentage and the median of the normal tissue array

print("Data hereafter was obtained using a cancer threshold of {}x(median of the cancer array), which corresponds to {:.4f}, and a minimum cancer patient coverage of {}%, as well as a normal tissue threshold of {}x(median of the normal array), which corresponds to {:.4f}, and a maximum normal tissue sample coverage of {}%".format(thres_canc,thres_canc_rsem,cov_canc,thres_norm,thres_norm_rsem,cov_norm), file=output_file)

print("Protein, cancer coverage (%), normal tissue coverage (%)", file=output_file)

patient_num=len(all_cancer[0]) #calculates total number of patients

sample_num=len(all_norm[0]) #calculates total number of normal tissue samples

canc_index=0

prot_count=0

for prot_i in all_cancer:

patient_count=0

for patient_i in prot_i:

if patient_i>=thres_canc_rsem:

patient_count+=1 #counts patients where the protein is "expressed" (i.e., at or above the threshold)

```
        cov_canc_prot=patient_count/patient_num*100 #calculates the percentage of cancer
patients where the protein is "expressed"
        if cov_canc_prot>=cov_canc: #checks if protein satisfies minimum patient coverage
            prot_norm_index=prot_norm.index(prot_cancer[canc_index])
            sample_count=0
            for sample_i in all_norm[prot_norm_index]:
                if sample_i>=thres_norm_rsem:
                    sample_count+=1 #counts normal tissue samples where the protein is "expressed"
(i.e., at or above the threshold)
            cov_norm_prot=sample_count/sample_num*100 #calculates normal tissue coverage
            if cov_norm_prot<=cov_norm: #checks if protein satisfies maximum normal tissue
coverage
                prot_count+=1 #counts number of protein hits
                print(prot_cancer[canc_index],file=output_prot) #saves protein name in corresponding
file

print("{},{:.2f},{:.2f}".format(prot_cancer[canc_index],cov_canc_prot,cov_norm_prot),file=out
put_file) #saves protein name, cancer coverage and normal tissue coverage in corresponding file
        canc_index+=1
    output_file.close()
```

```
    output_prot.close()


    while compliance_str=="N":
        compliance_temp=input('{} proteins identified. Are the threshold and patient coverage acceptable? (Answer "Y" or "N"):'.format(prot_count))
        compliance_temp=compliance_temp.upper()
        if compliance_temp=='Y':
            compliance_str=compliance_temp
        elif compliance_temp=="N":
            output_file=open(input("Enter a new output file name (proteins+coverage in cancer and normal tissues):"), "w")
            output_prot=open(input("Enter a new output file name (protein names only):"),"w")
            print("Please set new thresholds and minimum coverages.")
            break
        else:
            print('Please answer "Y" or "N" only.')


cancer_file.close()
norm_file.close()
```

7.1.3. Script 3: Combination generator

```
data_input=open(input("Enter protein list file name:"),'r')
data_output=open(input("Enter output file name for the combinations:"),'w')

k_int=0 #starts a variable for the number of proteins per combinations
while k_int==0: #starts a loop
    k_temp=int(input("How many proteins to pick (2 or 3 only)?")) #asks how many proteins to pick
    if k_temp==3 or k_temp==2:
        k_int=k_temp #stores the number of proteins per combination in the k_int variable only if it's 2 or 3
    else: #ensures only bivalent and trivalent combinations are generated
        print("This script only generates 2- or 3-protein combinations. Choose again.")

prot_list=[]
for line_str in data_input:
    prot_list.append(line_str.strip()) #stores each protein in the file into the list "prot_list" after getting rid of the "enter" at the end of each line
```

```
prot_list.sort() #sorts the proteins in the list, this is for aesthetic purposes mostly

comb_count=0 #starts a variable to count the number of combinations generated


for prot_a in prot_list: #goes through the protein list and picks the first protein in the combination

    for prot_b in prot_list[prot_list.index(prot_a)+1:]: #to get the second protein in the combination, it moves one position over in the protein list

        if k_int==3: #for trivalent combinations

            for prot_c in prot_list[prot_list.index(prot_b)+1:]:#to get the third protein in the combination, it moves one more position over in the protein list

                print("{},{},{}".format(prot_a,prot_b,prot_c), file=data_output) #puts all three proteins together into a combination, proteins are separated by commas

                comb_count+=1 #adds 1 to the total number of combinations

        else: #for bivalent combinations

            print("{},{}".format(prot_a,prot_b), file=data_output) #puts both proteins together into a combinations, proteins are separated by a comma

            comb_count+=1 #adds 1 to the total number of combinations
```

```
print("{} proteins were input and {} combinations were
generated.".format(len(prot_list),comb_count)) #prints the total number of proteins and
combinations generated.


data_input.close() #closes the files

data_output.close()
```

7.1.4. Script 4: Combination evaluator

```
print("General guidelines:")
print("1. Protein combinations must be in one file, one protein combination per line, as comma-separated values.")
print("2. Patient/sample list must have one patient/sample per line, starting with protein names and then followed by expression values, all separated by commas.")
print("3. Output file names must be different each time the script is run. Otherwise, they will be overwritten.")
print('4. Threshold for expression must be provided for the script. The file generated with the "single target identification" script contains this value.')
print("5. File names should include full directly unless they are placed in the same folder as the script. Do not forget to write the file's extension")

data_file=open(input("Enter name of file with list of combinations:"),'r') #asks for the file with the list of combinations
samples_file=open(input("Enter name of file with list of patients/samples:"),'r')
output_file=open(input("Enter output file name:"), "w")
thres_float=float(input("Enter threshold:")) #obtain defined threshold

all_comb=[] #generates a list for combinations. After the next two lines, it becomes a list of lists of proteins
for line_str in data_file:
```

```
    all_comb.append(line_str.strip().split(',')) #removes '\n' at the end of each line and splits the line string into a list, then adds the list to the list of combinations

all_samples=[] #generates a list of expression values
prot_list=[] #generates a list of protein names
for sample_str in samples_file:
    sample_list=sample_str.strip().split(',') #splits line string by commas
    prot_list.append(sample_list[0]) #stores the first item in the list (protein name) in the protein list
    sample_i=[]
    for exp_i in sample_list[1:]: #converts expression values to a binary system where 1=expressed (equal or higher than threshold) and 0=not expressed (under threshold)
        if float(exp_i)>=thres_float:
            sample_i.append(1)
        else:
            sample_i.append(0)
    all_samples.append(sample_i) #adds all expression values for each protein to all_samples list

k_int=len(all_comb[0]) #checks if the combination is bivalent or trivalent
if k_int==2:
    print("Data hereafter is organized as follows: protein a, protein b, % coverage of the bivalent combination (% of patients/samples where both proteins are expressed)",file=output_file)
elif k_int==3:
```

```
    print("Data hereafter is organized as follows: protein a, protein b, protein c, % coverage as a trivalent binder (% of patients/samples where all three proteins are expressed), % coverage of bivalent subsets (% of patients where at least two of the proteins are expressed)",file=output_file)
sample_number=len(all_samples[0]) #determine total number of samples/patients

for comb_i in all_comb: #checks one combination at a time
    prot_a_index=prot_list.index(comb_i[0])
    prot_b_index=prot_list.index(comb_i[1])
    sample_i=0
    sample_cov_biv=0
    if k_int==2:
        while sample_i<sample_number: #check one sample/patient at a time
            if all_samples[prot_a_index][sample_i]+all_samples[prot_b_index][sample_i]==2:
                sample_cov_biv+=1
            sample_i+=1

print("{},{},{:.2f}".format(comb_i[0],comb_i[1],sample_cov_biv/sample_number*100),file=output_file)
    elif k_int==3:
        sample_cov_tri=0
        prot_c_index=prot_list.index(comb_i[2])
        while sample_i<sample_number: #check one sample/patient at a time
```

```
            if
all_samples[prot_a_index][sample_i]+all_samples[prot_b_index][sample_i]+all_samples[prot_c
_index][sample_i]==3:
                sample_cov_biv+=1
                sample_cov_tri+=1
            elif
all_samples[prot_a_index][sample_i]+all_samples[prot_b_index][sample_i]+all_samples[prot_c
_index][sample_i]==2:
                sample_cov_biv+=1
            sample_i+=1

print("{},{},{},{:.2f},{:.2f}".format(comb_i[0],comb_i[1],comb_i[2],sample_cov_tri/sample_nu
mber*100,sample_cov_biv/sample_number*100),file=output_file)
    else:
        print("Program only works on bivalent or trivalent combinations!")
        break

samples_file.close()
data_file.close()
output_file.close()

print("Done. {} combinations were analyzed across {}
samples/patients.".format(len(all_comb),sample_number))
```

7.1.5. Script 5: Ranking based on likelihood to generate good combinations

```
print("General guidelines:")
print("1. Protein list must contain one protein per line.")
print("2. Protein combinations must be in one file, one protein combination per line, as comma-separated values.")
print("3. Output file names must be different each time the script is run. Otherwise, they will be overwritten.")
print("4. File names should include full directly unless they are placed in the same folder as the script. Do not forget to write the file's extension.")

#data_biv=open(input("Enter name of file with list of bivalent combinations:"),'r') #asks for the file with the list of bivalent combinations
#data_tri=open(input("Enter name of file with list of trivalent combinations:"),'r') #asks for the file with the list of trivalent combinations

data_biv=open("mel_biv.csv",'r')
data_tri=open("mel_tri.csv",'r')

comb_biv=[] #generates a list for combinations. After the next two lines, it becomes a list of lists of proteins
for line_str in data_biv:
```

```
    comb_biv.append(line_str.strip().split(',')) #removes '\n' at the end of each line and splits the line string into a list, then adds the list to the list of combinations

comb_tri=[]
for line_str in data_tri:
    comb_tri.append(line_str.strip().split(',')) #removes '\n' at the end of each line and splits the line string into a list, then adds the list to the list of combinations

prot_temp=[]
for comb_i in comb_biv: #generates list of proteins
    for prot_i in comb_i:
        prot_temp.append(prot_i)

prot=list(set(prot_temp)) #removes duplicates

const=""
while const=="":
    thres_int=int(input('How many combinations to count from the top (e.g. "10" stands for best 10 combinations that the protein is part of):')) #obtain defined threshold
    output_file=open(input("Enter output file name:"), "w")
    #best_num_int=10 #these two lines left here for debugging
    #output_file=open("mel_best10.txt",'w')
```

```
    count_biv=[] #starts a list that will add up the "best_num_int"(e.g. 5 best, 10 best, etc.) bivalent combinations
    count_tri=[]
    for prot_i in prot: #checks one combination at a time
        count_i=0 #how many times the protein has been found in a combination starting from the top
        top_biv=0 #variable to which bivalent combination's rank is added
        index=1
        for comb_i in comb_biv: #go combination by combination
            if count_i<thres_int: #check if the best_num_int bivalent combinations have already been found, if not, move on
                if prot_i in comb_i:
                    top_biv+=index #add combination's rank if the protein is part of the combination, protein with the best ranking has the lowest "prot_best_biv" value
                    count_i+=1
            else: #if the method has found the best_num_int bivalent combinations, break the "for"-loop so it doesn't check any more combinations
                break
            index+=1
        count_biv.append(top_biv)
        count_i=0
        top_tri=0 #variable to which trivalent combination's rank is added. See comments for bivalent combinations
```

```
        index=1
        for comb_i in comb_tri:
            if count_i<thres_int:
                if prot_i in comb_i:
                    top_tri+=index
                    count_i+=1
            else:
                break
            index+=1
        count_tri.append(top_tri)


    sorted_biv=sorted(count_biv) #sorts counts based on bivalent combinations
    sorted_tri=sorted(count_tri) #sorts counts based on trivalent combinations
    rank_biv=[1+sorted_biv.index(count) for count in count_biv] #makes a new list that contains the ranking. Basically, it goes through the count_biv file and retrieves each count, then finds the count in the list and returns its location (index). Finally, it adds 1 since lists are numbered starting in 0
    rank_tri=[1+sorted_tri.index(count) for count in count_tri] #same for trivalent combinations
    average_rank=[]
    index=0
    while index<len(prot):
        avg_rank=(rank_biv[index]+rank_tri[index])/2
```

```
        average_rank.append(avg_rank) #calculates the average of the rankings based on bivalent and trivalent combinations
        index+=1

    sorted_avg_rank=sorted(average_rank)
    rank_final=[1+sorted_avg_rank.index(rank) for rank in average_rank] #performs the final ranking

    index=0
    for prot_i in prot: #checks one combination at a time
        print("{},{}".format(prot_i,rank_final[index]),file=output_file) #prints into file
        index+=1

    output_file.close()

    while const=="": #checks if other best_num_int's are to be tested
        const_temp=input('Change "n" best combinations and rescreen? (Y/N only):')
        if const_temp.upper()=="Y": #if yes, get out of this loop and return to the main loop
            break
        elif const_temp.upper()=="N": #if no, change the variable that controls this and the main loop, thus closing the loop and script
            const=const_temp
        else:
```

print('Answer "Y" or "N" only') #makes sure any other entry is ignored and the nested while loop is rerun

data_biv.close()

data_tri.close()

7.2. Appendix B

7.2.1. Curated surfaceome

Table 7.1. Surface protein-encoding genes represented in the curated surfaceome.

A4GALT	CSF3R	IL17RD	OR2W3	SIRPA
AAK1	CSK	IL18R1	OR2W5	SIRPB1
ABCA1	CSMD2	IL18RAP	OR2Y1	SIRPG
ABCA12	CSMD3	IL1R1	OR2Z1	SIT1
ABCA13	CSPG4	IL1R2	OR3A1	SKAP1
ABCA2	CSPG5	IL1RAP	OR3A2	SLAMF1
ABCA4	CTLA4	IL1RAPL1	OR3A3	SLAMF6
ABCA5	CTNNA1	IL1RAPL2	OR3A4	SLAMF7
ABCA6	CTNNB1	IL1RL1	OR4A5	SLAMF8
ABCA7	CTNS	IL1RL2	OR4B1	SLAMF9
ABCA8	CTTN	IL20RA	OR4C3	SLC10A1
ABCB1	CUTA	IL21R	OR4C6	SLC10A2
ABCB11	CUZD1	IL22RA1	OR4D1	SLC10A5
ABCB4	CX3CL1	IL23R	OR4D2	SLC11A1
ABCB5	CX3CR1	IL27RA	OR4D5	SLC11A2
ABCB6	CXADR	IL28RA	OR4D6	SLC12A1
ABCC1	CXCL16	IL2RA	OR4D9	SLC12A2
ABCC2	CXCR1	IL2RB	OR4E2	SLC12A3
ABCC3	CXCR2	IL2RG	OR4F3	SLC12A4
ABCC4	CXCR3	IL31RA	OR4F4	SLC12A5

ABCC5	CXCR4	IL3RA	OR4F5	SLC12A6
ABCC6	CXCR5	IL4R	OR4F6	SLC12A7
ABCC8	CXCR6	IL5RA	OR4K1	SLC12A8
ABCC9	CXCR7	IL6R	OR4K2	SLC12A9
ABCG1	CYB561D2	IL6ST	OR4K5	SLC13A1
ABCG2	CYBB	IL7R	OR4L1	SLC13A2
ABCG5	CYBRD1	IL9R	OR4M1	SLC13A5
ABCG8	CYSLTR1	ILDR1	OR4M2	SLC14A1
ABI1	CYSLTR2	INADL	OR4N2	SLC14A2
ACCN2	CYTH1	INSR	OR4N4	SLC15A1
ACE	CYTH2	INSRR	OR4N5	SLC15A2
ACE2	DAG1	IQGAP1	OR4P4	SLC15A3
ACHE	DARC	IRS4	OR4Q3	SLC15A4
ACKR3	DCBLD1	ISLR	OR4S1	SLC16A1
ACP2	DCBLD2	ISLR2	OR4S2	SLC16A10
ACSL6	DCC	ITCH	OR4X1	SLC16A2
ACTB	DCHS1	ITFG1	OR4X2	SLC16A3
ACTN1	DCHS2	ITGA1	OR51B2	SLC16A4
ACTN4	DCT	ITGA10	OR51B4	SLC16A5
ACVR1	DDR1	ITGA11	OR51B6	SLC16A6
ACVR1B	DDR2	ITGA2	OR51E1	SLC16A7
ACVR1C	DGCR2	ITGA2B	OR51F2	SLC16A8
ACVR2A	DGKG	ITGA3	OR51I1	SLC17A1

ACVR2B	DGKQ	ITGA4	OR51I2	SLC17A2
ACVRL1	DGKZ	ITGA5	OR51Q1	SLC17A3
ADA	DIRC2	ITGA6	OR51S1	SLC17A4
ADAM10	DISP1	ITGA7	OR52A1	SLC17A5
ADAM11	DISP2	ITGA8	OR52A4	SLC17A7
ADAM12	DLG1	ITGA9	OR52A5	SLC18A1
ADAM15	DLG2	ITGAD	OR52E4	SLC18A2
ADAM17	DLG4	ITGAE	OR52E8	SLC19A1
ADAM18	DLG5	ITGAL	OR52H1	SLC19A2
ADAM19	DLK1	ITGAM	OR52I1	SLC19A3
ADAM2	DLL1	ITGAV	OR52K1	SLC1A1
ADAM21	DLL4	ITGAX	OR52N4	SLC1A3
ADAM22	DMD	ITGB1	OR52N5	SLC1A4
ADAM23	DNAJA4	ITGB2	OR56A1	SLC1A5
ADAM28	DNAJC5	ITGB3	OR5A1	SLC1A6
ADAM29	DNER	ITGB4	OR5A2	SLC1A7
ADAM30	DOC2B	ITGB5	OR5AN1	SLC20A1
ADAM7	DOCK4	ITGB6	OR5AR1	SLC20A2
ADAM8	DPEP1	ITGB7	OR5B12	SLC22A11
ADAM9	DPP10	ITGB8	OR5B2	SLC22A12
ADAMTSL4	DPP4	ITM2A	OR5B3	SLC22A13
ADAP1	DPP6	ITM2B	OR5C1	SLC22A14
ADAP2	DPY19L4	ITM2C	OR5D13	SLC22A16

ADCY1	DRD1	ITPR1	OR5F1	SLC22A18
ADCY2	DRD2	ITPR2	OR5H1	SLC22A2
ADCY3	DRD3	ITPR3	OR5H2	SLC22A23
ADCY4	DRD4	ITPRIP	OR5H6	SLC22A3
ADCY5	DRD5	ITSN1	OR5I1	SLC22A4
ADCY6	DRP2	IYD	OR5J2	SLC22A5
ADCY7	DSC1	JAG1	OR5K1	SLC22A6
ADCY8	DSC2	JAG2	OR5K2	SLC22A7
ADCY9	DSC3	JAM2	OR5K3	SLC22A8
ADCYAP1R1	DSCAM	JAM3	OR5K4	SLC22A9
ADD3	DSCAML1	JPH1	OR5L1	SLC23A1
ADIPOR1	DSEL	JPH2	OR5L2	SLC23A2
ADIPOR2	DSG1	JPH3	OR5M1	SLC24A1
ADORA1	DSG2	JPH4	OR5M3	SLC24A2
ADORA2A	DSG3	JTB	OR5M8	SLC24A4
ADORA2B	DSG4	JUP	OR5M9	SLC24A6
ADORA3	DUOX1	KANK1	OR5P2	SLC26A1
ADRA1A	DUOX2	KCNA1	OR5P3	SLC26A11
ADRA1B	DUOXA1	KCNA10	OR5R1	SLC26A2
ADRA1D	DUSP15	KCNA2	OR5T1	SLC26A3
ADRA2A	DVL1	KCNA3	OR5T2	SLC26A4
ADRA2B	DYSF	KCNA4	OR5T3	SLC26A5
ADRA2C	ECE1	KCNA5	OR5V1	SLC26A6

ADRB1	ECEL1	KCNA6	OR5W2	SLC26A7
ADRB2	EDA	KCNA7	OR6A2	SLC26A8
ADRB3	EDA2R	KCNB1	OR6B1	SLC27A4
ADRBK2	EDAR	KCNB2	OR6B2	SLC28A1
ADRM1	EDNRA	KCNC1	OR6B3	SLC28A2
AGER	EDNRB	KCNC2	OR6C1	SLC29A1
AGRN	EFNA1	KCNC3	OR6C2	SLC29A2
AGTR1	EFNA2	KCNC4	OR6C3	SLC29A4
AGTR2	EFNA3	KCND1	OR6C4	SLC2A1
AHNAK	EFNA4	KCND2	OR6C6	SLC2A10
AJAP1	EFNA5	KCND3	OR6F1	SLC2A11
AKAP11	EFNB1	KCNE1	OR6K2	SLC2A12
AKAP12	EFNB2	KCNE1L	OR6K3	SLC2A13
AKT1	EFNB3	KCNE2	OR6K6	SLC2A14
AKT2	EFR3A	KCNE3	OR6M1	SLC2A2
AKTIP	EGF	KCNE4	OR6N1	SLC2A3
ALCAM	EGFR	KCNG1	OR6N2	SLC2A4
ALG10	EHD1	KCNG2	OR6Q1	SLC2A5
ALG10B	EHD2	KCNG3	OR6S1	SLC2A6
ALK	EHD3	KCNG4	OR6T1	SLC2A7
ALOX5AP	EHD4	KCNH1	OR6V1	SLC2A8
ALPI	ELANE	KCNH2	OR6X1	SLC2A9
ALPL	ELFN2	KCNH3	OR6Y1	SLC30A1

ALPP	ELMO1	KCNH4	OR7A5	SLC30A5
ALPPL2	ELN	KCNH8	OR7C1	SLC30A8
AMBP	ELOVL4	KCNJ1	OR7C2	SLC31A1
AMER1	ELOVL6	KCNJ10	OR7D2	SLC31A2
AMICA1	ELTD1	KCNJ11	OR7D4	SLC34A1
AMIGO1	EMB	KCNJ12	OR7G1	SLC34A2
AMIGO2	EMP1	KCNJ13	OR7G2	SLC34A3
ANK2	EMP2	KCNJ14	OR7G3	SLC35B1
ANK3	EMP3	KCNJ15	OR8A1	SLC36A1
ANKH	EMR1	KCNJ16	OR8B2	SLC36A2
ANKRA2	EMR2	KCNJ2	OR8B3	SLC36A4
ANO1	EMR3	KCNJ3	OR8B4	SLC37A4
ANO10	ENAH	KCNJ4	OR8B8	SLC38A1
ANO2	ENG	KCNJ5	OR8D1	SLC38A2
ANO3	ENO1	KCNJ6	OR8D2	SLC38A3
ANO4	ENOX1	KCNJ8	OR8D4	SLC38A5
ANO5	ENOX2	KCNJ9	OR8G1	SLC38A6
ANO6	ENPEP	KCNK1	OR8G2	SLC38A7
ANO7	ENPP1	KCNK12	OR8G5	SLC38A9
ANO8	ENPP3	KCNK13	OR8H1	SLC39A1
ANO9	ENPP4	KCNK17	OR8H2	SLC39A10
ANPEP	ENTPD1	KCNK18	OR8H3	SLC39A14
ANTXR1	ENTPD2	KCNK2	OR8I2	SLC39A2

ANXA1	ENTPD3	KCNK3	OR8J1	SLC39A3
ANXA13	ENTPD7	KCNK4	OR8J3	SLC39A4
AOC3	EPB41	KCNK5	OR8K1	SLC39A5
AP2A1	EPB41L1	KCNK6	OR8K3	SLC39A6
AP2A2	EPB41L2	KCNK7	OR8K5	SLC39A7
AP2B1	EPB41L3	KCNK9	OR8S1	SLC39A8
AP2M1	EPB42	KCNMA1	OR8U1	SLC3A1
AP2S1	EPCAM	KCNMB1	OR8U8	SLC3A2
APBB1	EPHA1	KCNMB2	OR9A2	SLC40A1
APC	EPHA2	KCNMB3	OR9A4	SLC41A3
APCDD1	EPHA3	KCNMB4	OR9G1	SLC43A1
APLNR	EPHA4	KCNN1	OR9G4	SLC43A3
APLP1	EPHA5	KCNN2	OR9G9	SLC44A1
APLP2	EPHA7	KCNN3	OR9I1	SLC44A2
APMAP	EPHA8	KCNN4	OR9K2	SLC44A4
APOH	EPHB1	KCNQ1	OR9Q1	SLC44A5
APP	EPHB2	KCNQ2	OR9Q2	SLC45A2
APR3	EPHB3	KCNQ3	ORAI1	SLC45A4
AQP1	EPHB4	KCNQ4	OSBPL6	SLC46A1
AQP10	EPHB6	KCNQ5	OSCAR	SLC46A2
AQP2	EPM2A	KCNS1	OSCP1	SLC46A3
AQP3	EPN1	KCNS2	OSMR	SLC47A1
AQP4	EPOR	KCNS3	OSTM1	SLC47A2

AQP5	EPS15	KCNT1	OTOF	SLC4A1
AQP6	EPS8	KCNT2	OTOP1	SLC4A10
AQP7	ERAP1	KCNV1	OTOP3	SLC4A2
AQP8	ERBB2	KCNV2	OXER1	SLC4A3
AQP9	ERBB2IP	KDR	OXGR1	SLC4A4
ARAP1	ERBB3	KEL	OXTR	SLC4A5
ARAP3	ERBB4	KIAA0195	P2RX1	SLC4A7
AREG	EREG	KIAA0247	P2RX2	SLC4A8
ARF1	ERLIN1	KIAA0319L	P2RX3	SLC5A1
ARF6	ERMAP	KIAA0644	P2RX4	SLC5A2
ARHGEF1	ERMP1	KIAA1161	P2RX5	SLC5A3
ARHGEF12	ESAM	KIAA1324	P2RX6	SLC5A4
ARL4C	ESR1	KIAA1324L	P2RX7	SLC5A5
ARRB1	ESYT1	KIAA1467	P2RY1	SLC5A6
ARRB2	ETNK1	KIAA1919	P2RY10	SLC5A7
ARSE	EVI2A	KIDINS220	P2RY11	SLC5A8
ARSF	EVI2B	KIR2DL1	P2RY12	SLC6A1
ART3	EXOC7	KIR2DL2	P2RY14	SLC6A11
ARVCF	EZR	KIR2DL3	P2RY2	SLC6A12
ASAH2	F11R	KIR2DL5A	P2RY4	SLC6A13
ASAP2	F2R	KIR2DS1	P2RY6	SLC6A14
ASGR1	F2RL1	KIR2DS2	P2RY8	SLC6A15
ASGR2	F2RL2	KIR2DS3	PAG1	SLC6A16

ASIC1	F2RL3	KIR2DS4	PAK1	SLC6A18
ASPSCR1	F3	KIR2DS5	PALM	SLC6A19
ASTN1	F5	KIR3DL1	PALM2	SLC6A2
ATG9A	FAIM2	KIR3DL2	PALM3	SLC6A20
ATP10A	FAIM3	KIR3DL3	PANX1	SLC6A3
ATP10D	FAM171A1	KIRREL	PANX2	SLC6A4
ATP12A	FAM174A	KIRREL2	PANX3	SLC6A5
ATP13A4	FAM174B	KIRREL3	PAQR4	SLC6A6
ATP1A1	FAM38A	KISS1R	PAQR7	SLC6A7
ATP1A2	FAM38B	KIT	PAQR8	SLC6A8
ATP1A3	FAM57A	KITLG	PARD3	SLC6A9
ATP1A4	FAP	KL	PARM1	SLC7A1
ATP1B1	FAS	KLK7	PARVA	SLC7A10
ATP1B2	FASLG	KLRA1	PARVG	SLC7A11
ATP1B3	FAT1	KLRB1	PAWR	SLC7A13
ATP1B4	FAT2	KLRC1	PCDH1	SLC7A2
ATP2B1	FAT4	KLRC2	PCDH10	SLC7A3
ATP2B2	FBLN2	KLRC4	PCDH11X	SLC7A4
ATP2B3	FCAMR	KLRF1	PCDH11Y	SLC7A5
ATP2B4	FCAR	KLRG1	PCDH12	SLC7A6
ATP2C2	FCER1A	KLRK1	PCDH15	SLC7A7
ATP4A	FCER2	KREMEN1	PCDH17	SLC7A8
ATP4B	FCGR1A	KREMEN2	PCDH18	SLC7A9

ATP6AP1	FCGR2A	KRT1	PCDH19	SLC8A1
ATP6AP2	FCGR2B	L1CAM	PCDH20	SLC8A2
ATP6V0A1	FCGR2C	LAG3	PCDH7	SLC8A3
ATP6V0A2	FCGR3A	LAIR1	PCDH8	SLC9A1
ATP6V0A4	FCGR3B	LAMA2	PCDH9	SLC9A2
ATP6V0C	FCGRT	LAMA4	PCDHA1	SLC9A3
ATP6V0E1	FCN1	LAMA5	PCDHA10	SLC9A3R2
ATP7A	FCRL1	LAMC1	PCDHA11	SLC9A4
ATP8A2	FCRL2	LAMC2	PCDHA13	SLC9A5
ATP8B2	FCRL3	LAMP1	PCDHA2	SLC9A8
ATRN	FCRL4	LAMP2	PCDHA3	SLC9A9
ATRNL1	FCRL5	LAMP3	PCDHA4	SLCO1A2
AVPR1A	FCRL6	LAPTM4A	PCDHA5	SLCO1B1
AVPR1B	FES	LASP1	PCDHA6	SLCO1B3
AVPR2	FEZ1	LAT	PCDHA7	SLCO1C1
AXIN1	FFAR1	LAX1	PCDHA8	SLCO2A1
AXL	FFAR2	LAYN	PCDHA9	SLCO2B1
BACE1	FFAR3	LCK	PCDHAC1	SLCO3A1
BACE2	FGD1	LCT	PCDHAC2	SLCO4A1
BAI1	FGFR1	LCTL	PCDHB1	SLCO4C1
BAI2	FGFR2	LDLR	PCDHB10	SLIT2
BAI3	FGFR3	LEPR	PCDHB11	SLIT3
BAMBI	FGFR4	LEPROTL1	PCDHB12	SLITRK1

BASP1	FGFRL1	LETM1	PCDHB13	SLITRK2
BCAM	FGR	LGR4	PCDHB14	SLITRK3
BCAN	FHIT	LGR5	PCDHB15	SLITRK4
BCAR1	FKRP	LGR6	PCDHB16	SLITRK5
BCHE	FLNA	LHCGR	PCDHB2	SLITRK6
BDKRB1	FLNB	LHFP	PCDHB3	SLMAP
BDKRB2	FLOT1	LIFR	PCDHB4	SMAGP
BEST1	FLOT2	LILRA1	PCDHB5	SMAP1
BEST2	FLRT1	LILRA5	PCDHB6	SMO
BEST3	FLRT2	LILRB1	PCDHB7	SMPD1
BEST4	FLRT3	LILRB2	PCDHB8	SMPD2
BFSP1	FLT1	LILRB3	PCDHB9	SMPD4
BFSP2	FLT3	LILRB4	PCDHGA10	SNAP23
BLNK	FLT3LG	LILRB5	PCDHGA11	SNTA1
BMPR1A	FLT4	LIM2	PCDHGA12	SNX5
BMPR1B	FLVCR1	LIMS2	PCDHGA5	SNX9
BMPR2	FLVCR2	LIN7A	PCDHGA6	SOAT2
BMX	FMN1	LIN7B	PCDHGA7	SOCS7
BOC	FMN2	LINGO1	PCDHGA8	SORCS1
BRAF	FMNL1	LMAN2L	PCDHGA9	SORCS2
BRI3BP	FMNL3	LMBR1	PCDHGB1	SORCS3
BRS3	FNDC4	LMBR1L	PCDHGB4	SORL1
BSCL2	FNDC5	LMBRD1	PCDHGB5	SORT1

BSG	FOLH1	LMBRD2	PCDHGB7	SPAM1
BSND	FOLR1	LMF2	PCDHGC3	SPHK1
BST1	FOLR2	LMLN	PCDHGC4	SPINT2
BST2	FPR1	LMTK2	PCDHGC5	SPN
BTC	FPR2	LNPEP	PDCD1	SPPL2A
BTK	FPR3	LPAR1	PDCD1LG2	SPPL2B
BTLA	FRAS1	LPAR2	PDE3A	SPPL3
BTN1A1	FREM2	LPAR3	PDE4A	SRC
BTN2A1	FRMD6	LPAR4	PDE4D	SRD5A3
BTN3A3	FSCN1	LPAR5	PDE6A	SSFA2
BVES	FSHR	LPAR6	PDGFC	SSPN
C10ORF54	FURIN	LPHN1	PDGFRA	SSR1
C10ORF72	FXYD1	LPHN2	PDGFRB	SSR2
C11ORF24	FXYD2	LPHN3	PDLIM5	SSTR1
C14ORF37	FXYD3	LPP	PDPK1	SSTR2
C16ORF54	FXYD4	LRBA	PDPN	SSTR3
C18ORF19	FXYD5	LRFN1	PDZK1IP1	SSTR4
C19ORF59	FYN	LRFN2	PEAR1	SSTR5
C19ORF6	FZD1	LRFN3	PEBP1	ST14
C1ORF27	FZD10	LRIG1	PECAM1	ST3GAL5
C1ORF85	FZD2	LRIG2	PERP	ST8SIA4
C1QBP	FZD3	LRIG3	PGAP1	STAB1
C3AR1	FZD4	LRP1	PGAP3	STAB2

C5AR1	FZD5	LRP10	PGRMC2	STARD4
C6ORF25	FZD6	LRP11	PHLDB2	STEAP1
C7ORF42	FZD7	LRP12	PI4KA	STEAP2
C9ORF5	FZD8	LRP1B	PICK1	STEAP4
C9ORF91	FZD9	LRP2	PIGG	STIM1
CA12	G3BP1	LRP3	PIGN	STOM
CA14	G6B	LRP4	PIGR	STOML1
CA4	GAB2	LRP5	PIK3CG	STOML2
CA9	GABBR1	LRP6	PILRA	STRA6
CABP1	GABBR2	LRP8	PILRB	STS
CABP7	GABRA1	LRPAP1	PIM1	STT3A
CACHD1	GABRA2	LRRC15	PIP5K1A	STT3B
CACNA1B	GABRA3	LRRC25	PKD1	STX1A
CACNA1C	GABRA4	LRRC26	PKD1L1	STX2
CACNA1D	GABRA5	LRRC32	PKD1L2	STX3
CACNA1E	GABRA6	LRRC4	PKD2	STX4
CACNA1I	GABRB1	LRRC4C	PKD2L1	STX6
CACNA2D1	GABRB3	LRRC7	PKD2L2	STX8
CACNA2D2	GABRD	LRRC8A	PKDREJ	SUCNR1
CACNG1	GABRE	LRRC8B	PKHD1	SULF1
CACNG2	GABRG1	LRRC8C	PKHD1L1	SUSD1
CACNG3	GABRG2	LRRC8D	PKP4	SUSD2
CACNG4	GABRP	LRRN1	PLA2R1	SUSD3

CACNG5	GABRQ	LRRN4CL	PLAU	SV2A
CACNG6	GABRR1	LSAMP	PLAUR	SVIL
CACNG7	GABRR2	LSP1	PLB1	SYK
CACNG8	GALR1	LSR	PLCB1	SYMPK
CADM1	GALR2	LTB	PLCD3	SYNC
CADM2	GALR3	LTB4R	PLCD4	SYNE2
CADM3	GAPDH	LTB4R2	PLCE1	SYNGR2
CADM4	GAS1	LTBR	PLCH2	SYNGR3
CALCR	GCA	LTC4S	PLD1	SYNGR4
CALCRL	GCGR	LTF	PLD3	SYNJ2
CALD1	GDAP1	LTK	PLD4	SYNPO
CALR	GDE1	LY6D	PLEK2	SYPL1
CALY	GDPD2	LY6E	PLEKHA4	SYT2
CANT1	GDPD5	LY6H	PLIN4	SYT3
CANX	GEM	LY6K	PLLP	SYT5
CAP2	GFRA1	LY75	PLP1	SYT8
CAPNS1	GFRA2	LY9	PLPP1	SYT9
CASD1	GFRA3	LYN	PLSCR1	SYTL1
CASR	GFRA4	LYPD3	PLVAP	SYTL2
CATSPER1	GGCX	LYVE1	PLXDC1	SYVN1
CATSPER4	GGT1	LZTS1	PLXNA1	TAAR1
CAV1	GGT5	M6PR	PLXNA2	TAAR2
CAV2	GGT7	MACF1	PLXNA3	TAAR5

CAV3	GHR	MADD	PLXNB1	TAAR6
CBL	GHRHR	MAEA	PLXNB2	TAAR8
CBLN1	GHSR	MAG	PLXNB3	TAAR9
CCBP2	GIPR	MAGI1	PLXNC1	TACR1
CCDC107	GJA1	MAGI3	PLXND1	TACR2
CCDC109B	GJA3	MAGT1	PMEPA1	TACR3
CCKAR	GJA4	MAL	PMP22	TACSTD2
CCKBR	GJA5	MAL2	PNKD	TAS1R1
CCNY	GJA8	MALL	PODXL	TAS1R2
CCR1	GJA9	MAN2A2	PODXL2	TAS2R1
CCR10	GJB1	MAPT	PON1	TAS2R16
CCR2	GJB2	MARCH2	PON2	TAS2R38
CCR3	GJB3	MARCH3	PPAP2A	TAS2R60
CCR4	GJB4	MARCKS	PPAP2B	TAS2R7
CCR5	GJB5	MARCO	PPAP2C	TAS2R8
CCR6	GJB6	MARK2	PPFIBP1	TAS2R9
CCR7	GJB7	MAS1	PPIB	TBXA2R
CCR8	GJC1	MAS1L	PPL	TCAF2
CCR9	GJC2	MAST1	PPP1R16B	TCHP
CCRL2	GJC3	MAST2	PPP3R1	TCIRG1
CD101	GJD2	MBOAT2	PPYR1	TEC
CD109	GJD3	MBOAT7	PREX1	TEK
CD14	GLG1	MBP	PRF1	TES

CD151	GLIPR1	MBTPS2	PRG2	TF
CD160	GLP1R	MC1R	PRIMA1	TFPI
CD163	GLP2R	MC2R	PRKCA	TFPI2
CD163L1	GLRA1	MC3R	PRKD1	TFR2
CD164	GLRA2	MC4R	PRLHR	TFRC
CD177	GLRA3	MC5R	PRLR	TGFA
CD180	GLRB	MCAM	PRND	TGFBR1
CD19	GMIP	MCF2L	PRNP	TGFBR2
CD1A	GML	MCHR1	PROCR	TGFBR3
CD1B	GNA11	MCHR2	PROKR1	TGM2
CD1C	GNAI1	MCOLN1	PROKR2	THBD
CD1D	GNAI2	MCOLN3	PROM1	THBS1
CD2	GNAI3	MDGA1	PROM2	THBS2
CD200	GNAQ	MDGA2	PRPH2	THEM4
CD200R1	GNAS	MEGF8	PRR7	THSD1
CD207	GNB4	MEP1A	PRRG1	THY1
CD209	GNRHR	MEP1B	PRRG2	TIAM1
CD22	GOPC	MERTK	PRRG3	TIE1
CD226	GP1BA	MET	PRRT1	TIGIT
CD24	GP1BB	MFAP3	PRRT2	TIMD4
CD244	GP2	MFAP3L	PRRT3	TIMP1
CD247	GP5	MFGE8	PRSS8	TLCD1
CD248	GP9	MFI2	PRTG	TLN1

CD27	GPA33	MFRP	PRX	TLN2
CD274	GPBAR1	MFSD4	PSAP	TLR1
CD276	GPC1	MFSD8	PSCA	TLR10
CD28	GPC2	MGAM	PSD4	TLR2
CD300A	GPC3	MGAT5B	PSEN1	TLR3
CD300C	GPC4	MIB1	PTAFR	TLR4
CD300E	GPC5	MICA	PTCH1	TLR5
CD300LF	GPC6	MICB	PTCH2	TLR6
CD302	GPER	MIP	PTGDR	TLR7
CD320	GPI	MKLN1	PTGER1	TLR8
CD33	GPM6A	MLC1	PTGER2	TM2D1
CD34	GPM6B	MLLT4	PTGER3	TM2D2
CD36	GPNMB	MLNR	PTGER4	TM2D3
CD37	GPR1	MMD	PTGFR	TM4SF1
CD38	GPR101	MME	PTGFRN	TM4SF20
CD3D	GPR107	MMP14	PTGIR	TM4SF4
CD3E	GPR109A	MMP15	PTGS2	TM4SF5
CD3G	GPR109B	MMP16	PTH1R	TM7SF3
CD4	GPR110	MMP17	PTH2R	TM7SF4
CD40	GPR112	MMP24	PTK2B	TM9SF1
CD40LG	GPR113	MMP25	PTK7	TM9SF3
CD44	GPR116	MMRN1	PTOV1	TMC1
CD46	GPR119	MOG	PTPN13	TMC2

CD47	GPR12	MPDZ	PTPN22	TMC5
CD48	GPR124	MPHOSPH9	PTPN4	TMC7
CD5	GPR125	MPL	PTPRA	TMCO3
CD52	GPR126	MPP5	PTPRB	TMED1
CD53	GPR128	MPZ	PTPRC	TMED10
CD55	GPR132	MPZL1	PTPRCAP	TMED7
CD58	GPR135	MPZL2	PTPRD	TMEFF1
CD59	GPR137B	MR1	PTPRE	TMEFF2
CD6	GPR143	MRAP	PTPRF	TMEM101
CD63	GPR146	MRAP2	PTPRG	TMEM104
CD68	GPR15	MRC1	PTPRH	TMEM106B
CD69	GPR150	MRC2	PTPRJ	TMEM106C
CD7	GPR155	MRGPRD	PTPRK	TMEM115
CD70	GPR156	MRGPRF	PTPRM	TMEM123
CD72	GPR160	MRGPRX1	PTPRN	TMEM132A
CD74	GPR17	MRGPRX2	PTPRR	TMEM132B
CD79A	GPR172A	MRGPRX3	PTPRS	TMEM132E
CD79B	GPR172B	MRGPRX4	PTPRU	TMEM139
CD80	GPR173	MS4A1	PTPRZ1	TMEM144
CD81	GPR174	MS4A2	PTRF	TMEM156
CD82	GPR176	MSLN	PVR	TMEM160
CD83	GPR18	MSN	PVRL1	TMEM161A
CD84	GPR182	MSR1	PVRL2	TMEM176B

CD86	GPR19	MST1R	PVRL3	TMEM179B
CD8A	GPR20	MTM1	PVRL4	TMEM183A
CD8B	GPR21	MTMR1	PXK	TMEM194A
CD9	GPR22	MTNR1A	QRFPR	TMEM2
CD93	GPR25	MTNR1B	QSOX1	TMEM200B
CD96	GPR26	MTUS1	QSOX2	TMEM206
CD97	GPR27	MUC1	RAB13	TMEM222
CDC42	GPR3	MUC13	RAB18	TMEM25
CDC42EP1	GPR31	MUC15	RAB23	TMEM26
CDC42EP4	GPR32	MUC16	RAB3A	TMEM27
CDC42SE1	GPR34	MUC17	RAB3B	TMEM30A
CDCP1	GPR35	MUC20	RAB3C	TMEM30B
CDH1	GPR37	MUC21	RAB3D	TMEM47
CDH10	GPR39	MUC4	RAB5A	TMEM50A
CDH11	GPR4	MUSK	RAB5B	TMEM55B
CDH12	GPR42	MXRA8	RAB5C	TMEM61
CDH13	GPR44	MYH9	RAB8A	TMEM63A
CDH15	GPR45	MYO10	RAB8B	TMEM63B
CDH16	GPR50	MYO1B	RAC1	TMEM63C
CDH17	GPR52	MYO1C	RAC3	TMEM67
CDH18	GPR55	MYO1G	RAET1G	TMEM87A
CDH19	GPR56	MYOF	RAF1	TMEM87B
CDH2	GPR6	NAALADL2	RALA	TMEM8A

CDH20	GPR61	NAGPA	RAMP1	TMIGD2
CDH22	GPR62	NALCN	RAMP2	TMPRSS2
CDH23	GPR63	NCAM1	RAMP3	TMPRSS4
CDH24	GPR64	NCAM2	RAP1A	TMPRSS5
CDH26	GPR65	NCK1	RAP1B	TMPRSS6
CDH3	GPR68	NCLN	RAPH1	TMPRSS9
CDH4	GPR75	NCR1	RAPSN	TMUB1
CDH5	GPR78	NCR2	RASA3	TNC
CDH6	GPR82	NCR3	RASD1	TNF
CDH8	GPR83	NCS1	RASD2	TNFRSF10B
CDH9	GPR84	NCSTN	RDX	TNFRSF10C
CDHR1	GPR85	NEDD4	RECK	TNFRSF10D
CDHR2	GPR87	NEDD9	RELL1	TNFRSF11A
CDHR5	GPR88	NEGR1	REM2	TNFRSF12A
CDIPT	GPR98	NEO1	RET	TNFRSF13B
CDK5	GPRC5A	NETO1	RFNG	TNFRSF13C
CDON	GPRC5B	NEU1	RFTN1	TNFRSF14
CDSN	GPRC5C	NF2	RGS1	TNFRSF17
CEACAM1	GPRC5D	NFAM1	RGS10	TNFRSF18
CEACAM3	GRAMD1A	NFASC	RGS14	TNFRSF19
CEACAM4	GRASP	NGFR	RGS2	TNFRSF1A
CEACAM5	GRB7	NID2	RGS7	TNFRSF1B
CEACAM6	GRIA1	NINJ1	RGS8	TNFRSF21

CEACAM8	GRIA2	NINJ2	RHAG	TNFRSF25
CECR6	GRIA3	NIPA1	RHBDF2	TNFRSF4
CELSR1	GRIA4	NIPA2	RHBDL1	TNFRSF8
CELSR2	GRID1	NISCH	RHBG	TNFRSF9
CELSR3	GRID2	NKG7	RHCE	TNFSF11
CFC1	GRIK1	NKTR	RHCG	TNFSF12
CFTR	GRIK2	NLGN1	RHD	TNFSF13
CGN	GRIK3	NLGN2	RHO	TNFSF13B
CHIC1	GRIK4	NLGN3	RHOA	TNFSF14
CHIC2	GRIK5	NLGN4X	RHOB	TNFSF15
CHL1	GRIN1	NMBR	RHOC	TNFSF8
CHODL	GRIN2A	NMUR1	RHOD	TNFSF9
CHRM2	GRIN2B	NMUR2	RHOF	TNMD
CHRM3	GRIN2C	NOS1	RHOG	TOLLIP
CHRM4	GRIN2D	NOS3	RHOJ	TOR1AIP2
CHRM5	GRIN3A	NOTCH1	RHOQ	TPBG
CHRNA1	GRK5	NOTCH2	RIC3	TPCN1
CHRNA2	GRM1	NOX1	RIC8A	TPO
CHRNA3	GRM2	NOX3	RIMS2	TPSG1
CHRNA4	GRM3	NOX4	RIPK1	TPTE
CHRNA5	GRM4	NOXO1	RIT1	TRAT1
CHRNA7	GRM5	NPBWR1	RIT2	TRDN
CHRNB1	GRM6	NPBWR2	RND1	TREM1

CHRNB2	GRM7	NPC1	RNF130	TREM2
CHRNB3	GRM8	NPFFR2	RNF139	TREML1
CHRNB4	GRPR	NPHS1	RNF144B	TRHDE
CHRND	GSG1L	NPHS2	RNF145	TRHR
CHRNE	GUCY2C	NPR1	RNF149	TRIP6
CHRNG	GUCY2D	NPR2	RNF43	TRPA1
CHST7	GUCY2F	NPR3	RNF5	TRPC1
CIB1	GYPA	NPSR1	ROBO1	TRPC3
CKAP4	GYPB	NPTN	ROBO4	TRPC4
CLCA1	GYPC	NPTX1	ROM1	TRPC5
CLCA2	GYPE	NPTX2	ROR1	TRPC6
CLCA4	HAS1	NPTXR	ROR2	TRPC7
CLCN1	HAS2	NPY1R	ROS1	TRPM1
CLCN2	HAS3	NPY2R	RP2	TRPM2
CLCN4	HAVCR1	NPY5R	RPN2	TRPM3
CLCN5	HAVCR2	NRCAM	RRAD	TRPM4
CLCN6	HBEGF	NRG1	RRAS	TRPM5
CLCN7	HCK	NRG2	RRAS2	TRPM6
CLCNKA	HCLS1	NRG3	RRH	TRPM7
CLCNKB	HCN1	NRG4	RTN4	TRPM8
CLDN1	HCN2	NRP1	RTN4R	TRPV1
CLDN11	HCN3	NRP2	RTP1	TRPV2
CLDN12	HCN4	NRXN1	RTP2	TRPV3

CLDN14	HCRTR1	NRXN2	RUSC1	TRPV4
CLDN15	HCRTR2	NRXN3	RXFP1	TRPV5
CLDN16	HCST	NT5E	RXFP2	TRPV6
CLDN17	HEG1	NTM	RXFP3	TSHR
CLDN18	HEPACAM	NTNG1	RXFP4	TSPAN1
CLDN19	HEPACAM2	NTNG2	RYK	TSPAN13
CLDN2	HEPH	NTRK1	RYR2	TSPAN15
CLDN20	HFE	NTRK2	RYR3	TSPAN2
CLDN23	HFE2	NTRK3	S100A10	TSPAN3
CLDN3	HHATL	NTSR1	S100P	TSPAN31
CLDN4	HHIP	NTSR2	S1PR1	TSPAN33
CLDN5	HHLA2	OCLN	S1PR2	TSPAN4
CLDN6	HLA-A	ODC1	S1PR3	TSPAN6
CLDN7	HLA-B	ODZ1	S1PR4	TSPAN7
CLDN8	HLA-C	ODZ3	S1PR5	TSPAN8
CLDN9	HLA-DMA	OLR1	SC4MOL	TSPAN9
CLDND1	HLA-DMB	OPCML	SCAMP1	TTC17
CLEC12A	HLA-DOA	OPN1LW	SCAMP2	TTC7B
CLEC14A	HLA-DPA1	OPN1MW	SCAMP3	TTYH1
CLEC17A	HLA-DPB1	OPN1SW	SCAMP4	TTYH2
CLEC1A	HLA-DQA1	OPN3	SCARA3	TTYH3
CLEC1B	HLA-DQA2	OPN4	SCARA5	TUB
CLEC2B	HLA-DQB1	OPRD1	SCARB1	TUBB

CLEC2D	HLA-DQB2	OPRM1	SCARB2	TULP3
CLEC4A	HLA-DRA	OR10A6	SCIN	TXNDC15
CLEC4F	HLA-DRB1	OR10G7	SCN10A	TYRO3
CLEC4G	HLA-DRB3	OR10J3	SCN11A	TYROBP
CLEC5A	HLA-DRB4	OR10V1	SCN1A	TYRP1
CLIC1	HLA-DRB5	OR11A1	SCN1B	UBE2B
CLIC2	HLA-E	OR13C4	SCN2A	UBIAD1
CLIC4	HLA-F	OR13D1	SCN2B	ULBP1
CLIC5	HLA-G	OR13H1	SCN3A	ULBP2
CLIC6	HM13	OR1A1	SCN4A	ULBP3
CLIP3	HMGB1	OR1A2	SCN4B	UMODL1
CLN3	HOMER3	OR1B1	SCN5A	UNC5A
CLN6	HRH1	OR1C1	SCN7A	UNC5B
CLPTM1	HRH2	OR1D2	SCN8A	UNC5C
CLPTM1L	HRH3	OR1D4	SCN9A	UPK1A
CLRN1	HRH4	OR1D5	SCNN1A	UPK1B
CLSTN1	HS3ST3B1	OR1E1	SCNN1B	UPK2
CLSTN2	HS6ST1	OR1E2	SCNN1D	UPK3B
CLSTN3	HSPB1	OR1F1	SCNN1G	USH2A
CLTB	HSPD1	OR1G1	SCRIB	UST
CLTC	HSPG2	OR1I1	SCTR	UTRN
CMKLR1	HTR1A	OR1J1	SDC1	UTS2R
CMTM2	HTR1B	OR1J2	SDC2	VAMP1

CMTM4	HTR1D	OR1J4	SDC3	VAMP2
CMTM6	HTR1E	OR1K1	SDC4	VAMP3
CMTM7	HTR1F	OR1L1	SDCBP	VAMP5
CMTM8	HTR2A	OR1L3	SDK1	VAMP8
CNGA1	HTR2B	OR1L4	SDK2	VANGL1
CNGA3	HTR2C	OR1L6	SDPR	VANGL2
CNGA4	HTR3A	OR1L8	SEL1L	VAPA
CNGB3	HTR3B	OR1M1	SELE	VAPB
CNIH2	HTR3C	OR1N1	SELL	VASN
CNIH3	HTR4	OR1N2	SELP	VASP
CNKSR2	HTR5A	OR1Q1	SELPLG	VCAM1
CNNM1	HTR6	OR1S1	SEMA3C	VDAC1
CNNM2	HTR7	OR1S2	SEMA4A	VEZT
CNNM3	HTRA1	OR2A1	SEMA4B	VIL1
CNNM4	HVCN1	OR2A2	SEMA4C	VIPR1
CNR1	HYAL2	OR2A4	SEMA4D	VIPR2
CNR2	ICAM1	OR2A5	SEMA4F	VN1R2
CNTFR	ICAM2	OR2A7	SEMA4G	VN1R4
CNTN1	ICAM3	OR2B2	SEMA5A	VN1R5
CNTN2	ICAM4	OR2B6	SEMA6A	VNN1
CNTN3	ICAM5	OR2C1	SEMA6B	VNN2
CNTN4	ICOS	OR2C3	SEMA6D	VNN3
CNTNAP1	ICOSLG	OR2D2	SEMA7A	VSIG10

CNTNAP2	IDE	OR2D3	SEPT3	VSIG2
CNTNAP3	IFIT5	OR2F1	SERINC1	VSIG4
COBL	IFITM1	OR2F2	SERINC3	VSTM1
COL17A1	IFITM3	OR2G2	SERPINH1	VTCN1
COL6A2	IFNAR1	OR2G3	SEZ6	WDR13
COLEC12	IFNAR2	OR2G6	SEZ6L2	WDR6
CORIN	IFNGR1	OR2H1	SFTPC	WWP1
CORO2B	IFNGR2	OR2H2	SGCA	XCR1
CP	IFRD1	OR2J2	SGCB	XG
CPD	IGDCC4	OR2J3	SGCD	XK
CPM	IGF1R	OR2K2	SGCE	XKR3
CPNE6	IGF2R	OR2L2	SGCZ	XKR8
CR1	IGHD	OR2L3	SGK1	XKRY
CR2	IGHM	OR2L8	SGMS2	XPR1
CRB1	IGSF1	OR2M2	SH2B2	YES1
CRB2	IGSF11	OR2M3	SHH	YIPF3
CRB3	IGSF3	OR2M4	SIDT1	YIPF4
CRELD1	IGSF6	OR2M5	SIDT2	ZACN
CRHR1	IGSF8	OR2M7	SIGIRR	ZAN
CRHR2	IL10RA	OR2T1	SIGLEC1	ZAP70
CRIM1	IL10RB	OR2T11	SIGLEC10	ZDHHC22
CRK	IL11RA	OR2T2	SIGLEC12	ZDHHC23
CRLF2	IL12RB1	OR2T3	SIGLEC15	ZDHHC5

CRTAM	IL12RB2	OR2T4	SIGLEC5	ZP1
CRTAP	IL13RA1	OR2T5	SIGLEC6	ZP2
CSF1	IL13RA2	OR2T6	SIGLEC7	ZP3
CSF1R	IL17RA	OR2T8	SIGLEC8	ZP4
CSF2RA	IL17RB	OR2V2	SIGLEC9	-
CSF2RB	IL17RC	OR2W1	SILV	-

7.2.2. Experimental surfaceome

Table 7.2. Surface protein-encoding genes represented in the experimental surfaceome.

A2M	CPD	HS3ST5	OR2G2	SLC14A1
A4GALT	CPE	HS6ST1	OR2H2	SLC14A2
AADAT	CPLX1	HSD11B1	OR2K2	SLC15A1
AAK1	CPLX2	HSD17B10	OR2T11	SLC15A2
AARS	CPM	HSD17B2	OR2T8	SLC15A3
ABCA1	CPNE6	HSH2D	OR2Y1	SLC15A4
ABCA12	CPPED1	HSP90AA1	OR3A1	SLC16A1
ABCA13	CPVL	HSP90AB1	OR3A2	SLC16A10
ABCA2	CPZ	HSP90B1	OR3A3	SLC16A2
ABCA4	CR1	HSPA13	OR4F6	SLC16A3
ABCA5	CR2	HSPA4	OR4N2	SLC16A4
ABCA6	CRB1	HSPA5	OR51B2	SLC16A5
ABCA7	CRB2	HSPA8	OR51B4	SLC16A6
ABCA8	CRB3	HSPA9	OR51B6	SLC16A7
ABCB1	CREB5	HSPB1	OR51E1	SLC16A8
ABCB11	CREG1	HSPB8	OR51F2	SLC17A1
ABCB4	CRELD1	HSPD1	OR51I1	SLC17A2
ABCB5	CRELD2	HSPG2	OR51I2	SLC17A3
ABCB6	CRHBP	HTR1A	OR51Q1	SLC17A4
ABCC1	CRHR1	HTR1B	OR51S1	SLC17A5
ABCC10	CRHR2	HTR1D	OR52A1	SLC17A7

ABCC12	CRIM1	HTR1E	OR52A4	SLC18A1
ABCC2	CRIP2	HTR1F	OR52A5	SLC18A2
ABCC3	CRK	HTR2A	OR52E4	SLC18A3
ABCC4	CRLF1	HTR2B	OR52E8	SLC19A1
ABCC5	CRLF2	HTR2C	OR52H1	SLC19A2
ABCC6	CRLF3	HTR3A	OR52I1	SLC19A3
ABCC8	CRNN	HTR3B	OR52K1	SLC1A1
ABCC9	CROCC	HTR3C	OR52N4	SLC1A2
ABCG1	CRTAM	HTR4	OR52N5	SLC1A3
ABCG2	CRTAP	HTR5A	OR56A1	SLC1A4
ABCG5	CRTC1	HTR6	OR5AN1	SLC1A5
ABCG8	CRYAB	HTR7	OR5AR1	SLC1A6
ABHD3	CRYL1	HTRA1	OR5B12	SLC1A7
ABI1	CSF1	HTT	OR5B2	SLC20A1
ABI3BP	CSF1R	HUNK	OR5D13	SLC20A2
ABL1	CSF2RA	HYAL2	OR5H6	SLC22A11
ABLIM3	CSF2RB	HYDIN	OR5I1	SLC22A12
ACAA1	CSF3R	HYI	OR5P2	SLC22A13
ACAM	CSGALNACT2	HYLS1	OR5P3	SLC22A14
ACAP2	CSK	HYOU1	OR6A2	SLC22A16
ACCN1	CSNK1D	IBTK	OR6C1	SLC22A18
ACCN2	CSNK1G2	ICAM1	OR6F1	SLC22A2
ACCN3	CSNK2A1	ICAM2	OR6V1	SLC22A23

ACE	CSNK2A2	ICAM3	OR7C2	SLC22A3
ACE2	CSPG4	ICAM4	OR7G3	SLC22A4
ACHE	CSPG5	ICAM5	OR8A1	SLC22A5
ACKR3	CST6	ICOS	OR8D2	SLC22A6
ACLY	CTAG2	ICOSLG	OR8S1	SLC22A7
ACOT9	CTLA4	IDE	OR9G1	SLC22A8
ACP2	CTNNA1	IDH2	ORAI1	SLC22A9
ACSL6	CTNNA3	IFFO2	ORC1	SLC23A1
ACTB	CTNNB1	IFI30	ORM1	SLC23A2
ACTN1	CTNNBIP1	IFIT5	OSBP2	SLC24A1
ACTN4	CTNND1	IFITM1	OSBPL10	SLC24A2
ACTR6	CTNND2	IFITM3	OSBPL3	SLC24A4
ACVR1	CTNS	IFNAR1	OSBPL6	SLC24A6
ACVR1B	CTSA	IFNAR2	OSBPL7	SLC26A1
ACVR1C	CTSB	IFNGR1	OSCAR	SLC26A11
ACVR2A	CTSC	IFNGR2	OSCP1	SLC26A2
ACVR2B	CTSD	IFRD1	OSGEP	SLC26A3
ACVRL1	CTSE	IGDCC3	OSMR	SLC26A4
ADA	CTSF	IGDCC4	OSR2	SLC26A5
ADAM10	CTSG	IGF1R	OSTM1	SLC26A6
ADAM11	CTSK	IGF2R	OTOA	SLC26A7
ADAM12	CTSL1	IGFBP3	OTOP1	SLC26A8
ADAM15	CTSL2	IGFBP7	OTOP3	SLC27A1

ADAM17	CTSS	IGFL2	OXGR1	SLC27A4
ADAM18	CTSV	IGHA1	OXT	SLC27A6
ADAM19	CTTN	IGHD	OXTR	SLC28A1
ADAM2	CTXN1	IGHG1	P2RX1	SLC28A2
ADAM21	CUBN	IGHG2	P2RX2	SLC29A1
ADAM22	CUL5	IGHG4	P2RX3	SLC29A2
ADAM23	CUTA	IGHM	P2RX4	SLC29A4
ADAM28	CUZD1	IGLON5	P2RX5	SLC2A1
ADAM29	CX3CL1	IGSF1	P2RX6	SLC2A10
ADAM30	CX3CR1	IGSF11	P2RX7	SLC2A11
ADAM33	CXADR	IGSF3	P2RY1	SLC2A12
ADAM7	CXCL16	IGSF6	P2RY10	SLC2A13
ADAM8	CXCR1	IGSF8	P2RY11	SLC2A14
ADAM9	CXCR2	IKBKB	P2RY12	SLC2A2
ADAMTS15	CXCR3	IKBKG	P2RY14	SLC2A3
ADAMTS2	CXCR4	IKIP	P2RY2	SLC2A4
ADAMTSL4	CXCR5	IKZF5	P2RY4	SLC2A5
ADAP1	CXCR6	IL10RA	P2RY6	SLC2A6
ADAP2	CXCR7	IL10RB	P2RY8	SLC2A7
ADCK5	CYB561D2	IL11RA	P4HA1	SLC2A8
ADCY1	CYB5R3	IL12RB1	P4HA2	SLC2A9
ADCY2	CYBA	IL12RB2	PA2G4	SLC30A1
ADCY3	CYBB	IL13	PACSIN2	SLC30A4

ADCY4	CYBRD1	IL13RA1	PAFAH1B2	SLC30A5
ADCY5	CYFIP2	IL13RA2	PAFAH2	SLC30A8
ADCY6	CYP20A1	IL16	PAG1	SLC31A1
ADCY7	CYR61	IL17D	PAIP1	SLC31A2
ADCY8	CYSLTR1	IL17RA	PAK1	SLC34A1
ADCY9	CYSLTR2	IL17RB	PAK4	SLC34A2
ADCYAP1R1	CYTH1	IL17RC	PAK6	SLC34A3
ADD1	CYTH2	IL17RD	PALLD	SLC35B1
ADD2	CYTH3	IL18R1	PALM	SLC36A1
ADD3	CYTIP	IL18RAP	PALM2	SLC36A2
ADGRG2	DAAM1	IL1A	PALM2-AKAP2	SLC36A4
ADGRL2	DAB2	IL1B	PALM3	SLC37A4
ADH1A	DAG1	IL1R1	PALMD	SLC38A1
ADH1B	DAGLB	IL1R2	PANX1	SLC38A2
ADH1C	DAPK1	IL1RAP	PANX2	SLC38A3
ADH7	DAPP1	IL1RAPL1	PANX3	SLC38A4
ADI1	DARC	IL1RAPL2	PAPPA	SLC38A5
ADIPOR1	DBH	IL1RL1	PAQR4	SLC38A6
ADIPOR2	DBN1	IL1RL2	PAQR7	SLC38A7
ADORA1	DCAF16	IL20RA	PAQR8	SLC38A9
ADORA2A	DCBLD1	IL21R	PARD3	SLC39A1
ADORA2B	DCBLD2	IL22RA1	PARD3B	SLC39A10

ADORA3	DCC	IL26	PARD6A	SLC39A14
ADRA1A	DCHS1	IL27RA	PARD6B	SLC39A2
ADRA1B	DCHS2	IL28RA	PARD6G	SLC39A3
ADRA1D	DCLK1	IL2RA	PARK2	SLC39A4
ADRA2A	DCN	IL2RB	PARM1	SLC39A5
ADRA2B	DCT	IL2RG	PARP6	SLC39A6
ADRA2C	DCTN6	IL31RA	PARVA	SLC39A7
ADRB1	DCUN1D2	IL3RA	PARVG	SLC39A8
ADRB2	DDAH1	IL4R	PATE1	SLC3A1
ADRB3	DDIAS	IL5RA	PAWR	SLC3A2
ADRBK1	DDR1	IL6R	PBXIP1	SLC40A1
ADRBK2	DDR2	IL6ST	PCBD2	SLC41A3
ADRM1	DDX60L	IL7R	PCBP2	SLC43A1
ADSS	DEF6	IL9R	PCBP3	SLC43A3
AEBP1	DFFA	ILDR1	PCDH1	SLC44A1
AER61	DFNB31	IMPAD1	PCDH10	SLC44A2
AFAP1	DGCR2	INADL	PCDH11X	SLC44A4
AFAP1L2	DGKA	INHBE	PCDH11Y	SLC44A5
AGER	DGKB	INO80B	PCDH12	SLC45A2
AGPS	DGKD	INPP5A	PCDH15	SLC45A4
AGRN	DGKG	INPP5B	PCDH17	SLC46A1
AGTR1	DGKQ	INPP5J	PCDH18	SLC46A2
AGTR2	DGKZ	INPP5K	PCDH19	SLC46A3

AHCYL1	DHCR24	INPPL1	PCDH20	SLC47A1
AHNAK	DHPS	INSR	PCDH7	SLC47A2
AHNAK2	DHRS7B	INSRR	PCDH8	SLC4A1
AHSG	DHX36	INTS12	PCDH9	SLC4A10
AJAP1	DIAPH1	INVS	PCDHA1	SLC4A2
AK1	DIAPH3	IPCEF1	PCDHA10	SLC4A3
AKAP10	DIO3	IQGAP1	PCDHA11	SLC4A4
AKAP11	DIRAS1	IQGAP2	PCDHA13	SLC4A5
AKAP12	DIRAS2	IRAK4	PCDHA2	SLC4A7
AKAP13	DIRAS3	IRS4	PCDHA3	SLC4A8
AKAP2	DIRC2	ISLR	PCDHA4	SLC5A1
AKAP5	DIS3L	ITCH	PCDHA5	SLC5A2
AKAP7	DISP1	ITFG1	PCDHA6	SLC5A3
AKR1A1	DISP2	ITFG3	PCDHA7	SLC5A4
AKR1B10	DKK3	ITGA1	PCDHA8	SLC5A5
AKR1B15	DLEU1	ITGA10	PCDHA9	SLC5A6
AKT1	DLG1	ITGA11	PCDHAC1	SLC5A7
AKT2	DLG2	ITGA2	PCDHAC2	SLC5A8
AKT3	DLG3	ITGA2B	PCDHB1	SLC6A1
AKTIP	DLG4	ITGA3	PCDHB10	SLC6A11
ALCAM	DLG5	ITGA4	PCDHB11	SLC6A12
ALDH3A1	DLGAP2	ITGA5	PCDHB12	SLC6A13
ALG10	DLGAP4	ITGA6	PCDHB13	SLC6A14

ALG10B	DLK1	ITGA7	PCDHB14	SLC6A15
ALK	DLL1	ITGA8	PCDHB15	SLC6A16
ALOX15	DLL3	ITGA9	PCDHB16	SLC6A18
ALOX15B	DLL4	ITGAD	PCDHB2	SLC6A19
ALOX5AP	DLST	ITGAE	PCDHB3	SLC6A2
ALOXE3	DMBT1	ITGAL	PCDHB4	SLC6A20
ALPI	DMD	ITGAM	PCDHB5	SLC6A3
ALPL	DMTF1	ITGAV	PCDHB6	SLC6A4
ALPP	DMWD	ITGAX	PCDHB7	SLC6A5
ALPPL2	DMXL2	ITGB1	PCDHB8	SLC6A6
ALS2CR12	DNAAF1	ITGB1BP2	PCDHB9	SLC6A7
AMACR	DNAJA4	ITGB2	PCDHGA10	SLC6A8
AMBP	DNAJB4	ITGB3	PCDHGA11	SLC6A9
AMDHD1	DNAJC5	ITGB4	PCDHGA12	SLC7A1
AMER1	DNAJC6	ITGB5	PCDHGA5	SLC7A10
AMHR2	DNAJC9	ITGB6	PCDHGA6	SLC7A11
AMICA1	DNASE1L1	ITGB7	PCDHGA7	SLC7A13
AMIGO1	DNER	ITGB8	PCDHGA8	SLC7A2
AMIGO2	DNM1	ITIH2	PCDHGA9	SLC7A3
AMOT	DNM2	ITK	PCDHGB1	SLC7A4
AMOTL1	DOC2B	ITM2A	PCDHGB4	SLC7A5
AMPH	DOCK1	ITM2B	PCDHGB5	SLC7A6
ANGPT1	DOCK2	ITPKB	PCDHGB7	SLC7A7

ANGPTL1	DOCK4	ITPR1	PCDHGC3	SLC7A8
ANGPTL2	DOCK8	ITPR2	PCDHGC4	SLC7A9
ANGPTL4	DOCK9	ITPR3	PCDHGC5	SLC8A1
ANK1	DOK1	ITPRIP	PCLO	SLC8A2
ANK2	DOK2	ITSN1	PCMTD1	SLC8A3
ANK3	DOK6	ITSN2	PCYOX1	SLC9A1
ANKEF1	DPEP1	IVL	PDAP1	SLC9A2
ANKH	DPM1	IYD	PDCD1	SLC9A3
ANKIB1	DPM3	JAG1	PDCD1LG2	SLC9A3R1
ANKLE2	DPP10	JAG2	PDCD7	SLC9A3R2
ANKRA2	DPP4	JAK2	PDE3A	SLC9A4
ANKRD13A	DPP6	JAM2	PDE4A	SLC9A5
ANKRD35	DPP7	JAM3	PDE4B	SLC9A6
ANKRD37	DPY19L4	JKAMP	PDE4D	SLC9A7
ANKRD42	DPYSL2	JMJD4	PDE6A	SLC9A8
ANKRD44	DQX1	JMJD6	PDE6B	SLC9A9
ANKS1B	DRD1	JPH1	PDGFC	SLCO1A2
ANO2	DRD2	JPH2	PDGFRA	SLCO1B1
ANO5	DRD3	JPH3	PDGFRB	SLCO1B3
ANO6	DRD5	JPH4	PDGFRL	SLCO1C1
ANO7	DRP2	JTB	PDHX	SLCO2A1
ANO9	DSC1	JUB	PDIA3	SLCO2B1
ANP32A	DSC2	JUP	PDIA4	SLCO3A1

ANPEP	DSC3	KAL1	PDLIM1	SLCO4A1
ANTXR1	DSCAM	KANK1	PDLIM5	SLCO4C1
ANXA1	DSCAML1	KANK3	PDPK1	SLIT2
ANXA13	DSEL	KANSL2	PDPN	SLIT3
ANXA2	DSG1	KARS	PDZD2	SLITRK1
ANXA5	DSG2	KATNB1	PDZD3	SLITRK2
ANXA6	DSG3	KBTBD10	PDZD8	SLITRK3
ANXA7	DSG4	KCNA1	PDZK1	SLITRK4
ANXA8L2	DSP	KCNA10	PDZK1IP1	SLITRK5
AOC3	DST	KCNA2	PEAR1	SLITRK6
AP1G1	DUOX1	KCNA3	PEBP1	SLK
AP2A1	DUOX2	KCNA4	PEBP4	SLMAP
AP2A2	DUOXA1	KCNA5	PECAM1	SMAD1
AP2B1	DUS1L	KCNA6	PEMT	SMAD7
AP2M1	DUSP15	KCNA7	PERP	SMAGP
AP2S1	DVL1	KCNB1	PFAS	SMAP1
AP3B1	DXO	KCNB2	PFDN1	SMO
APBA1	DYNC1H1	KCNC2	PFN1	SMOC1
APBA2	DYNC2H1	KCNC3	PGAP1	SMPD1
APBA3	DYSF	KCNC4	PGAP3	SMPD2
APBB1	EBAG9	KCND1	PGCP	SMPD4
APBB1IP	EBI3	KCND2	PGF	SMPDL3B
APC	ECE1	KCND3	PGLYRP2	SNAP23

APCDD1	ECEL1	KCNE1	PGLYRP4	SNAP25
APLF	EDA	KCNE1L	PGM5	SNAP29
APLNR	EDA2R	KCNE2	PGRMC1	SNAP91
APLP1	EDAR	KCNE3	PGRMC2	SNCG
APLP2	EDEM3	KCNE4	PHACTR4	SNPH
APMAP	EDNRA	KCNF1	PHB	SNRK
APOB	EDNRB	KCNG2	PHB2	SNRPD2
APOD	EEF1A1	KCNG3	PHEX	SNTA1
APOH	EEF1D	KCNG4	PHF20L1	SNX1
APOM	EEF1G	KCNH1	PHGDH	SNX15
APP	EEF2	KCNH2	PHKA1	SNX2
APR3	EEPD1	KCNH3	PHLDB2	SNX26
AQP1	EFCAB1	KCNH4	PHLDB3	SNX5
AQP10	EFCAB2	KCNH8	PI16	SNX6
AQP2	EFCAB7	KCNIP2	PI3	SNX9
AQP3	EFEMP1	KCNIP3	PI4K2A	SOAT2
AQP4	EFHC1	KCNJ1	PI4K2B	SOCS5
AQP5	EFNA1	KCNJ10	PI4KA	SOCS7
AQP6	EFNA2	KCNJ11	PI4KB	SOD1
AQP7	EFNA3	KCNJ12	PICALM	SORBS1
AQP8	EFNA4	KCNJ13	PICK1	SORBS2
AQP9	EFNA5	KCNJ14	PIGG	SORBS3
AR	EFNB1	KCNJ15	PIGN	SORCS1

ARAP1	EFNB2	KCNJ16	PIGR	SORCS2
ARAP3	EFNB3	KCNJ2	PIGS	SORCS3
AREG	EFR3A	KCNJ3	PIGT	SORD
ARF1	EGF	KCNJ4	PIGU	SORL1
ARF3	EGFR	KCNJ5	PIGW	SORT1
ARF4	EHBP1	KCNJ6	PIK3AP1	SOS1
ARF5	EHBP1L1	KCNJ8	PIK3C2B	SPA17
ARF6	EHD1	KCNJ9	PIK3C2G	SPACA4
ARFRP1	EHD2	KCNK1	PIK3CA	SPAG1
ARG1	EHD3	KCNK12	PIK3CG	SPAG16
ARHGAP18	EHD4	KCNK13	PIK3R1	SPAG9
ARHGAP19	EIF2B1	KCNK17	PIK3R2	SPAM1
ARHGAP21	EIF2B2	KCNK18	PIK3R5	SPANXB1
ARHGAP22	EIF5	KCNK2	PIKFYVE	SPATA12
ARHGAP23	EIF5B	KCNK5	PILRA	SPATA9
ARHGAP24	ELANE	KCNK6	PILRB	SPATC1L
ARHGAP29	ELAVL4	KCNK7	PIM1	SPERT
ARHGAP33	ELF5	KCNMA1	PIP4K2A	SPG11
ARHGAP36	ELFN1	KCNMB1	PIP4K2B	SPG20
ARHGAP40	ELFN2	KCNMB2	PIP5K1A	SPHK1
ARHGAP5	ELMO1	KCNMB3	PIP5K1C	SPINK13
ARHGEF1	ELMO2	KCNMB4	PITPNM3	SPINK2
ARHGEF11	ELMO3	KCNN1	PKD1	SPINT2

ARHGEF12	ELN	KCNN2	PKD1L1	SPN
ARHGEF15	ELOVL4	KCNN3	PKD1L2	SPPL2A
ARHGEF26	ELOVL6	KCNN4	PKD1L3	SPPL2B
ARHGEF28	ELTD1	KCNQ2	PKD2	SPPL3
ARHGEF5	EMB	KCNQ3	PKD2L1	SPRED2
ARHGEF7	EMD	KCNQ4	PKD2L2	SPRED3
ARHGEF9	EMID2	KCNQ5	PKDREJ	SPRR1A
ARID2	EMILIN1	KCNS1	PKHD1	SPRR3
ARID4A	EMILIN2	KCNS2	PKHD1L1	SPRY1
ARID5B	EMP1	KCNS3	PKM2	SPRY2
ARL15	EMP2	KCNT2	PKP2	SPRY3
ARL4C	EMP3	KCNV1	PKP3	SPSB4
ARL6IP1	EMR1	KCNV2	PKP4	SPTA1
ARL6IP5	EMR2	KCTD16	PLA2G2F	SPTAN1
ARMC12	EMR3	KDELC2	PLA2G3	SPTB
ARMCX1	EN1	KDR	PLA2G5	SPTBN1
ARMCX2	EN2	KEL	PLA2R1	SPTBN2
ARMCX3	ENAH	KIAA0090	PLAA	SQRDL
ARMCX6	ENG	KIAA0195	PLAT	SRC
ARPC1B	ENKD1	KIAA0247	PLAU	SRCIN1
ARPC5	ENO1	KIAA0319L	PLAUR	SRD5A3
ARR3	ENO2	KIAA0494	PLBD1	SRI
ARRB1	ENO3	KIAA0556	PLBD2	SRPK1

ARRB2	ENOX1	KIAA0644	PLCB1	SRPRB
ARSB	ENOX2	KIAA0895	PLCD1	SRPX
ARSE	ENPEP	KIAA0922	PLCD3	SRRM3
ARSF	ENPP1	KIAA1026	PLCD4	SSB
ART1	ENPP2	KIAA1161	PLCE1	SSFA2
ART3	ENPP3	KIAA1324	PLCH2	SSH3
ART4	ENPP4	KIAA1324L	PLCL2	SSPN
ARVCF	ENPP7	KIAA1467	PLD1	SSR1
ASAH1	ENTPD1	KIAA1522	PLD2	SSR2
ASAH2	ENTPD2	KIAA1524	PLD3	SSTR1
ASAP1	ENTPD3	KIAA1549	PLD4	SSTR2
ASAP2	ENTPD6	KIAA1919	PLDN	SSTR3
ASGR1	ENTPD7	KIAA1958	PLEC1	SSTR4
ASGR2	EPB41	KIDINS220	PLEK	SSTR5
ASH1L	EPB41L1	KIF13A	PLEK2	SSX2IP
ASH2L	EPB41L2	KIF17	PLEKHA1	ST14
ASIC1	EPB41L3	KIF19	PLEKHA4	ST3GAL1
ASPH	EPB41L4A	KIF21A	PLEKHB1	ST3GAL3
ASPM	EPB41L4B	KIF26A	PLEKHG3	ST3GAL4
ASPN	EPB41L5	KIF26B	PLEKHG6	ST3GAL5
ASPSCR1	EPB42	KIF5A	PLEKHH2	ST3GAL6
ASTN1	EPB49	KIF5C	PLEKHO1	ST6GAL1
ATF6B	EPCAM	KIR2DL1	PLEKHO2	ST6GALNAC3

ATG13	EPDR1	KIR2DL2	PLIN2	ST6GALNAC4
ATG3	EPHA1	KIR2DL3	PLIN4	ST8SIA1
ATG7	EPHA2	KIR2DL4	PLK3	ST8SIA2
ATG9A	EPHA3	KIR2DL5A	PLLP	ST8SIA4
ATIC	EPHA4	KIR2DS1	PLOD1	ST8SIA5
ATL1	EPHA5	KIR2DS2	PLOD2	STAB1
ATOX1	EPHA7	KIR2DS3	PLOD3	STAB2
ATP10A	EPHA8	KIR2DS4	PLP1	STAMBP
ATP10B	EPHB1	KIR2DS5	PLP2	STAMBPL1
ATP10D	EPHB2	KIR3DL1	PLPP1	STAP2
ATP11B	EPHB3	KIR3DL2	PLPPR5	STARD4
ATP12A	EPHB4	KIR3DL3	PLS3	STAT2
ATP13A3	EPHB6	KIR3DS1	PLSCR1	STAT3
ATP13A4	EPM2A	KIRREL	PLSCR2	STBD1
ATP1A1	EPN1	KIRREL2	PLSCR4	STEAP1
ATP1A2	EPOR	KIRREL3	PLTP	STEAP2
ATP1A3	EPR1	KISS1R	PLVAP	STEAP4
ATP1A4	EPS15	KIT	PLXDC1	STIL
ATP1B1	EPS15L1	KITLG	PLXNA1	STIM1
ATP1B2	EPS8	KL	PLXNA2	STIM2
ATP1B3	EPS8L2	KLC2	PLXNA3	STIP1
ATP1B4	EPX	KLF9	PLXNA4	STK17A
ATP2B1	ERAP1	KLHL41	PLXNB1	STOM

ATP2B2	ERAP2	KLK7	PLXNB2	STOML1
ATP2B3	ERBB2	KLRA1	PLXNB3	STOML2
ATP2B4	ERBB2IP	KLRB1	PLXNC1	STOML3
ATP2C2	ERBB3	KLRC1	PLXND1	STON1
ATP4A	ERBB4	KLRC2	PMEPA1	STON2
ATP4B	ERC1	KLRC4	PMP2	STRN
ATP5B	ERGIC2	KLRD1	PMP22	STS
ATP6AP1	ERLIN1	KLRF1	PNCK	STT3A
ATP6AP2	ERMAP	KLRG1	PNKD	STT3B
ATP6V0A1	ERMP1	KLRK1	PNN	STX12
ATP6V0A2	ERO1L	KMT2D	PNPLA6	STX18
ATP6V0A4	ERO1LB	KNG1	PNPLA8	STX1A
ATP6V0C	ERVWE1	KNTC1	PODXL	STX1B1
ATP6V0E1	ESAM	KPNA2	PODXL2	STX2
ATP6V1A	ESR1	KPTN	POFUT2	STX3
ATP6V1B1	ESYT1	KRAS	POLE	STX4
ATP6V1E1	ETNK1	KREMEN1	POMT1	STX6
ATP6V1F	EVA1A	KREMEN2	POMT2	STX8
ATP6V1G3	EVI2A	KRT1	PON1	STXBP1
ATP7A	EVI2B	KRT13	PON2	STXBP2
ATP8A2	EVL	KRT7	POR	STXBP4
ATP8B2	EVPL	KRT75	POSTN	STXBP5
ATPIF1	EWSR1	KRT77	POTEB	STYK1

ATRN	EXOC1	KRTAP25-1	POTEE	SUCLG1
ATRNL1	EXOC2	KSR1	POTEG	SUCLG2
ATRX	EXOC3	L1CAM	POU2AF1	SUCNR1
ATXN10	EXOC3L1	L3MBTL1	PPAP2A	SUGT1
ATXN3	EXOC4	LACE1	PPAP2B	SULF1
ATXN7L3	EXOC7	LAD1	PPAP2C	SULF2
AVEN	EXOC8	LAG3	PPFIA1	SUMF1
AVPI1	EXTL2	LAIR1	PPFIA2	SUPT6H
AVPR1A	EXTL3	LALBA	PPFIA3	SUSD1
AVPR1B	EZR	LAMA1	PPFIBP1	SUSD2
AVPR2	F11	LAMA2	PPFIBP2	SUSD5
AXIN1	F11R	LAMA3	PPHLN1	SV2A
AXL	F13A1	LAMA4	PPIAL4A	SVIL
AZU1	F2	LAMA5	PPIB	SWAP70
B2M	F2R	LAMB1	PPIL2	SYCP2
B3GALNT1	F2RL1	LAMB2	PPL	SYK
B3GALT2	F2RL2	LAMB3	PPM1A	SYMPK
B3GALTL	F2RL3	LAMC1	PPM1F	SYNC
B3GNT1	F3	LAMC2	PPM1L	SYNE2
B3GNT2	F5	LAMC3	PPP1CA	SYNGR2
B4GALNT2	FAAH	LAMP1	PPP1CB	SYNGR3
B4GALT1	FABP2	LAMP2	PPP1R12A	SYNGR4
B4GALT3	FABP5	LAMP3	PPP1R12B	SYNJ2

B4GALT5	FADD	LAMTOR1	PPP1R16A	SYNPO
BACE1	FAIM	LANCL1	PPP1R16B	SYNRG
BACE2	FAIM2	LANCL2	PPP1R18	SYPL1
BAG3	FAIM3	LAPTM4A	PPP1R9B	SYT13
BAI1	FAM101B	LASP1	PPP2R1B	SYT14
BAI2	FAM126A	LASS2	PPP2R2B	SYT16
BAI3	FAM129A	LASS6	PPP3R1	SYT2
BAIAP2	FAM129B	LAT	PPP3R2	SYT4
BAIAP2L1	FAM163B	LAX1	PPP4C	SYT5
BAMBI	FAM171A1	LAYN	PPP5C	SYT9
BANK1	FAM171A2	LBP	PPP6R3	SYTL1
BASP1	FAM174A	LCK	PPT1	SYTL2
BAT2L2	FAM174B	LCP1	PPT2	SYVN1
BBOF1	FAM189B	LCP2	PPTC7	TAAR1
BCAM	FAM193A	LCT	PPYR1	TAAR2
BCAN	FAM20B	LCTL	PRAP1	TAAR5
BCAR1	FAM210B	LDHC	PRC1	TAAR6
BCAR3	FAM213B	LDLR	PRDX2	TAAR8
BCHE	FAM222A	LDLRAP1	PRDX6	TAAR9
BCL2L10	FAM38A	LEMD2	PRELP	TACR1
BDKRB1	FAM38B	LENG1	PREX1	TACR2
BDKRB2	FAM45A	LEPR	PREX2	TACR3
BDNF	FAM57A	LEPRE1	PRF1	TACSTD2

BEND6	FAM72A	LEPROTL1	PRG2	TAGLN2
BEST1	FAM72B	LETM1	PRICKLE3	TAP1
BFSP1	FAM72C	LGALS1	PRIMA1	TAPBPL
BFSP2	FAM72D	LGALS13	PRKCA	TAS1R1
BGN	FAM84B	LGALS3	PRKCB	TAS1R2
BICD1	FAM89B	LGALS3BP	PRKCD	TAS2R1
BIN1	FANCG	LGALS4	PRKCDBP	TAS2R16
BIRC2	FAP	LGALS7	PRKCG	TAS2R3
BIRC3	FARS2	LGALS8	PRKCH	TAS2R38
BIRC6	FAS	LGALS9	PRKCI	TAS2R4
BLNK	FASLG	LGMN	PRKCQ	TAS2R5
BLVRB	FASN	LGR4	PRKCSH	TAS2R60
BMP1	FAT1	LGR5	PRKCZ	TAS2R7
BMPR1A	FAT3	LGR6	PRKD1	TAS2R8
BMPR1B	FAT4	LGTN	PRKD2	TAS2R9
BMPR2	FBLN2	LHCGR	PRKD3	TASP1
BMX	FBLN5	LHFP	PRKG2	TAZ
BNIP3	FBLN7	LIFR	PRLHR	TBC1D10A
BOC	FBN1	LILRA1	PRLR	TBC1D24
BPI	FBN2	LILRA4	PRNP	TBC1D30
BRAF	FBP2	LILRA5	PROCR	TBC1D3F
BRI3BP	FBRSL1	LILRA6	PROKR1	TBC1D9B
BRPF1	FBXL2	LILRB1	PROKR2	TBXA2R

BRS3	FBXO16	LILRB2	PROM1	TCAF2
BSCL2	FBXO38	LILRB3	PROM2	TCHP
BSG	FBXO41	LILRB4	PRPH2	TCIRG1
BSND	FCAMR	LILRB5	PRR7	TCL1A
BST1	FCAR	LIM2	PRRG1	TCP11
BST2	FCER1A	LIMA1	PRRG2	TCRB
BTD	FCER1G	LIMD1	PRRG3	TCTEX1D4
BTK	FCER2	LIMS2	PRRT2	TCTN2
BTLA	FCGBP	LIN7A	PRRT3	TDP1
BTN1A1	FCGR1A	LIN7B	PRRT4	TEC
BTN2A1	FCGR1C	LINGO1	PRSS21	TEK
BTN3A3	FCGR2A	LIPA	PRSS23	TES
BTRC	FCGR2B	LIPC	PRSS35	TESK1
BVES	FCGR2C	LIPG	PRSS8	TEX14
BZW2	FCGR3A	LIPH	PRTG	TEX9
C10ORF54	FCGR3B	LIPI	PRTN3	TF
C10ORF72	FCGRT	LLGL1	PRX	TFPI
C11ORF24	FCHSD2	LMAN2	PSAP	TFPI2
C11ORF30	FCN1	LMAN2L	PSCA	TFR2
C12ORF49	FCRL1	LMBR1	PSD	TFRC
C12ORF73	FCRL2	LMBR1L	PSD4	TGFA
C12ORF76	FCRL3	LMBRD1	PSEN1	TGFB1
C14ORF1	FCRL5	LMBRD2	PSG1	TGFBR1

C14ORF37	FCRL6	LMCD1	PSG2	TGFBR2
C14ORF93	FCRLB	LMF2	PSMC5	TGFBR3
C16ORF54	FERMT2	LMLN	PSMC6	TGM2
C16ORF74	FERMT3	LMNA	PSMD9	TGOLN2
C17ORF28	FES	LMTK2	PSMG1	THBD
C17ORF58	FEZ1	LNPEP	PSTPIP1	THBS1
C17ORF62	FFAR1	LNX2	PSTPIP2	THBS2
C18ORF19	FFAR2	LONRF1	PTAFR	THBS3
C18ORF21	FFAR3	LONRF3	PTBP1	THEM4
C19ORF33	FGA	LOR	PTCH1	THOP1
C19ORF48	FGB	LOX	PTCH2	THSD1
C19ORF59	FGD1	LOXL4	PTCHD4	THSD4
C19ORF6	FGF2	LPAR1	PTCRA	THSD7A
C19ORF71	FGFR1	LPAR2	PTEN	THY1
C1GALT1C1	FGFR2	LPAR3	PTGDR	TIAM1
C1ORF116	FGFR3	LPAR4	PTGER1	TICAM1
C1ORF159	FGFR4	LPAR5	PTGER2	TIE1
C1ORF27	FGFRL1	LPAR6	PTGER3	TIGD7
C1ORF35	FGL2	LPHN1	PTGER4	TIGIT
C1ORF74	FGR	LPHN2	PTGFR	TIMD4
C1ORF85	FHIT	LPHN3	PTGFRN	TIMM17B
C1ORF9	FHL1	LPIN2	PTGIR	TIMP1
C1QBP	FHL3	LPL	PTGIS	TJAP1

C1QTNF1	FICD	LPP	PTGS2	TJP1
C1QTNF2	FIGN	LRBA	PTH1R	TJP2
C2ORF72	FILIP1	LRCH2	PTH2R	TJP3
C3	FILIP1L	LRFN1	PTK2B	TLL1
C3AR1	FKBP10	LRIG1	PTK6	TLL2
C3ORF14	FKBP14	LRIG2	PTK7	TLN1
C3ORF38	FKBP2	LRP1	PTOV1	TLN2
C4A	FKBP7	LRP10	PTP4A2	TLR1
C4BPA	FKBP9	LRP11	PTP4A3	TLR10
C4BPB	FKRP	LRP12	PTPN11	TLR2
C4ORF22	FKTN	LRP1B	PTPN12	TLR3
C5AR1	FLAD1	LRP2	PTPN13	TLR4
C5ORF45	FLCN	LRP3	PTPN20A	TLR5
C6ORF120	FLG2	LRP4	PTPN22	TLR6
C6ORF25	FLNA	LRP5	PTPN3	TLR7
C7ORF42	FLNB	LRP6	PTPN4	TLR8
C7ORF43	FLNC	LRPAP1	PTPN7	TLR9
C8ORF37	FLOT1	LRPPRC	PTPRA	TM2D1
C8ORF4	FLOT2	LRRC15	PTPRB	TM2D2
C9ORF5	FLRT1	LRRC17	PTPRC	TM2D3
C9ORF91	FLRT2	LRRC25	PTPRCAP	TM4SF1
CA12	FLRT3	LRRC26	PTPRD	TM4SF20
CA14	FLT1	LRRC32	PTPRE	TM4SF4

CA2	FLT3	LRRC4	PTPRF	TM4SF5
CA4	FLT3LG	LRRC45	PTPRG	TM7SF3
CA9	FLT4	LRRC4B	PTPRH	TM7SF4
CABP1	FLVCR1	LRRC4C	PTPRJ	TM9SF1
CACHD1	FLVCR2	LRRC69	PTPRK	TM9SF3
CACNA1B	FMN1	LRRC7	PTPRM	TMC1
CACNA1C	FMN2	LRRC8A	PTPRN	TMC2
CACNA1D	FMNL1	LRRC8B	PTPRN2	TMC5
CACNA1E	FMNL2	LRRC8C	PTPRQ	TMC7
CACNA1I	FMNL3	LRRC8D	PTPRR	TMCO3
CACNA2D1	FN1	LRRFIP2	PTPRS	TMED1
CACNA2D2	FNDC4	LRRIQ3	PTPRT	TMED10
CACNA2D3	FNDC5	LRRN1	PTPRU	TMED2
CACNA2D4	FOLH1	LRRN4CL	PTPRZ1	TMED4
CACNB1	FOLR1	LSAMP	PTRF	TMED7
CACNB2	FOLR2	LSP1	PTTG1IP	TMED9
CACNB4	FOLR3	LSR	PURG	TMEFF1
CACNG1	FOXN3	LST1	PVR	TMEFF2
CACNG2	FPR1	LTA	PVRL1	TMEM101
CACNG3	FPR2	LTB	PVRL2	TMEM104
CACNG4	FPR3	LTB4R	PVRL3	TMEM105
CACNG5	FRAS1	LTB4R2	PVRL4	TMEM106B
CACNG6	FREM2	LTBP1	PXDC1	TMEM106C

CACNG7	FRMD3	LTBP3	PXDN	TMEM11
CACNG8	FRMD4B	LTBR	PXK	TMEM115
CADM1	FRMD6	LTC4S	PXN	TMEM123
CADM2	FRS2	LTF	PYGL	TMEM131
CADM3	FRS3	LTK	QRFPR	TMEM132A
CADM4	FRZB	LUM	QSOX1	TMEM132B
CADPS	FSCN1	LUZP2	QSOX2	TMEM132E
CADPS2	FSCN2	LY6D	RAB11B	TMEM139
CALB1	FSD1L	LY6E	RAB11FIP1	TMEM144
CALCR	FSHR	LY6G5B	RAB11FIP2	TMEM150A
CALCRL	FSTL1	LY6G6C	RAB11FIP3	TMEM156
CALD1	FTCD	LY6G6D	RAB11FIP5	TMEM160
CALM1	FUCA1	LY6H	RAB13	TMEM161A
CALM2	FURIN	LY6K	RAB17	TMEM176B
CALM3	FUT11	LY75	RAB18	TMEM179B
CALML3	FUT4	LY86	RAB20	TMEM181
CALML5	FUT5	LY9	RAB23	TMEM183A
CALR	FUT6	LYN	RAB27A	TMEM183B
CALU	FXYD1	LYPD3	RAB27B	TMEM194A
CALY	FXYD2	LYVE1	RAB31	TMEM2
CAMK1G	FXYD3	LZTFL1	RAB32	TMEM200B
CAMK2A	FXYD4	LZTS1	RAB33A	TMEM206
CAMK2D	FXYD5	LZTS2	RAB34	TMEM218

CAMK2G	FXYD7	M6PR	RAB3A	TMEM222
CANT1	FYB	MACF1	RAB3B	TMEM229A
CANX	FYN	MADCAM1	RAB3C	TMEM245
CAP2	FZD1	MADD	RAB3D	TMEM25
CAPN1	FZD10	MAEA	RAB3GAP2	TMEM254
CAPN10	FZD2	MAG	RAB43	TMEM258
CAPN2	FZD3	MAGED1	RAB4A	TMEM26
CAPNS1	FZD4	MAGED4	RAB5A	TMEM266
CAPRIN1	FZD5	MAGI1	RAB5B	TMEM27
CAPS	FZD6	MAGI3	RAB5C	TMEM30A
CARD10	FZD7	MAL	RAB6A	TMEM30B
CARD11	FZD8	MALL	RAB8A	TMEM47
CARD14	FZD9	MAMDC2	RAB8B	TMEM50A
CASC4	G3BP1	MAN2A1	RABAC1	TMEM55B
CASD1	G6B	MAN2A2	RABL2A	TMEM61
CASK	GAA	MAN2B1	RAC1	TMEM63A
CASP4	GAB1	MAN2B2	RAC2	TMEM63B
CASP7	GAB2	MAP1A	RAC3	TMEM63C
CASQ2	GABARAP	MAP1B	RAD52	TMEM67
CASR	GABBR1	MAP2K1	RAET1G	TMEM87A
CATSPER1	GABBR2	MAP2K4	RAF1	TMEM87B
CATSPER4	GABRA1	MAP2K5	RAI2	TMEM8A
CAV1	GABRA2	MAP2K7	RALA	TMEM8B

CAV2	GABRA3	MAP3K12	RALBP1	TMEM9
CAV3	GABRA4	MAP3K6	RALGAPA2	TMEM97
CBL	GABRA5	MAP3K7	RALGAPB	TMIGD2
CBLN1	GABRA6	MAP3K7IP2	RALGDS	TMPRSS11D
CBWD1	GABRB1	MAP4	RALGPS2	TMPRSS2
CBWD2	GABRB3	MAP4K2	RALY	TMPRSS3
CBWD3	GABRD	MAP4K5	RAMP1	TMPRSS4
CBWD5	GABRE	MAPK8IP1	RAMP2	TMPRSS5
CBWD7	GABRG1	MAPKAP1	RAMP3	TMPRSS6
CC2D1A	GABRG2	MAPT	RAP1A	TMPRSS9
CCBE1	GABRP	MARCH10	RAP1B	TMX1
CCBP2	GABRQ	MARCH2	RAP1GAP	TMX3
CCDC107	GABRR1	MARCH3	RAP2A	TMX4
CCDC109B	GABRR2	MARCH7	RAP2B	TNC
CCDC112	GADL1	MARCKS	RAPGEF2	TNF
CCDC115	GAGE12H	MARCO	RAPGEF6	TNFRSF10B
CCDC124	GAK	MARK2	RAPH1	TNFRSF10C
CCDC126	GALNS	MAS1	RAPSN	TNFRSF10D
CCDC144NL	GALNT1	MAS1L	RASA1	TNFRSF11A
CCDC15	GALNT10	MAST1	RASA2	TNFRSF11B
CCDC154	GALNT11	MAST2	RASA3	TNFRSF12A
CCDC24	GALNT15	MBOAT2	RASA4	TNFRSF13B
CCDC62	GALNT5	MBOAT7	RASAL1	TNFRSF13C

CCDC69	GALNT8	MBP	RASAL2	TNFRSF14
CCDC8	GALR1	MBTPS2	RASD1	TNFRSF17
CCDC80	GALR2	MC1R	RASD2	TNFRSF18
CCDC88A	GAP43	MC2R	RASGRF1	TNFRSF19
CCDC93	GAPDH	MC3R	RASGRF2	TNFRSF1A
CCDC97	GAPDHS	MC4R	RASGRP1	TNFRSF1B
CCKAR	GAS1	MC5R	RASGRP2	TNFRSF21
CCKBR	GAS2	MCA32	RASGRP3	TNFRSF25
CCL26	GAS2L1	MCAM	RASGRP4	TNFRSF4
CCNB1	GAS7	MCC	RASSF3	TNFRSF8
CCND3	GAS8	MCEMP1	RASSF5	TNFRSF9
CCNT2	GBA	MCF2	RC3H1	TNFSF10
CCNY	GBAS	MCF2L	RC3H2	TNFSF11
CCNYL1	GC	MCF2L2	RCCD1	TNFSF12
CCR1	GCA	MCHR1	RCN1	TNFSF13
CCR10	GCC1	MCHR2	RCN3	TNFSF13B
CCR2	GCC2	MCOLN1	RDH12	TNFSF14
CCR3	GCGR	MCOLN3	RDH8	TNFSF15
CCR4	GCLM	MDGA1	RDX	TNFSF18
CCR5	GDA	ME1	RECK	TNFSF4
CCR6	GDAP1	MEGF8	RELL1	TNFSF8
CCR7	GDE1	MEN1	RELN	TNFSF9
CCR8	GDF15	MEP1A	REM2	TNMD

CCR9	GDI2	MEP1B	REPS1	TNS1
CCRN4L	GDPD2	MERTK	RERG	TNS4
CCT3	GDPD5	MET	RET	TNXB
CCT4	GEM	METAP1	RFNG	TOB1
CD101	GFOD1	METAP2	RFPL1	TOLLIP
CD109	GFRA1	METTL15	RFPL2	TOM1
CD14	GFRA2	METTL9	RFPL3	TOR1AIP1
CD151	GFRA3	MFAP3	RFTN1	TOR1AIP2
CD160	GFRA4	MFAP3L	RGL4	TOR1B
CD163	GGCX	MFAP4	RGP1	TOR3A
CD163L1	GGH	MFAP5	RGS1	TPBG
CD164	GGT1	MFGE8	RGS10	TPCN1
CD177	GGT3P	MFI2	RGS13	TPD52L1
CD180	GGT5	MFNG	RGS14	TPI1
CD19	GGT7	MFRP	RGS16	TPO
CD1A	GGTLC2	MFSD2A	RGS17	TPP1
CD1B	GHR	MFSD4	RGS19	TPP2
CD1C	GHRHR	MFSD6	RGS2	TPRN
CD1D	GHSR	MFSD8	RGS20	TPSG1
CD2	GINM1	MGAM	RGS4	TPT1
CD200	GIPC1	MGAT4B	RGS5	TPTE
CD200R1	GIPR	MGAT5	RGS7	TRAC
CD207	GIT2	MGAT5B	RGS8	TRADD

CD209	GJA1	MIA3	RGS9	TRAF2
CD22	GJA3	MIB1	RHAG	TRAF3
CD226	GJA4	MICA	RHBDF2	TRAF6
CD24	GJA5	MICALCL	RHBDL1	TRAF7
CD244	GJA8	MICB	RHBG	TRAIP
CD247	GJA9	MINPP1	RHCE	TRAK2
CD248	GJB1	MIP	RHCG	TRAP1
CD27	GJB2	MISP	RHD	TRAPPC13
CD274	GJB3	MKLN1	RHEB	TRAPPC4
CD276	GJB4	MKRN3	RHO	TRAT1
CD28	GJB5	MKS1	RHOA	TRBC1
CD2AP	GJB6	MLANA	RHOB	TRBV12-3
CD300A	GJB7	MLC1	RHOBTB2	TRDN
CD300C	GJC1	MLF2	RHOC	TREH
CD300E	GJC2	MLLT4	RHOD	TREM1
CD300LF	GJC3	MLNR	RHOF	TREM2
CD302	GJD2	MMACHC	RHOG	TREML1
CD320	GJD3	MMD	RHOJ	TRGV9
CD33	GLA	MME	RHOQ	TRHDE
CD34	GLB1	MMGT1	RIC3	TRHR
CD36	GLCE	MMP10	RIC8A	TRIB1
CD37	GLG1	MMP13	RIC8B	TRIM16
CD38	GLIPR1	MMP14	RILPL1	TRIM16L

CD3D	GLIS2	MMP15	RIMS2	TRIM21
CD3E	GLO1	MMP16	RIN1	TRIM38
CD3G	GLP1R	MMP17	RIOK2	TRIM4
CD4	GLP2R	MMP19	RIOK3	TRIM44
CD40	GLRA1	MMP2	RIPK1	TRIM71
CD40LG	GLRA2	MMP24	RIT1	TRIML2
CD44	GLRA3	MMP25	RIT2	TRIP10
CD46	GLRB	MMP9	RNASE2	TRIP6
CD47	GLRX	MMRN1	RND1	TRMT1
CD48	GLT25D1	MOCS1	RND3	TRMT10B
CD5	GLT25D2	MORN4	RNF114	TRMT13
CD52	GLT8D1	MOXD1	RNF13	TRO
CD53	GLT8D2	MPDZ	RNF130	TROVE2
CD55	GLTPD2	MPHOSPH9	RNF139	TRPA1
CD58	GMFB	MPI	RNF144B	TRPC1
CD59	GMIP	MPL	RNF145	TRPC3
CD6	GML	MPO	RNF146	TRPC4
CD63	GMPPB	MPP1	RNF149	TRPC5
CD68	GMPS	MPP2	RNF165	TRPC6
CD69	GNA11	MPP5	RNF166	TRPC7
CD7	GNA12	MPP6	RNF5	TRPM1
CD70	GNA13	MPV17L2	RNPEP	TRPM2
CD72	GNA14	MPZ	ROBO1	TRPM3

CD74	GNA15	MPZL1	ROBO4	TRPM4
CD79A	GNAI1	MPZL2	ROM1	TRPM5
CD79B	GNAI2	MR1	ROR1	TRPM6
CD80	GNAI3	MRAP	ROR2	TRPM7
CD81	GNAL	MRAS	ROS1	TRPM8
CD82	GNAO1	MRC1	RP11-192H23.4	TRPV1
CD83	GNAQ	MRC2	RP11-298I3.5	TRPV2
CD84	GNAS	MRGPRD	RP11-385D13.1	TRPV3
CD86	GNAT1	MRGPRF	RP2	TRPV4
CD8A	GNAT2	MRGPRX1	RP5-862P8.2	TRPV5
CD8B	GNAZ	MRGPRX2	RP5-877J2.1	TRPV6
CD9	GNB1	MRGPRX3	RPIA	TSC2
CD93	GNB2	MRGPRX4	RPL17	TSG101
CD96	GNB2L1	MRPL42	RPL18	TSHR
CD97	GNB3	MRPL44	RPL19	TSPAN1
CDAN1	GNB4	MRVI1	RPL21	TSPAN13
CDC42	GNG10	MS4A1	RPL27A	TSPAN14
CDC42EP1	GNG7	MS4A14	RPL28	TSPAN15
CDC42EP2	GNG8	MS4A2	RPL29	TSPAN18
CDC42EP4	GNGT1	MSLN	RPL3	TSPAN2
CDC42EP5	GNGT2	MSN	RPL32	TSPAN3
CDC42SE1	GNPTAB	MSR1	RPL35	TSPAN31
CDC42SE2	GNPTG	MST1R	RPL35A	TSPAN33

CDCP1	GNRHR	MTCP1	RPL36A	TSPAN4
CDH1	GNS	MTDH	RPL36A-HNRNPH2	TSPAN6
CDH10	GOLIM4	MTM1	RPL36AL	TSPAN7
CDH11	GOLM1	MTMR1	RPL37A	TSPAN8
CDH12	GOLPH3L	MTMR14	RPL9	TSPAN9
CDH13	GOPC	MTMR3	RPLP2	TTC17
CDH15	GP1BA	MTNR1A	RPN1	TTC23
CDH16	GP1BB	MTNR1B	RPN2	TTC7B
CDH17	GP2	MTOR	RPS15A	TTLL12
CDH18	GP5	MTPN	RPS23	TTLL5
CDH19	GP9	MTUS1	RPS6KB1	TTMA
CDH2	GPA33	MUC1	RPSA	TTYH2
CDH20	GPBAR1	MUC13	RRAD	TTYH3
CDH22	GPBP1	MUC15	RRAGA	TUB
CDH23	GPBP1L1	MUC16	RRAS	TUBA1B
CDH24	GPC1	MUC20	RRAS2	TUBA4A
CDH26	GPC2	MUC4	RRH	TUBB
CDH3	GPC3	MUSK	RSAD1	TUBB2C
CDH4	GPC4	MXRA5	RSBN1L	TUBB4
CDH5	GPC5	MXRA8	RSC1A1	TUBB6
CDH6	GPC6	MYD88	RSL1D1	TULP3
CDH8	GPER	MYH11	RSPH3	TWF2

CDH9	GPI	MYH9	RTN4	TXLNA
CDHR1	GPM6A	MYL2	RTN4R	TXLNB
CDHR2	GPM6B	MYL4	RTN4RL1	TXNDC11
CDHR5	GPNMB	MYL6B	RTN4RL2	TXNDC15
CDIPT	GPR1	MYLK	RTP1	TXNDC5
CDK5	GPR101	MYO10	RTP2	TXNIP
CDK5R1	GPR107	MYO15A	RTP4	TXNRD1
CDK5R2	GPR109A	MYO1B	RUFY1	TYR
CDKN1B	GPR109B	MYO1C	RUSC1	TYRO3
CDKN2A	GPR110	MYO1G	RUVBL1	TYROBP
CDON	GPR112	MYO3A	RXFP1	TYRP1
CDR1	GPR113	MYO5B	RXFP2	TYW3
CDSN	GPR116	MYO5C	RXFP3	UAP1
CDV3	GPR119	MYO7A	RXFP4	UBA1
CEACAM1	GPR12	MYO9A	RYK	UBA3
CEACAM3	GPR124	MYOF	RYR2	UBA52
CEACAM4	GPR125	MYOT	RYR3	UBAC1
CEACAM5	GPR126	MYSM1	S100A10	UBAP1
CEACAM6	GPR128	N4BP2	S100A11	UBASH3B
CEACAM7	GPR132	N6AMT2	S100A14	UBE2A
CEACAM8	GPR135	NAAA	S100A16	UBE2B
CECR6	GPR137B	NAALAD2	S100A2	UBE2C
CELSR1	GPR143	NAALADL1	S100A4	UBE2D2

CELSR2	GPR146	NAALADL2	S100A7	UBE2D3
CELSR3	GPR15	NAB1	S100A8	UBE2D4
CEP162	GPR150	NAGA	S100A9	UBE2N
CEP295	GPR155	NAGLT1	S100P	UBE2S
CEP55	GPR156	NAGLU	S1PR1	UBE3C
CEP57L1	GPR160	NAGPA	S1PR2	UBIAD1
CEP89	GPR17	NAIF1	S1PR3	UBQLN2
CERCAM	GPR172A	NALCN	S1PR4	UBR2
CERK	GPR172B	NAP1L2	S1PR5	UBTD1
CES1	GPR173	NAPG	SAMHD1	UBXN11
CFAP157	GPR174	NARFL	SC4MOL	UCP3
CFAP97	GPR176	NAV2	SC5D	UGGT1
CFC1	GPR18	NBEA	SC65	UGGT2
CFD	GPR182	NCAM1	SCAMP1	UGT1A10
CFH	GPR19	NCAM2	SCAMP2	UGT1A9
CFL1	GPR25	NCEH1	SCAMP3	UGT8
CFL2	GPR26	NCF2	SCAMP4	UHRF1BP1L
CFLAR	GPR27	NCF4	SCAP	ULBP1
CFTR	GPR3	NCK1	SCARA3	ULBP2
CGN	GPR31	NCKAP1L	SCARA5	ULBP3
CHD9	GPR32	NCKAP5	SCARB1	UMOD
CHGA	GPR34	NCKIPSD	SCARB2	UMODL1
CHIC2	GPR37	NCL	SCARF1	UNC119

CHL1	GPR39	NCLN	SCARF2	UNC13A
CHMP2A	GPR4	NCOA1	SCEL	UNC13B
CHN1	GPR42	NCR1	SCFD1	UNC13D
CHN2	GPR44	NCR2	SCGN	UNC5A
CHODL	GPR45	NCR3	SCIN	UNC5B
CHPF2	GPR50	NCR3LG1	SCN10A	UNC5C
CHRM2	GPR52	NCS1	SCN11A	UNC84A
CHRM3	GPR55	NCSTN	SCN1A	UNC84B
CHRM4	GPR56	NDC1	SCN1B	UNQ870/PRO1886
CHRM5	GPR6	NEDD4	SCN2A	UPK1A
CHRNA1	GPR63	NEDD4L	SCN2B	UPK1B
CHRNA2	GPR64	NEDD9	SCN3A	UPK2
CHRNA3	GPR65	NEGR1	SCN4A	UPK3A
CHRNA4	GPR68	NELF	SCN4B	UPK3B
CHRNA5	GPR75	NELL1	SCN5A	UQCC1
CHRNA7	GPR85	NEO1	SCN7A	UQCRC2
CHRNB1	GPR98	NEU1	SCN8A	USP2
CHRNB2	GPRC5A	NEU3	SCN9A	USP20
CHRNB3	GPRC5B	NEURL	SCNN1A	USP21
CHRNB4	GPRC5C	NEXN	SCNN1B	USP4
CHRND	GPRC5D	NF2	SCNN1D	USP6
CHRNE	GPRIN1	NFAM1	SCNN1G	UST
CHRNG	GPX5	NFASC	SCOC	UTP20

CHST11	GRAMD1A	NGEF	SCP2	UTRN
CHST12	GRASP	NGFR	SCRIB	UTS2
CHST3	GRB14	NID2	SCTR	UTS2R
CHST6	GRB2	NIN	SCUBE1	VAMP1
CHST7	GRB7	NINJ1	SCUBE3	VAMP2
CHSY3	GREM1	NINJ2	SCYL3	VAMP3
CHUK	GRIA1	NIPAL4	SDC1	VAMP5
CIB1	GRIA2	NIPSNAP3A	SDC2	VAMP7
CKAP4	GRIA3	NISCH	SDC3	VAMP8
CKAP5	GRIA4	NKTR	SDC4	VANGL1
CLC	GRID1	NKX2-8	SDCBP	VANGL2
CLCA1	GRID2	NLGN1	SDE2	VAPA
CLCA2	GRIK1	NLGN2	SDK1	VAPB
CLCA3	GRIK2	NLGN3	SDK2	VASN
CLCN1	GRIK3	NLGN4X	SDPR	VASP
CLCN2	GRIK4	NLN	SEC16B	VAV1
CLCN3	GRIK5	NLRX1	SEC22B	VCAM1
CLCN4	GRIN1	NMB	SECTM1	VCAN
CLCN5	GRIN2A	NMBR	SEL1L	VCL
CLCN6	GRIN2B	NME2	SEL1L3	VCPIP1
CLCN7	GRIN2C	NMT1	SELE	VDAC1
CLCNKA	GRIN2D	NMT2	SELENBP1	VDAC2
CLCNKB	GRIN3A	NMUR1	SELL	VEGFC

CLDN1	GRK5	NOMO3	SELP	VEZT
CLDN11	GRK6	NOP56	SELPLG	VIL1
CLDN12	GRK7	NOS1	SEMA3A	VIM
CLDN14	GRM1	NOS2	SEMA3C	VIPR1
CLDN15	GRM2	NOS3	SEMA3D	VIPR2
CLDN16	GRM3	NOSTRIN	SEMA3F	VNN1
CLDN17	GRM4	NOTCH1	SEMA4A	VNN2
CLDN18	GRM5	NOTCH2	SEMA4B	VNN3
CLDN19	GRM6	NOTCH3	SEMA4C	VPREB1
CLDN2	GRM7	NOTCH4	SEMA4D	VPREB3
CLDN20	GRM8	NOX3	SEMA4F	VPS11
CLDN23	GRN	NOX4	SEMA4G	VPS13B
CLDN3	GRPR	NOXO1	SEMA5A	VPS26B
CLDN4	GSR	NPBWR1	SEMA6A	VPS35
CLDN5	GSTO1	NPBWR2	SEMA6B	VPS36
CLDN6	GTSE1	NPC1	SEMA6C	VPS37C
CLDN7	GUCY1B2	NPC1L1	SEMA6D	VPS41
CLDN8	GUCY2C	NPFF	SEMA7A	VPS8
CLDN9	GUCY2D	NPFFR2	SEPT3	VSIG10
CLDND1	GUCY2F	NPHP1	SEPT8	VSIG2
CLEC12A	GUSB	NPHS1	SERINC1	VSIG4
CLEC14A	GXYLT1	NPHS2	SERINC3	VSNL1
CLEC17A	GYPA	NPL	SERPINA1	VSTM1

CLEC1A	GYPB	NPR1	SERPINA3	VTCN1
CLEC1B	GYPC	NPR2	SERPINA5	VTN
CLEC2B	GYPE	NPR3	SERPINB2	VWA5B1
CLEC2D	GZMH	NPSR1	SERPINB3	VWA7
CLEC4A	HAPLN3	NPTN	SERPINB5	VWF
CLEC4D	HARBI1	NPTX1	SERPINE2	WARS2
CLEC4E	HAS1	NPTX2	SERPINF1	WAS
CLEC4F	HAS2	NPTXR	SERPING1	WASF1
CLEC4G	HAS3	NPY1R	SERPINH1	WASF2
CLEC5A	HAUS1	NPY2R	SERPINI1	WBSCR17
CLEC6A	HAUS4	NPY5R	SET	WDR12
CLIC1	HAVCR1	NQO1	SEZ6L2	WDR13
CLIC2	HAVCR2	NR3C1	SFRP4	WDR53
CLIC4	HBEGF	NRAS	SFRS17A	WDR6
CLIC5	HCFC2	NRBP1	SFRS3	WIPF1
CLIC6	HCK	NRCAM	SFSWAP	WNT11
CLIP3	HCLS1	NRD1	SFTPC	WNT5A
CLN3	HCN1	NRG1	SFXN1	WNT5B
CLN5	HCN2	NRG3	SGCA	WNT8A
CLN6	HCN4	NRP1	SGCB	WTIP
CLN8	HCRTR1	NRP2	SGCD	WWP1
CLNS1A	HCRTR2	NRXN1	SGCE	XG
CLPTM1	HCST	NRXN2	SGCZ	XIAP

CLPTM1L	HDAC8	NRXN3	SGK1	XIRP1
CLSTN1	HDLBP	NSD1	SGK196	XK
CLSTN2	HEG1	NSF	SGK494	XKRY
CLSTN3	HELZ	NSFL1C	SGMS1	XPC
CLTA	HENMT1	NSMAF	SGMS2	XPNPEP2
CLTB	HEPACAM	NT5E	SGPL1	XPO6
CLTC	HEPACAM2	NTM	SGSM3	XPR1
CLU	HEPH	NTN3	SH2B2	XRCC6
CMKLR1	HEXA	NTNG1	SH2B3	XRCC6BP1
CMTM2	HEXB	NTNG2	SH2D1A	XRN1
CMTM4	HFE	NTRK1	SH3BGR	XYLT2
CMTM6	HFE2	NTRK2	SH3BGRL3	YAP1
CMTM7	HGF	NTRK3	SH3D19	YBX1
CMTM8	HGSNAT	NTSR1	SH3D21	YES1
CNGA1	HHATL	NTSR2	SH3PXD2B	YIPF4
CNGA2	HHIP	NUCB2	SHANK1	YOD1
CNGA3	HHIPL1	NUDCD3	SHARPIN	YTHDC1
CNGA4	HHLA2	NUDT16L1	SHC1	YWHAQ
CNGB3	HIF1A	NUMB	SHH	YWHAZ
CNKSR2	HIF3A	NUP210	SHROOM2	ZAN
CNNM1	HINT1	NUP35	SHROOM3	ZAP70
CNNM2	HIP1R	NUP58	SHTN1	ZBED3
CNNM3	HK1	NUS1	SI	ZBTB33

CNNM4	HLA-A	NXF2	SIDT1	ZBTB8OS
CNP	HLA-B	NXPH4	SIDT2	ZC3HAV1
CNR1	HLA-C	NYX	SIGIRR	ZCCHC16
CNR2	HLA-DMA	OAS2	SIGLEC1	ZCCHC5
CNTFR	HLA-DMB	OAS3	SIGLEC10	ZDHHC22
CNTN1	HLA-DOA	OBSCN	SIGLEC11	ZDHHC23
CNTN2	HLA-DPA1	OCLN	SIGLEC12	ZDHHC5
CNTN3	HLA-DPB1	ODC1	SIGLEC15	ZFP90
CNTN4	HLA-DQA1	ODZ1	SIGLEC5	ZFYVE28
CNTNAP1	HLA-DQA2	ODZ2	SIGLEC6	ZGPAT
CNTNAP2	HLA-DQB1	ODZ3	SIGLEC7	ZMYND10
CNTNAP3	HLA-DQB2	ODZ4	SIGLEC8	ZMYND19
COBL	HLA-DRA	OGFOD1	SIGLEC9	ZNF10
COBLL1	HLA-DRB1	OGN	SIGMAR1	ZNF18
COL11A1	HLA-DRB3	OGT	SIL1	ZNF185
COL12A1	HLA-DRB4	OIP5	SILV	ZNF2
COL13A1	HLA-DRB5	OLFM1	SIN3B	ZNF20
COL14A1	HLA-E	OLFML2A	SIPA1	ZNF227
COL16A1	HLA-F	OLFML2B	SIPA1L1	ZNF229
COL17A1	HLA-G	OLFML3	SIRPA	ZNF276
COL1A1	HMCN1	OLR1	SIRPB1	ZNF277
COL1A2	HMGB1	OMG	SIRPG	ZNF337
COL23A1	HMGCS1	OPCML	SIT1	ZNF397

COL29A1	HMHA1	OPHN1	SKAP1	ZNF407
COL2A1	HMMR	OPN1LW	SLA	ZNF503
COL3A1	HN1L	OPN1MW	SLA2	ZNF510
COL4A3BP	HNRNPK	OPN1SW	SLAMF1	ZNF571
COL5A1	HNRNPM	OPN3	SLAMF6	ZNF598
COL5A2	HNRNPU	OPN4	SLAMF7	ZNF599
COL5A3	HOMER3	OPRD1	SLAMF8	ZNF620
COL6A1	HP	OPRM1	SLAMF9	ZNF646
COL6A2	HPCAL1	OR10A6	SLC10A1	ZNF655
COL6A3	HPN	OR10G7	SLC10A2	ZNF688
COL6A6	HPS3	OR10J3	SLC10A5	ZNF706
COL7A1	HPSE2	OR10V1	SLC11A1	ZNF717
COLEC12	HPX	OR11A1	SLC11A2	ZNF783
COLQ	HRAS	OR13C4	SLC12A1	ZNF785
COMMD4	HRG	OR13D1	SLC12A2	ZNF79
COMTD1	HRH1	OR13H1	SLC12A3	ZNF839
COPRS	HRH2	OR1D2	SLC12A4	ZNF844
COPS5	HRH3	OR1E2	SLC12A5	ZP1
COQ7	HRH4	OR1F1	SLC12A6	ZP3
CORIN	HRNR	OR1G1	SLC12A7	ZP4
CORO1B	HRSP12	OR1M1	SLC12A8	ZSCAN22
CORO2A	HS2ST1	OR2A4	SLC12A9	ZSCAN29
CORO2B	HS3ST2	OR2B6	SLC13A1	ZSWIM1

COX6B2	HS3ST3A1	OR2C1	SLC13A2	ZSWIM8
CP	HS3ST3B1	OR2C3	SLC13A5	ZYX

7.2.3. Comprehensive surfaceome

Table 7.3. Surface protein-encoding genes represented in the comprehensive surfaceome.

1C16	CTR2	ITM2C	OR4DB	SIN3B
2B11	CTR3	ITPKB	OR4E1	SIPA1
2B13	CTSA	ITPR1	OR4E2	SIPA1L1
2B14	CTSB	ITPR2	OR4F15	SIR2
2B17	CTSC	ITPR3	OR4F16	SIRPA
2B18	CTSD	ITPRIP	OR4F17	SIRPB1
2B19	CTSE	ITPRIPL1	OR4F21	SIRPG
2B1A	CTSF	ITPRIPL2	OR4F29	SIT1
2B1B	CTSG	ITSN1	OR4F3	SKAP1
2B1C	CTSK	ITSN2	OR4F4	SL9A1
2B1D	CTSL1	IVL	OR4F5	SL9A3
2B1E	CTSL2	IYD	OR4F6	SL9A4
2B1F	CTSR1	IYD1	OR4K1	SL9C1
2B1G	CTSR2	IZUM3	OR4K13	SLA
4F2	CTSR3	IZUMO1	OR4K14	SLA2
5HT1A	CTSR4	JAG1	OR4K15	SLAF1
5HT1B	CTSRD	JAG2	OR4K17	SLAF5
5HT1D	CTSS	JAGN1	OR4K2	SLAF6
5HT1E	CTSV	JAK2	OR4K3	SLAMF1
5HT1F	CTTN	JAM1	OR4K5	SLAMF6
5HT2A	CTXN1	JAM2	OR4KD	SLAMF7

5HT2B	CTXN3	JAM3	OR4KE	SLAMF8
5HT2C	CUBN	JAML	OR4KF	SLAMF9
5HT3A	CUL5	JKAMP	OR4KH	SLAP2
5HT3B	CUTA	JMJD4	OR4L1	SLC10A1
5HT3C	CUZD1	JMJD6	OR4M1	SLC10A2
5HT3D	CX3C1	JOS1	OR4M2	SLC10A3
5HT3E	CX3CL1	JPH1	OR4N2	SLC10A4
5HT4R	CX3CR1	JPH2	OR4N4	SLC10A5
5HT5A	CXA1	JPH3	OR4N5	SLC10A6
5HT6R	CXA10	JPH4	OR4P4	SLC10A7
5HT7R	CXA3	JTB	OR4Q2	SLC11A1
5NTD	CXA4	JUB	OR4Q3	SLC11A2
A2M	CXA5	JUNO	OR4S1	SLC12A1
A4GALT	CXA8	JUP	OR4S2	SLC12A2
AA1R	CXA9	K0319	OR4X1	SLC12A3
AA2AR	CXADR	K1324	OR4X2	SLC12A4
AA2BR	CXAR	K2C1	OR51A2	SLC12A5
AA3R	CXB1	K319L	OR51A4	SLC12A6
AAAT	CXB2	KAL1	OR51A7	SLC12A7
AADACL1	CXB3	KALM	OR51B2	SLC12A8
AADAT	CXB4	KANK1	OR51B4	SLC12A9
AAK1	CXB5	KANK3	OR51B5	SLC13A1
AAMP	CXB6	KANSL2	OR51B6	SLC13A2

AARS	CXB7	KAP0	OR51D1	SLC13A3
AATM	CXCL16	KAP1	OR51E1	SLC13A4
AB1IP	CXCR1	KAP2	OR51E2	SLC13A5
ABCA1	CXCR2	KAP3	OR51F1	SLC14A1
ABCA10	CXCR3	KAPCA	OR51F2	SLC14A2
ABCA12	CXCR4	KAPCB	OR51G1	SLC15A1
ABCA13	CXCR5	KARS	OR51G2	SLC15A2
ABCA2	CXCR6	KATNB1	OR51I1	SLC15A3
ABCA3	CXCR7	KBTBD10	OR51I2	SLC15A4
ABCA4	CXD2	KC1D	OR51L1	SLC16A1
ABCA5	CXD3	KCAB1	OR51M1	SLC16A10
ABCA6	CXD4	KCAB2	OR51Q1	SLC16A11
ABCA7	CXE1	KCC1G	OR51S1	SLC16A12
ABCA8	CXG1	KCC2A	OR51T1	SLC16A13
ABCA9	CXG2	KCC2D	OR51V1	SLC16A14
ABCB1	CXG3	KCD12	OR52A1	SLC16A2
ABCB11	CXL16	KCD16	OR52A4	SLC16A3
ABCB4	CXORF1	KCIP1	OR52A5	SLC16A4
ABCB5	CXORF59	KCIP2	OR52B2	SLC16A5
ABCB6	CXX1	KCIP4	OR52B4	SLC16A6
ABCC1	CY24A	KCJ10	OR52B6	SLC16A7
ABCC10	CY24B	KCJ12	OR52D1	SLC16A8
ABCC12	CYB561	KCJ16	OR52E2	SLC16A9

ABCC2	CYB561D1	KCJ18	OR52E4	SLC17A1
ABCC3	CYB561D2	KCMA1	OR52E5	SLC17A2
ABCC4	CYB5R3	KCNA1	OR52E6	SLC17A3
ABCC5	CYBA	KCNA10	OR52E8	SLC17A4
ABCC6	CYBASC3	KCNA2	OR52H1	SLC17A5
ABCC8	CYBB	KCNA3	OR52I1	SLC17A6
ABCC9	CYBRD1	KCNA4	OR52I2	SLC17A7
ABCD2	CYFIP2	KCNA5	OR52J3	SLC17A8
ABCG1	CYH1	KCNA6	OR52K1	SLC18A1
ABCG2	CYH2	KCNA7	OR52K2	SLC18A2
ABCG4	CYH3	KCNB1	OR52L1	SLC18A3
ABCG5	CYH4	KCNB2	OR52M1	SLC19A1
ABCG8	CYHR1	KCNC1	OR52N1	SLC19A2
ABHD1	CYLD	KCNC2	OR52N2	SLC19A3
ABHD12	CYP20A1	KCNC3	OR52N4	SLC1A1
ABHD13	CYP2U1	KCNC4	OR52N5	SLC1A2
ABHD3	CYP51A1	KCND1	OR52R1	SLC1A3
ABI1	CYR61	KCND2	OR52W1	SLC1A4
ABI3BP	CYS1	KCND3	OR56A1	SLC1A5
ABL1	CYSLTR1	KCNE1	OR56A3	SLC1A6
ABLIM3	CYSLTR2	KCNE1L	OR56A4	SLC1A7
ACAA1	CYTH1	KCNE2	OR56B4	SLC20A1
ACAM	CYTH2	KCNE3	OR5A1	SLC20A2

ACAP2	CYTH3	KCNE4	OR5A2	SLC22A1
ACBD4	CYTIP	KCNF1	OR5AC2	SLC22A10
ACBD5	CYYR1	KCNG1	OR5AK2	SLC22A11
ACCN1	DAAM1	KCNG2	OR5AN1	SLC22A12
ACCN2	DAB2	KCNG3	OR5AP2	SLC22A13
ACCN3	DAB2P	KCNG4	OR5AR1	SLC22A14
ACCN5	DAD1	KCNH1	OR5AS1	SLC22A15
ACE	DAF	KCNH2	OR5AU1	SLC22A16
ACE2	DAG1	KCNH3	OR5B12	SLC22A17
ACES	DAGLA	KCNH4	OR5B17	SLC22A18
ACH10	DAGLB	KCNH5	OR5B2	SLC22A2
ACHA	DAPK1	KCNH6	OR5B21	SLC22A23
ACHA2	DAPP1	KCNH7	OR5B3	SLC22A25
ACHA3	DARC	KCNH8	OR5BC	SLC22A3
ACHA4	DBH	KCNIP2	OR5BH	SLC22A4
ACHA5	DBN1	KCNIP3	OR5BL	SLC22A5
ACHA6	DBNL	KCNJ1	OR5C1	SLC22A6
ACHA7	DC2	KCNJ10	OR5D13	SLC22A7
ACHA9	DCAF16	KCNJ11	OR5D14	SLC22A8
ACHB	DCAKD	KCNJ12	OR5D16	SLC22A9
ACHB2	DCBLD1	KCNJ13	OR5D18	SLC23A1
ACHB3	DCBLD2	KCNJ14	OR5DD	SLC23A2
ACHB4	DCC	KCNJ15	OR5DE	SLC23A3

ACHD	DCE2	KCNJ16	OR5DG	SLC24A1
ACHE	DCHS1	KCNJ2	OR5DI	SLC24A2
ACHG	DCHS2	KCNJ3	OR5F1	SLC24A3
ACK1	DCLK1	KCNJ4	OR5G3	SLC24A4
ACKR2	DCN	KCNJ5	OR5H1	SLC24A5
ACKR3	DCNL3	KCNJ6	OR5H14	SLC24A6
ACKR4	DCST2	KCNJ8	OR5H15	SLC25A17
ACLY	DCSTP	KCNJ9	OR5H2	SLC25A3
ACM1	DCT	KCNK1	OR5H6	SLC25A34
ACM2	DCTN6	KCNK10	OR5H8	SLC25A4
ACM3	DCUN1D2	KCNK12	OR5I1	SLC25A5
ACM4	DDAH1	KCNK13	OR5J2	SLC26A1
ACM5	DDIAS	KCNK15	OR5K1	SLC26A10
ACMSD	DDR1	KCNK16	OR5K2	SLC26A11
ACOT9	DDR2	KCNK17	OR5K3	SLC26A2
ACP1	DDX58	KCNK18	OR5K4	SLC26A3
ACP2	DDX60L	KCNK2	OR5L1	SLC26A4
ACPT	DEF6	KCNK3	OR5L2	SLC26A5
ACSL5	DEFI6	KCNK4	OR5M1	SLC26A6
ACSL6	DEGS1	KCNK5	OR5M10	SLC26A7
ACTB	DEGS2	KCNK6	OR5M11	SLC26A8
ACTHR	DEMA	KCNK7	OR5M3	SLC26A9
ACTN1	DEN1A	KCNK9	OR5M8	SLC27A1

ACTN4	DEN4C	KCNKI	OR5M9	SLC27A3
ACTR6	DEND	KCNMA1	OR5MA	SLC27A4
ACV1B	DENND1B	KCNMB1	OR5MB	SLC27A6
ACVL1	DESM	KCNMB2	OR5P2	SLC28A1
ACVR1	DESP	KCNMB3	OR5P3	SLC28A2
ACVR1B	DFFA	KCNMB4	OR5R1	SLC28A3
ACVR1C	DFNB31	KCNN1	OR5T1	SLC29A1
ACVR2A	DGAT2	KCNN2	OR5T2	SLC29A2
ACVR2B	DGAT2L4	KCNN3	OR5T3	SLC29A3
ACVRL1	DGCR2	KCNN4	OR5V1	SLC29A4
ACY3	DGKA	KCNQ1	OR5W2	SLC2A1
ADA	DGKB	KCNQ2	OR6A2	SLC2A10
ADA10	DGKD	KCNQ3	OR6B1	SLC2A11
ADA12	DGKG	KCNQ4	OR6B2	SLC2A12
ADA1A	DGKH	KCNQ5	OR6B3	SLC2A13
ADA1B	DGKQ	KCNS1	OR6C1	SLC2A14
ADA1D	DGKZ	KCNS2	OR6C2	SLC2A2
ADA23	DGLA	KCNS3	OR6C3	SLC2A3
ADA28	DGLB	KCNT1	OR6C4	SLC2A4
ADA2A	DHB7	KCNT2	OR6C6	SLC2A5
ADA2B	DHCR24	KCNU1	OR6C65	SLC2A6
ADA2C	DHH	KCNV1	OR6C68	SLC2A7
ADAM10	DHPS	KCNV2	OR6C70	SLC2A8

ADAM11	DHRS7B	KCTD16	OR6C74	SLC2A9
ADAM12	DHX36	KCTD7	OR6C75	SLC30A1
ADAM15	DIAP1	KCTD8	OR6C76	SLC30A10
ADAM17	DIAPH1	KDELC2	OR6F1	SLC30A2
ADAM18	DIAPH3	KDR	OR6J1	SLC30A3
ADAM19	DIO3	KEL	OR6K2	SLC30A4
ADAM2	DIRA1	KELL	OR6K3	SLC30A5
ADAM20	DIRA2	KGP2	OR6K6	SLC30A6
ADAM21	DIRA3	KI2L1	OR6M1	SLC30A7
ADAM22	DIRAS1	KI2L2	OR6N1	SLC30A8
ADAM23	DIRAS2	KI2L3	OR6N2	SLC31A1
ADAM28	DIRAS3	KI2L4	OR6P1	SLC31A2
ADAM29	DIRC2	KI2LA	OR6Q1	SLC32A1
ADAM30	DIS3L	KI2LB	OR6S1	SLC33A1
ADAM32	DISC1	KI2S1	OR6T1	SLC34A1
ADAM33	DISP1	KI2S2	OR6V1	SLC34A2
ADAM7	DISP2	KI2S3	OR6X1	SLC34A3
ADAM8	DKFZP564O0823	KI2S4	OR6Y1	SLC35A1
ADAM9	DKK3	KI2S5	OR7A10	SLC35B1
ADAMTS15	DLEU1	KI3L1	OR7A17	SLC35B2
ADAMTS2	DLG1	KI3L2	OR7A2	SLC35B3
ADAMTSL4	DLG2	KI3L3	OR7A5	SLC35B4
ADAP1	DLG3	KI3S1	OR7AA	SLC35C2

ADAP2	DLG4	KIAA0090	OR7AH	SLC35D2
ADCK4	DLG5	KIAA0152	OR7C1	SLC35D3
ADCK5	DLGAP2	KIAA0195	OR7C2	SLC35E1
ADCY1	DLGAP4	KIAA0247	OR7D2	SLC35E4
ADCY2	DLGP1	KIAA0319	OR7D4	SLC35F2
ADCY3	DLGP2	KIAA0319L	OR7G1	SLC35F3
ADCY4	DLGP3	KIAA0494	OR7G2	SLC35F5
ADCY5	DLK1	KIAA0556	OR7G3	SLC36A1
ADCY6	DLK2	KIAA0644	OR83P	SLC36A2
ADCY7	DLL1	KIAA0746	OR8A1	SLC36A3
ADCY8	DLL3	KIAA0895	OR8B12	SLC36A4
ADCY9	DLL4	KIAA0922	OR8B2	SLC37A1
ADCYA	DLST	KIAA1024	OR8B3	SLC37A2
ADCYAP1R1	DMBT1	KIAA1026	OR8B4	SLC37A3
ADD1	DMD	KIAA1161	OR8B8	SLC37A4
ADD2	DMPK	KIAA1324	OR8BC	SLC38A1
ADD3	DMTF1	KIAA1324L	OR8D1	SLC38A10
ADDA	DMWD	KIAA1467	OR8D2	SLC38A11
ADDB	DMXL2	KIAA1522	OR8D4	SLC38A2
ADDG	DNAAF1	KIAA1524	OR8G1	SLC38A3
ADGRG2	DNAJA4	KIAA1529	OR8G2	SLC38A4
ADGRL2	DNAJB12	KIAA1549	OR8G5	SLC38A5
ADH1A	DNAJB4	KIAA1715	OR8H1	SLC38A6

ADH1B	DNAJC14	KIAA1919	OR8H2	SLC38A7
ADH1C	DNAJC15	KIAA1958	OR8H3	SLC38A9
ADH7	DNAJC22	KIAA2013	OR8I2	SLC39A1
ADI1	DNAJC25	KIBRA	OR8J1	SLC39A10
ADIPOR1	DNAJC30	KIDINS220	OR8J3	SLC39A11
ADIPOR2	DNAJC4	KIF13A	OR8K1	SLC39A12
ADORA1	DNAJC5	KIF17	OR8K3	SLC39A13
ADORA2A	DNAJC6	KIF19	OR8K5	SLC39A14
ADORA2B	DNAJC9	KIF21A	OR8S1	SLC39A2
ADORA3	DNASE1L1	KIF26A	OR8U1	SLC39A3
ADR1	DNER	KIF26B	OR8U8	SLC39A4
ADR2	DNJA3	KIF5A	OR8U9	SLC39A5
ADRA1A	DNJB4	KIF5C	OR9A1	SLC39A6
ADRA1B	DNJC5	KIR2DL1	OR9A2	SLC39A7
ADRA1D	DNJC9	KIR2DL2	OR9A4	SLC39A8
ADRA2A	DNM1	KIR2DL3	OR9G1	SLC39A9
ADRA2B	DNM2	KIR2DL4	OR9G4	SLC3A1
ADRA2C	DOC2B	KIR2DL5A	OR9G9	SLC3A2
ADRB1	DOCK1	KIR2DL5B	OR9I1	SLC40A1
ADRB2	DOCK2	KIR2DS1	OR9K2	SLC41A1
ADRB3	DOCK4	KIR2DS2	OR9Q1	SLC41A2
ADRBK1	DOCK8	KIR2DS3	OR9Q2	SLC41A3
ADRBK2	DOCK9	KIR2DS4	ORAI1	SLC43A1

ADRM1	DOK1	KIR2DS5	ORAI2	SLC43A2
ADSS	DOK2	KIR3DL1	ORAI3	SLC43A3
ADTRP	DOK3	KIR3DL2	ORC1	SLC44A1
AEBP1	DOK6	KIR3DL3	ORM1	SLC44A2
AER61	DOK7	KIR3DS1	ORMDL1	SLC44A3
AFAP1	DOLK	KIRR1	ORMDL3	SLC44A4
AFAP1L2	DPA1	KIRR2	OSBL3	SLC44A5
AG10B	DPB1	KIRR3	OSBL6	SLC45A2
AGER	DPCR1	KIRREL	OSBL7	SLC45A3
AGPAT2	DPEP1	KIRREL2	OSBP2	SLC45A4
AGPAT3	DPM1	KIRREL3	OSBPL10	SLC46A1
AGPAT4	DPM3	KISS1R	OSBPL3	SLC46A2
AGPAT5	DPP10	KISSR	OSBPL6	SLC46A3
AGPAT6	DPP4	KIT	OSBPL7	SLC47A1
AGPAT9	DPP6	KITLG	OSBPL8	SLC47A2
AGPS	DPP7	KKLC1	OSCAR	SLC4A1
AGRA2	DPP8	KL	OSCP1	SLC4A10
AGRB2	DPY19L2	KLB	OSGEP	SLC4A11
AGRB3	DPY19L3	KLC2	OSMR	SLC4A2
AGRD1	DPY19L4	KLF9	OSR2	SLC4A3
AGRE1	DPYSL2	KLH17	OSTA	SLC4A4
AGRE2	DQA1	KLHL31	OSTalpha	SLC4A5
AGRE3	DQA2	KLHL41	OSTB	SLC4A7

AGRE4	DQB1	KLK7	OSTbeta	SLC4A8
AGRF1	DQB2	KLOT	OSTM1	SLC4A9
AGRF5	DQX1	KLOTB	OTOA	SLC5A1
AGRG2	DRA	KLRA1	OTOAN	SLC5A10
AGRG3	DRAM	KLRB1	OTOF	SLC5A11
AGRIN	DRAM2	KLRC1	OTOG	SLC5A12
AGRL1	DRB3	KLRC2	OTOP1	SLC5A2
AGRL3	DRB4	KLRC3	OTOP2	SLC5A3
AGRL4	DRB5	KLRC4	OTOP3	SLC5A4
AGRN	DRD1	KLRD1	OX1R	SLC5A5
AGTR1	DRD2	KLRF1	OX2G	SLC5A6
AGTR2	DRD3	KLRF2	OX2R	SLC5A7
AGTRAP	DRD4	KLRG1	OXER1	SLC5A8
AHCYL1	DRD5	KLRG2	OXGR1	SLC5A9
AHNAK	DRP2	KLRK1	OXT	SLC6A1
AHNAK2	DSC1	KMO	OXTR	SLC6A11
AHSG	DSC2	KMT2D	OXYR	SLC6A12
AIF1	DSC3	KMT2E	P2RX1	SLC6A13
AIF1L	DSCAM	KNG1	P2RX2	SLC6A14
AIG1	DSCAML1	KNTC1	P2RX3	SLC6A15
AJAP1	DSCL1	KPB1	P2RX4	SLC6A16
AJUBA	DSEL	KPB2	P2RX5	SLC6A17
AK1	DSG1	KPBB	P2RX6	SLC6A18

AKA7A	DSG2	KPCA	P2RX7	SLC6A19
AKAP10	DSG3	KPCD	P2RY1	SLC6A2
AKAP11	DSG4	KPCD1	P2RY10	SLC6A20
AKAP12	DSP	KPCD2	P2RY11	SLC6A3
AKAP13	DST	KPCE	P2RY12	SLC6A4
AKAP2	DTBP1	KPCG	P2RY13	SLC6A5
AKAP5	DTNA	KPCT	P2RY14	SLC6A6
AKAP7	DUOX1	KPNA2	P2RY2	SLC6A7
AKR1A1	DUOX2	KPSH1	P2RY4	SLC6A8
AKR1B10	DUOXA1	KPTN	P2RY5	SLC6A9
AKR1B15	DUOXA2	KRAS	P2RY6	SLC7A1
AKT1	DUS15	KREMEN1	P2RY8	SLC7A10
AKT2	DUS1L	KREMEN2	P2Y10	SLC7A11
AKT3	DUSP15	KRIT1	P2Y11	SLC7A13
AKTIP	DUSTY	KRT1	P2Y12	SLC7A14
AL3B1	DVL1	KRT13	P2Y13	SLC7A2
ALCAM	DVL2	KRT7	P2Y14	SLC7A3
ALDH3A1	DXO	KRT75	P3C2A	SLC7A4
ALEX	DYHC2	KRT77	P3C2B	SLC7A5
ALG10	DYN2	KRTAP19-1	P3IP1	SLC7A6
ALG10B	DYNAP	KRTAP19-3	P4HA1	SLC7A7
ALG11	DYNC1H1	KRTAP25-1	P4HA2	SLC7A8
ALK	DYNC2H1	KRTCAP2	P4K2A	SLC7A9

ALKBH5	DYSF	KRTCAP3	P5I13	SLC8A1
ALOX15	DYST	KSR1	PA2G3	SLC8A2
ALOX15B	DYTN	KSYK	PA2G4	SLC8A3
ALOX5AP	E41L2	L1CAM	PACN1	SLC9A1
ALOXE3	E41L3	L3MBTL1	PACN2	SLC9A10
ALPI	EAA3	LACE1	PACN3	SLC9A11
ALPL	EBAG9	LAD1	PACR	SLC9A2
ALPP	EBI3	LAG3	PACSIN2	SLC9A3
ALPPL2	EBP	LAIR1	PAFAH1B2	SLC9A3R1
ALS2CR12	ECE1	LALBA	PAFAH2	SLC9A3R2
ALS2CR4	ECE2	LAMA1	PAG1	SLC9A4
AMAC1	ECEL1	LAMA2	PAIP1	SLC9A5
AMAC1L2	ECOP	LAMA3	PAK1	SLC9A6
AMACR	ECSCR	LAMA4	PAK4	SLC9A7
AMBP	EDA	LAMA5	PAK6	SLC9A8
AMDHD1	EDA2R	LAMB1	PALLD	SLC9A9
AMER1	EDAR	LAMB2	PALM	SLCO1A2
AMER2	EDEM3	LAMB3	PALM2	SLCO1B1
AMER3	EDNRA	LAMC1	PALM2-AKAP2	SLCO1B3
AMFR	EDNRB	LAMC2	PALM3	SLCO1C1
AMGO1	EEF1A1	LAMC3	PALMD	SLCO2A1
AMGO2	EEF1D	LAMP1	PAM	SLCO2B1

AMHR2	EEF1G	LAMP2	PANX1	SLCO3A1
AMICA1	EEF2	LAMP3	PANX2	SLCO4A1
AMIGO1	EEPD1	LAMP5	PANX3	SLCO4C1
AMIGO2	EFC1	LAMTOR1	PAP2D	SLCO5A1
AMIGO3	EFCAB1	LANC1	PAPPA	SLCO6A1
AMN	EFCAB2	LANC2	PAQR3	SLIK6
AMOT	EFCAB7	LANCL1	PAQR4	SLIT2
AMOTL1	EFEMP1	LANCL2	PAQR5	SLIT3
AMPH	EFHC1	LAP2	PAQR6	SLITRK1
AMPN	EFNA1	LAPTM4A	PAQR7	SLITRK2
AN13A	EFNA2	LAPTM4B	PAQR8	SLITRK3
AN13B	EFNA3	LAPTM5	PAQR9	SLITRK4
AN13D	EFNA4	LASP1	PAR1	SLITRK5
ANGPT1	EFNA5	LASS2	PAR2	SLITRK6
ANGPTL1	EFNB1	LASS6	PAR3	SLK
ANGPTL2	EFNB2	LAT	PAR4	SLMAP
ANGPTL4	EFNB3	LAT1	PAR6A	SMAD1
ANK1	EFR3A	LAT2	PAR6B	SMAD7
ANK2	EFR3B	LAX1	PAR6G	SMAGP
ANK3	EGF	LAYN	PARD3	SMAKA
ANKEF1	EGFR	LBN	PARD3B	SMAP1
ANKH	EHBP1	LBP	PARD6A	SMIM9
ANKIB1	EHBP1L1	LCAP	PARD6B	SMO

ANKLE2	EHD1	LCAT	PARD6G	SMOC1
ANKRA2	EHD2	LCK	PARK2	SMPD1
ANKRD13A	EHD3	LCP1	PARK7	SMPD2
ANKRD22	EHD4	LCP2	PARL	SMPD3
ANKRD35	EI24	LCT	PARM1	SMPD4
ANKRD37	EIF2B1	LCTL	PARP6	SMPDL3B
ANKRD42	EIF2B2	LDHC	PARVA	SMS2
ANKRD44	EIF5	LDLR	PARVB	SMUF1
ANKRD46	EIF5B	LDLRAD1	PARVG	SMUF2
ANKS1B	ELANE	LDLRAD3	PATE1	SN
ANM8	ELAVL4	LDLRAP1	PAWR	SNAP23
ANO1	ELF5	LECT1	PBXIP1	SNAP25
ANO10	ELFN1	LEMD2	PC11X	SNAP29
ANO2	ELFN2	LENG1	PC11Y	SNAP91
ANO3	ELMO1	LEPR	PCBD2	SNCG
ANO4	ELMO2	LEPRE1	PCBP2	SNP23
ANO5	ELMO3	LEPROT	PCBP3	SNP25
ANO6	ELN	LEPROTL1	PCD10	SNP29
ANO7	ELOVL4	LETM1	PCD12	SNPH
ANO8	ELOVL5	LETMD1	PCD15	SNRK
ANO9	ELOVL6	LFA3	PCD16	SNRPD2
ANP32A	ELOVL7	LFG2	PCD17	SNTA1
ANPEP	ELTD1	LGALS1	PCD18	SNTB1

ANPRB	EMB	LGALS13	PCD19	SNTG2
ANR27	EMCN	LGALS3	PCD20	SNX1
ANS1B	EMD	LGALS3BP	PCDA1	SNX13
ANTR1	EMID2	LGALS4	PCDA2	SNX14
ANTR2	EMILIN1	LGALS7	PCDA3	SNX15
ANTXR1	EMILIN2	LGALS8	PCDA4	SNX18
ANTXR2	EML2	LGALS9	PCDA5	SNX19
ANX13	EMP1	LGMN	PCDA6	SNX2
ANXA1	EMP2	LGR4	PCDA7	SNX20
ANXA13	EMP3	LGR5	PCDA8	SNX26
ANXA2	EMR1	LGR6	PCDA9	SNX5
ANXA5	EMR2	LGTN	PCDAA	SNX6
ANXA6	EMR3	LHCGR	PCDAB	SNX9
ANXA7	EN1	LHFP	PCDAC	SO1A2
ANXA8L2	EN113	LHFPL1	PCDAD	SO1B1
AOC2	EN2	LHFPL2	PCDB1	SO1B3
AOC3	ENAH	LHFPL3	PCDB2	SO1B7
AP180	ENG	LHFPL4	PCDB3	SO1C1
AP1G1	ENH1	LHFPL5	PCDB4	SO2A1
AP2A1	ENH3	LIFR	PCDB5	SO2B1
AP2A2	ENK13	LIGO1	PCDB6	SO3A1
AP2B1	ENK18	LILRA1	PCDB7	SO4A1
AP2M1	ENK19	LILRA4	PCDB8	SO4C1

AP2S1	ENK21	LILRA5	PCDB9	SO5A1
AP3B1	ENK24	LILRA6	PCDBA	SO6A1
APBA1	ENK25	LILRB1	PCDBB	SOAT2
APBA2	ENK6	LILRB2	PCDBC	SOCS5
APBA3	ENK7	LILRB3	PCDBD	SOCS7
APBB1	ENK8	LILRB4	PCDBE	SOD1
APBB1IP	ENK9	LILRB5	PCDBF	SORBS1
APC	ENKD1	LIM2	PCDBI	SORBS2
APCD1	ENO1	LIMA1	PCDC1	SORBS3
APCDD1	ENO2	LIMD1	PCDC2	SORCS1
APH1B	ENO3	LIME1	PCDG1	SORCS2
APJ	ENOA	LIMS1	PCDG2	SORCS3
APLD1	ENOG	LIMS2	PCDG3	SORD
APLF	ENOX1	LIN7A	PCDG4	SORL1
APLNR	ENOX2	LIN7B	PCDG5	SORT
APLP1	ENPEP	LIN7C	PCDG6	SORT1
APLP2	ENPP1	LINGO1	PCDG7	SOS1
APMAP	ENPP2	LINGO2	PCDG8	SPA17
APOB	ENPP3	LINGO4	PCDG9	SPACA1
APOBR	ENPP4	LIPA	PCDGA	SPACA4
APOD	ENPP5	LIPC	PCDGB	SPAG1
APOH	ENPP6	LIPG	PCDGC	SPAG16
APOL1	ENPP7	LIPH	PCDGD	SPAG9

APOL2	ENTP2	LIPI	PCDGE	SPAM1
APOL4	ENTP8	LIPL	PCDGF	SPANXB1
APOL6	ENTPD1	LIPS	PCDGG	SPATA12
APOLD1	ENTPD2	LIRA5	PCDGH	SPATA9
APOM	ENTPD3	LIRB1	PCDGI	SPATC1L
APP	ENTPD6	LIRB3	PCDGJ	SPEM1
APR3	ENTPD7	LIRB4	PCDGK	SPERT
AQP1	ENTPD8	LLGL1	PCDGL	SPG11
AQP10	ENVT1	LMAN2	PCDGM	SPG20
AQP11	EP15R	LMAN2L	PCDH1	SPHK1
AQP2	EPB41	LMBR1	PCDH10	SPINK13
AQP3	EPB41L1	LMBR1L	PCDH11X	SPINK2
AQP4	EPB41L2	LMBRD1	PCDH11Y	SPINT1
AQP5	EPB41L3	LMBRD2	PCDH12	SPINT2
AQP6	EPB41L4A	LMBRL	PCDH15	SPIR1
AQP7	EPB41L4B	LMCD1	PCDH17	SPIR2
AQP8	EPB41L5	LMF1	PCDH18	SPN
AQP9	EPB42	LMF2	PCDH19	SPNS1
AR	EPB49	LMLN	PCDH20	SPNS3
AR13B	EPCAM	LMNA	PCDH21	SPP2B
ARAID	EPDR1	LMTK2	PCDH7	SPPL2A
ARAP1	EPGN	LNPEP	PCDH8	SPPL2B
ARAP3	EPHA1	LNX2	PCDH9	SPPL3

ARBK1	EPHA10	LOC202459	PCDHA1	SPRE1
ARC	EPHA2	LOC203547	PCDHA10	SPRED2
AREG	EPHA3	LOC283874	PCDHA11	SPRED3
ARF1	EPHA4	LOC347487	PCDHA12	SPRN
ARF3	EPHA5	LOC389791	PCDHA13	SPRR1A
ARF4	EPHA7	LOC389832	PCDHA2	SPRR3
ARF5	EPHA8	LOC400464	PCDHA3	SPRY1
ARF6	EPHAA	LOC401152	PCDHA4	SPRY2
ARFRP1	EPHB1	LOC441294	PCDHA5	SPRY3
ARG1	EPHB2	LOC441426	PCDHA6	SPS1
ARG28	EPHB3	LOC57228	PCDHA7	SPSB4
ARG39	EPHB4	LOC650293	PCDHA8	SPT13
ARHG2	EPHB6	LOC92270	PCDHA9	SPTA1
ARHG4	EPM2A	LOH3CR2A	PCDHAC1	SPTAN1
ARHGAP18	EPN1	LONRF1	PCDHAC2	SPTB
ARHGAP19	EPOR	LONRF3	PCDHB1	SPTB2
ARHGAP21	EPR1	LOR	PCDHB10	SPTBN1
ARHGAP22	EPS15	LOX	PCDHB11	SPTBN2
ARHGAP23	EPS15L1	LOX15	PCDHB12	SPY2
ARHGAP24	EPS8	LOXL4	PCDHB13	SPY4
ARHGAP29	EPS8L2	LPAR1	PCDHB14	SQLE
ARHGAP33	EPSTI1	LPAR2	PCDHB15	SQRDL
ARHGAP36	EPX	LPAR3	PCDHB16	SRBS1

ARHGAP40	ERAP1	LPAR4	PCDHB2	SRBS2
ARHGAP5	ERAP2	LPAR5	PCDHB3	SRC
ARHGEF1	ERB1	LPAR6	PCDHB4	SRC8
ARHGEF11	ERBB2	LPCAT1	PCDHB5	SRCIN1
ARHGEF12	ERBB2IP	LPCAT2	PCDHB6	SRCN1
ARHGEF15	ERBB3	LPCAT3	PCDHB7	SRD5A2L2
ARHGEF26	ERBB4	LPCAT4	PCDHB8	SRD5A3
ARHGEF28	ERC1	LPD6B	PCDHB9	SRGP2
ARHGEF5	EREG	LPGAT1	PCDHGA1	SRI
ARHGEF7	ERGIC2	LPH	PCDHGA10	SRPK1
ARHGEF9	ERGIC3	LPHN1	PCDHGA11	SRPRB
ARHGI	ERLIN1	LPHN2	PCDHGA12	SRPX
ARHGP	ERMAP	LPHN3	PCDHGA2	SRRM3
ARID2	ERMP1	LPIN2	PCDHGA3	SSB
ARID4A	ERO1L	LPL	PCDHGA4	SSFA2
ARID5B	ERO1LB	LPP	PCDHGA5	SSH3
ARL15	ERRFI	LPPR2	PCDHGA6	SSPN
ARL4A	ERVWE1	LPPR4	PCDHGA7	SSR1
ARL4C	ESAM	LPXN	PCDHGA8	SSR2
ARL4D	ESR1	LRAD3	PCDHGA9	SSR3
ARL6	ESYT1	LRBA	PCDHGB1	SSR4
ARL6IP1	ESYT2	LRC26	PCDHGB2	SSR5
ARL6IP5	ESYT3	LRC38	PCDHGB3	SSTR1

ARL6IP6	ETBR2	LRC4B	PCDHGB4	SSTR2
ARMC12	ETNK1	LRC4C	PCDHGB5	SSTR3
ARMCX1	EVA1A	LRC52	PCDHGB6	SSTR4
ARMCX2	EVC	LRC55	PCDHGB7	SSTR5
ARMCX3	EVC2	LRC8A	PCDHGC3	SSX2IP
ARMCX6	EVI2A	LRC8B	PCDHGC4	ST14
ARPC1B	EVI2B	LRC8C	PCDHGC5	ST17B
ARPC5	EVL	LRC8D	PCFT	ST32A
ARR3	EVPL	LRC8E	PCLO	ST3GAL1
ARRB1	EWS	LRCH1	PCMTD1	ST3GAL3
ARRB2	EWSR1	LRCH2	PCNX	ST3GAL4
ARRD3	EXOC1	LRCH4	PCYOX1	ST3GAL5
ARRD4	EXOC2	LRFN1	PD1L1	ST3GAL6
ARSB	EXOC3	LRFN2	PD1L2	ST6GAL1
ARSE	EXOC3L1	LRFN3	PD2R	ST6GALNAC3
ARSF	EXOC4	LRFN4	PD2R2	ST6GALNAC4
ARSH	EXOC7	LRFN5	PDAP1	ST6GALNAC6
ART1	EXOC8	LRIG1	PDC10	ST7
ART3	EXTL2	LRIG2	PDCD1	ST7L
ART4	EXTL3	LRIG3	PDCD1LG2	ST8SIA1
ARV1	EZR	LRIT2	PDCD7	ST8SIA2
ARVCF	EZRI	LRIT3	PDE2A	ST8SIA4
ASAH1	F11	LRMP	PDE3A	ST8SIA5

ASAH2	F11R	LRP1	PDE3B	STAB1
ASAM	F120A	LRP10	PDE4A	STAB2
ASAP1	F126B	LRP11	PDE4B	STAMBP
ASAP2	F13A1	LRP12	PDE4D	STAMBPL1
ASGR1	F168B	LRP1B	PDE6A	STAP2
ASGR2	F2	LRP2	PDE6B	STARD3
ASH1L	F2R	LRP3	PDE6C	STARD3NL
ASH2L	F2RL1	LRP4	PDE9A	STARD4
ASIC1	F2RL2	LRP5	PDGFC	STAT2
ASIC2	F2RL3	LRP6	PDGFRA	STAT3
ASIC3	F3	LRP8	PDGFRB	STB5L
ASIC5	F5	LRPAP1	PDGFRL	STBD1
ASPH	FA21A	LRPPRC	PDHX	STEA2
ASPM	FA21C	LRRC15	PDIA1	STEA4
ASPN	FA2H	LRRC17	PDIA3	STEAP1
ASPSCR1	FA57A	LRRC19	PDIA4	STEAP2
ASTL	FA65B	LRRC24	PDIA6	STEAP3
ASTN1	FAAH	LRRC25	PDLI5	STEAP4
ASTN2	FABP2	LRRC26	PDLIM1	STIL
AT10A	FABP5	LRRC3	PDLIM5	STIM1
AT10D	FADD	LRRC32	PDPK1	STIM2
AT11A	FADS1	LRRC33	PDPN	STING
AT11C	FADS2	LRRC37A	PDZD2	STIP1

AT1A1	FADS3	LRRC37A2	PDZD3	STK10
AT1A2	FADS6	LRRC37A3	PDZD8	STK17A
AT1A3	FAIM	LRRC37B	PDZK1	STML2
AT1A4	FAIM2	LRRC3B	PDZK1IP1	STML3
AT1B1	FAIM3	LRRC4	PE2R1	STOM
AT1B2	FAK1	LRRC45	PE2R2	STOML1
AT1B3	FAK2	LRRC4B	PE2R3	STOML2
AT2B1	FAM101B	LRRC4C	PE2R4	STOML3
AT2B2	FAM105A	LRRC52	PEAR1	STON1
AT2B3	FAM126A	LRRC55	PEBP1	STON2
AT2B4	FAM129A	LRRC59	PEBP4	STRA6
AT8A1	FAM129B	LRRC66	PECA1	STRC
AT8A2	FAM134A	LRRC69	PECAM1	STRN
AT8B1	FAM134B	LRRC7	PEMT	STS
AT8B2	FAM134C	LRRC8A	PEPL	STT3A
AT8B4	FAM14B	LRRC8B	PERF	STT3B
ATAD1	FAM159A	LRRC8C	PERLD1	STX12
ATCAY	FAM162A	LRRC8D	PERP	STX17
ATF4	FAM163B	LRRC8E	PF2R	STX18
ATF6B	FAM171A1	LRRFIP2	PFAS	STX1A
ATG13	FAM171A2	LRRIQ3	PFDN1	STX1B
ATG3	FAM173B	LRRN1	PFN1	STX1B1
ATG7	FAM174A	LRRN2	PGAP1	STX2

ATG9A	FAM174B	LRRN3	PGAP3	STX3
ATG9B	FAM176A	LRRN4	PGCP	STX4
ATIC	FAM176B	LRRN4CL	PGF	STX6
ATL1	FAM189B	LRRT1	PGFRA	STX7
ATL2	FAM18B	LRRT2	PGFRB	STX8
ATL3	FAM18B2	LRRT3	PGLYRP2	STXB3
ATOX1	FAM193A	LRRT4	PGLYRP4	STXB5
ATP10A	FAM20B	LRRTM1	PGM5	STXBP1
ATP10B	FAM210B	LRRTM2	PGRMC1	STXBP2
ATP10D	FAM213B	LRRTM3	PGRMC2	STXBP4
ATP11A	FAM222A	LRRTM4	PHACTR4	STXBP5
ATP11B	FAM23B	LRTM1	PHAG1	STYK1
ATP11C	FAM26E	LRTM2	PHB	SUCLG1
ATP12A	FAM26F	LSAMP	PHB2	SUCLG2
ATP13A1	FAM38A	LSHR	PHEX	SUCNR1
ATP13A2	FAM38B	LSP1	PHF20L1	SUCR1
ATP13A3	FAM45A	LSR	PHGDH	SUGT1
ATP13A4	FAM57A	LST1	PHKA1	SUIS
ATP13A5	FAM57B	LST-3TM12	PHLDB2	SULF1
ATP1A1	FAM62A	LT4R1	PHLDB3	SULF2
ATP1A2	FAM62B	LT4R2	PHLP1	SUMF1
ATP1A3	FAM62C	LTA	PHTF2	SUMO1
ATP1A4	FAM70A	LTB	PI16	SUPT6H

ATP1B1	FAM70B	LTB4R	PI2R	SUSD1
ATP1B2	FAM72A	LTB4R2	PI3	SUSD2
ATP1B3	FAM72B	LTBP1	PI3R5	SUSD3
ATP1B4	FAM72C	LTBP3	PI3R6	SUSD4
ATP2A2	FAM72D	LTBR	PI42A	SUSD5
ATP2A3	FAM73A	LTC4S	PI42B	SV2A
ATP2B1	FAM74A1	LTF	PI4K2A	SV2B
ATP2B2	FAM84B	LTK	PI4K2B	SV2C
ATP2B3	FAM89B	LTOR1	PI4KA	SVIL
ATP2B4	FAM8A1	LUM	PI4KB	SVIP
ATP2C1	FANCG	LUZP2	PI51A	SVOP
ATP2C2	FANCM	LX15B	PI51C	SVOPL
ATP4A	FAP	LY66C	PICALM	SWAP70
ATP4B	FAR1	LY66D	PICK1	SWET1
ATP5B	FAR2	LY66F	PIEZ1	SWP70
ATP6AP1	FARP1	LY6D	PIGG	SYCP2
ATP6AP2	FARS2	LY6E	PIGM	SYCY1
ATP6V0A1	FAS	LY6G5B	PIGN	SYCY2
ATP6V0A2	FASLG	LY6G6C	PIGP	SYK
ATP6V0A4	FASN	LY6G6D	PIGQ	SYMPK
ATP6V0B	FAT	LY6G6F	PIGR	SYNC
ATP6V0C	FAT1	LY6H	PIGS	SYNE2
ATP6V0E1	FAT2	LY6K	PIGT	SYNG1

ATP6V0E2	FAT3	LY75	PIGU	SYNGR1
ATP6V1A	FAT4	LY86	PIGV	SYNGR2
ATP6V1B1	FATE1	LY9	PIGW	SYNGR3
ATP6V1E1	FBLN2	LYAM2	PIGX	SYNGR4
ATP6V1F	FBLN5	LYCAT	PIGY	SYNJ2
ATP6V1G3	FBLN7	LYN	PIGZ	SYNPO
ATP7A	FBN1	LYNX1	PIK3AP1	SYNPR
ATP7B	FBN2	LYPD1	PIK3C2B	SYNRG
ATP8A1	FBP1L	LYPD2	PIK3C2G	SYP
ATP8A2	FBP2	LYPD3	PIK3CA	SYPL1
ATP8B1	FBRSL1	LYPD4	PIK3CG	SYPL2
ATP8B2	FBSP1	LYPD5	PIK3IP1	SYT1
ATP8B3	FBXL2	LYPD8	PIK3R1	SYT10
ATP8B4	FBXO16	LYSMD3	PIK3R2	SYT13
ATP9A	FBXO38	LYSMD4	PIK3R5	SYT14
ATP9B	FBXO41	LYVE1	PIKFYVE	SYT15
ATPA	FCAMR	LZTFL1	PILRA	SYT16
ATPIF1	FCAR	LZTS1	PILRB	SYT2
ATRN	FCER1A	LZTS2	PIM1	SYT3
ATRNL1	FCER1G	LZTS3	PIP4K2A	SYT4
ATRX	FCER2	M3K7	PIP4K2B	SYT5
ATXN10	FCERA	M4K2	PIP5K1A	SYT6
ATXN3	FCERG	M6PR	PIP5K1C	SYT7

ATXN7L3	FCG2A	MACF1	PITPNM3	SYT8
AVEN	FCG2B	MADCAM1	PJA2	SYT9
AVPI1	FCG2C	MADD	PJCG6	SYTL1
AVPR1A	FCG3A	MAEA	PK1L1	SYTL2
AVPR1B	FCG3B	MAG	PK1L3	SYTL4
AVPR2	FCGBP	MAGD1	PK2L1	SYVN1
AVR2B	FCGR1	MAGE1	PK3CG	T131L
AXIN1	FCGR1A	MAGED1	PKD1	T150A
AXL	FCGR1B	MAGED4	PKD1L1	T150B
AZU1	FCGR1C	MAGEH1	PKD1L2	T150C
B2M	FCGR2A	MAGI1	PKD1L3	T2R16
B3AT	FCGR2B	MAGI2	PKD2	T2R43
B3GALNT1	FCGR2C	MAGI3	PKD2L1	T2R46
B3GALT2	FCGR3A	MAGT1	PKD2L2	T4S20
B3GALTL	FCGR3B	MAL	PKDREJ	TA2R
B3GNT1	FCGRB	MAL2	PKHA1	TA2R4
B3GNT2	FCGRC	MALD1	PKHA2	TAAR1
B4GALNT2	FCGRN	MALD2	PKHD1	TAAR2
B4GALT1	FCGRT	MALL	PKHD1L1	TAAR3
B4GALT3	FCHSD2	MAMDC2	PKHG5	TAAR5
B4GALT5	FCN1	MAMDC4	PKHH2	TAAR6
B4GT1	FCRL1	MAN2A1	PKHN1	TAAR8
BACE1	FCRL2	MAN2A2	PKHO1	TAAR9

BACE2	FCRL3	MAN2B1	PKM2	TACR1
BAG3	FCRL4	MAN2B2	PKN1	TACR2
BAI1	FCRL5	MANSC1	PKP2	TACR3
BAI2	FCRL6	MAP1A	PKP3	TACSTD1
BAI3	FCRLB	MAP1B	PKP4	TACSTD2
BAIAP2	FER	MAP2K1	PKR1	TAGLN2
BAIAP2L1	FER1L3	MAP2K4	PKR2	TANC1
BAK1	FER1L5	MAP2K5	PLA2G16	TAOK2
BAMBI	FER1L6	MAP2K7	PLA2G2F	TAOK3
BANK1	FERM1	MAP3K12	PLA2G3	TAP1
BASI	FERM2	MAP3K6	PLA2G5	TAPBPL
BASP1	FERMT2	MAP3K7	PLA2R	TAPT1
BAT1	FERMT3	MAP3K7IP2	PLA2R1	TARP
BAT2L2	FES	MAP4	PLAA	TAS1R1
BAT5	FEZ1	MAP4K2	PLAT	TAS1R2
BAX	FFAR1	MAP4K5	PLAU	TAS1R3
BBOF1	FFAR2	MAP7	PLAUR	TAS2R1
BBS1	FFAR3	MAPK8IP1	PLB1	TAS2R10
BBS2	FFAR4	MAPKAP1	PLBD1	TAS2R13
BBS4	FGA	MAPT	PLBD2	TAS2R14
BBS5	FGB	MARCH10	PLCB1	TAS2R16
BBS7	FGD1	MARCH2	PLCD1	TAS2R3
BCAM	FGD2	MARCH3	PLCD3	TAS2R38

BCAN	FGD5	MARCH4	PLCD4	TAS2R39
BCAP	FGF2	MARCH6	PLCE1	TAS2R4
BCAP31	FGFP1	MARCH7	PLCH2	TAS2R40
BCAR1	FGFR1	MARCH8	PLCL2	TAS2R41
BCAR3	FGFR2	MARCH9	PLD1	TAS2R42
BCHE	FGFR3	MARCKS	PLD2	TAS2R43
BCL2L10	FGFR4	MARCO	PLD3	TAS2R44
BCL2L2	FGFRL1	MARH1	PLD4	TAS2R45
BCR	FGL2	MARK1	PLDN	TAS2R46
BDKRB1	FGR	MARK2	PLDX1	TAS2R48
BDKRB2	FHIT	MARK3	PLEC1	TAS2R49
BDNF	FHL1	MARVELD1	PLEK	TAS2R5
BEND6	FHL3	MARVELD2	PLEK2	TAS2R50
BEST1	FIBCD1	MARVELD3	PLEKHA1	TAS2R60
BEST2	FICD	MAS	PLEKHA4	TAS2R7
BEST3	FIGN	MAS1	PLEKHB1	TAS2R8
BEST4	FILIP1	MAS1L	PLEKHG3	TAS2R9
BFAR	FILIP1L	MAST1	PLEKHG6	TASP1
BFSP1	FIT1	MAST2	PLEKHH2	TAU
BFSP2	FKBP10	MBOAT2	PLEKHO1	TAZ
BGN	FKBP14	MBOAT7	PLEKHO2	TBC1D10A
BI2L2	FKBP2	MBP	PLET1	TBC1D20
BICD1	FKBP7	MBTPS2	PLIN2	TBC1D24

BIK	FKBP8	MC1R	PLIN4	TBC1D30
BIN1	FKBP9	MC2R	PLK3	TBC1D3F
BIN2	FKRP	MC3R	PLLP	TBC1D4
BIRC2	FKTN	MC4R	PLOD1	TBC1D9B
BIRC3	FLAD1	MC5R	PLOD2	TBC30
BIRC6	FLCN	MCA32	PLOD3	TBC3A
BKRB1	FLG2	MCAM	PLP1	TBC3B
BKRB2	FLJ12684	MCART6	PLP2	TBC3C
BLCAP	FLJ16124	MCC	PLPL2	TBC3D
BLK	FLJ20160	MCEMP1	PLPP	TBC3E
BLNK	FLJ20489	MCF2	PLPP1	TBC3F
BLVRB	FLJ20674	MCF2L	PLPP3	TBC3G
BMP1	FLJ21511	MCF2L2	PLPPR5	TBC3H
BMPR1A	FLJ22167	MCHR1	PLPR5	TBC3I
BMPR1B	FLJ22222	MCHR2	PLRKT	TBC3K
BMPR2	FLJ23834	MCLN1	PLS1	TBC3L
BMR1B	FLJ35424	MCOLN1	PLS3	TBCD
BMX	FLJ35773	MCOLN2	PLSCR1	TBXA2R
BNIP1	FLJ35776	MCOLN3	PLSCR2	TCAF1
BNIP3	FLJ36166	MCTP1	PLSCR4	TCAF2
BOC	FLJ38482	MCTP2	PLSL	TCAM2
BPI	FLJ39653	MDGA1	PLTP	TCHP
BRAF	FLJ40235	MDGA2	PLVAP	TCIRG1

BRI3	FLJ41841	MDR1	PLXA1	TCL1A
BRI3BP	FLJ45244	MDR3	PLXA2	TCP11
BRPF1	FLJ45422	ME1	PLXA3	TCRB
BRS3	FLJ45803	MEG10	PLXA4	TCTA
BSCL2	FLJ46321	MEG11	PLXB1	TCTEX1D4
BSG	FLNA	MEGF10	PLXB2	TCTN2
BSND	FLNB	MEGF11	PLXB3	TCTN3
BST1	FLNC	MEGF8	PLXD1	TDGF1
BST2	FLOT1	MELK	PLXDC1	TDGF3
BT3A1	FLOT2	MELT	PLXDC2	TDP1
BT3A2	FLRT1	MEN1	PLXNA1	TEC
BT3A3	FLRT2	MEP1A	PLXNA2	TECTA
BTC	FLRT3	MEP1B	PLXNA3	TECTB
BTD	FLT1	MERL	PLXNA4	TEDDM1
BTK	FLT3	MERTK	PLXNB1	TEFF1
BTLA	FLT3L	MET	PLXNB2	TEK
BTN1A1	FLT3LG	METAP1	PLXNB3	TEN1
BTN2A1	FLT4	METAP2	PLXNC1	TEN2
BTN2A2	FLTOP	METTL15	PLXND1	TES
BTN2A3	FLVC1	METTL7A	PMEPA1	TESK1
BTN3A1	FLVC2	METTL9	PMP2	TESSP5
BTN3A2	FLVCR1	MFA3L	PMP22	TEST
BTN3A3	FLVCR2	MFAP3	PMP22CD	TEX14

BTNL3	FMN1	MFAP3L	PNCK	TEX2
BTNL8	FMN2	MFAP4	PNKD	TEX261
BTNL9	FMNL1	MFAP5	PNN	TEX28
BTRC	FMNL2	MFF	PNPLA2	TEX9
BVES	FMNL3	MFGE8	PNPLA6	TF
BY55	FMR1NB	MFI2	PNPLA8	TFPI
BZW2	FN1	MFNG	PODO	TFPI2
C10ORF111	FNBP1	MFRP	PODXL	TFR1
C10ORF26	FNDC3A	MFS4B	PODXL2	TFR2
C10ORF35	FNDC3B	MFSD1	POFUT2	TFRC
C10ORF39	FNDC4	MFSD10	POK9	TGBR3
C10ORF54	FNDC5	MFSD11	POLE	TGFA
C10ORF57	FOLH1	MFSD2	POMT1	TGFB1
C10ORF72	FOLR1	MFSD2A	POMT2	TGFBR1
C10ORF76	FOLR2	MFSD3	PON1	TGFBR2
C11ORF10	FOLR3	MFSD4	PON2	TGFBR3
C11ORF24	FOXN3	MFSD5	POPD1	TGFR1
C11ORF30	FPR1	MFSD6	POPD2	TGFR2
C11ORF75	FPR2	MFSD7	POPDC3	TGM2
C11ORF87	FPR3	MFSD8	POR	TGOLN2
C12ORF23	FR1L5	MFSD9	PORCN	TGON2
C12ORF49	FRAG1	MGA	POSTN	THBD
C12ORF53	FRAS1	MGAM	POTEB	THBS1

C12ORF63	FREM2	MGAT4B	POTED	THBS2
C12ORF67	FRITZ	MGAT5	POTEE	THBS3
C12ORF69	FRMD3	MGAT5B	POTEG	THEM4
C12ORF73	FRMD4B	MGC16703	POU2AF1	THOP1
C12ORF76	FRMD6	MGC34821	PP16A	THS7A
C13ORF16	FRPD1	MGC35440	PP16B	THSD1
C13ORF36	FRPD2	MGRN1	PP2BA	THSD4
C14ORF1	FRRS1	MIA3	PPAN-P2RY11	THSD7A
C14ORF100	FRS1L	MIB1	PPAP	THY1
C14ORF101	FRS2	MICA	PPAP2A	TIAM1
C14ORF132	FRS3	MICALCL	PPAP2B	TICAM1
C14ORF147	FRZB	MICB	PPAP2C	TIE1
C14ORF165	FSCN1	MIEN1	PPAPDC1A	TIE2
C14ORF180	FSCN2	MILK2	PPAPDC1B	TIGD7
C14ORF37	FSD1L	MILR1	PPAPDC2	TIGIT
C14ORF49	FSHR	MINK1	PPAPDC3	TIKI1
C14ORF83	FSTL1	MINPP1	PPB1	TIKI2
C14ORF93	FTCD	MIP	PPBI	TIMD4
C15ORF24	FUCA1	MISP	PPBN	TIMM17B
C163A	FURIN	MKLN1	PPBT	TIMM23
C163B	FUT11	MKRN3	PPFIA1	TIMP1
C16ORF54	FUT4	MKS1	PPFIA2	TIRAP

C16ORF58	FUT5	MKS3	PPFIA3	TJAP1
C16ORF74	FUT6	MLANA	PPFIBP1	TJP1
C16ORF79	FXYD1	MLC1	PPFIBP2	TJP2
C17ORF28	FXYD2	MLF2	PPHLN1	TJP3
C17ORF58	FXYD3	MLKL	PPIAL4A	TLCD1
C17ORF61	FXYD4	MLLT4	PPIB	TLL1
C17ORF62	FXYD5	MLNR	PPIL2	TLL2
C17ORF74	FXYD6	MMACHC	PPIP1	TLN1
C17ORF78	FXYD7	MMD	PPL	TLN2
C17ORF80	FYB	MMD2	PPM1A	TLR1
C18ORF1	FYN	MME	PPM1F	TLR10
C18ORF19	FZD1	MMGT1	PPM1L	TLR2
C18ORF21	FZD10	MMP10	PPP1CA	TLR3
C18ORF26	FZD2	MMP13	PPP1CB	TLR4
C18ORF45	FZD3	MMP14	PPP1R12A	TLR5
C18ORF55	FZD4	MMP15	PPP1R12B	TLR6
C19ORF12	FZD5	MMP16	PPP1R16A	TLR7
C19ORF15	FZD6	MMP17	PPP1R16B	TLR8
C19ORF18	FZD7	MMP19	PPP1R18	TLR9
C19ORF26	FZD8	MMP2	PPP1R3A	TM100
C19ORF28	FZD9	MMP24	PPP1R3F	TM102
C19ORF33	G32P1	MMP25	PPP1R9B	TM114
C19ORF41	G3BP1	MMP9	PPP2R1B	TM119

C19ORF46	G6B	MMRN1	PPP2R2B	TM11B
C19ORF48	G6PC2	MO2R1	PPP3R1	TM11D
C19ORF56	G6PC3	MOCS1	PPP3R2	TM11E
C19ORF59	GA113	MOES	PPP4C	TM127
C19ORF6	GAA	MOG	PPP5	TM198
C19ORF63	GAB1	MOGAT2	PPP5C	TM204
C19ORF71	GAB2	MOGAT3	PPP6R3	TM219
C1GALT1C1	GABARAP	MORN4	PPT1	TM231
C1ORF116	GABBR1	MOSPD1	PPT2	TM240
C1ORF130	GABBR2	MOSPD2	PPTC7	TM266
C1ORF159	GABR1	MOSPD3	PPYR1	TM2D1
C1ORF162	GABR2	MOT1	PQLC1	TM2D2
C1ORF2	GABRA1	MOT10	PQLC2	TM2D3
C1ORF210	GABRA2	MOT12	PQLC3	TM4SF1
C1ORF27	GABRA3	MOT14	PRAF1	TM4SF18
C1ORF32	GABRA4	MOT2	PRAF2	TM4SF19
C1ORF35	GABRA5	MOT3	PRAF3	TM4SF20
C1ORF71	GABRA6	MOT4	PRAME	TM4SF4
C1ORF74	GABRB1	MOT5	PRAP1	TM4SF5
C1ORF85	GABRB2	MOT6	PRAX	TM6SF1
C1ORF9	GABRB3	MOT7	PRC1	TM6SF2
C1ORF91	GABRD	MOT8	PRDX2	TM7S3
C1ORF95	GABRE	MOT9	PRDX6	TM7SF2

C1QBP	GABRG1	MOXD1	PRELP	TM7SF3
C1QTNF1	GABRG2	MPDU1	PREX1	TM7SF4
C1QTNF2	GABRG3	MPDZ	PREX2	TM9SF1
C20ORF103	GABRP	MPHOSPH9	PRF1	TM9SF2
C20ORF106	GABRQ	MPI	PRG2	TM9SF3
C20ORF107	GABRR1	MPL	PRG-3	TM9SF4
C20ORF108	GABRR2	MPO	PRICKLE3	TMBIM1
C20ORF24	GADL1	MPP1	PRIMA	TMBIM4
C20ORF3	GAGE12H	MPP2	PRIMA1	TMBIM6
C20ORF30	GAK	MPP5	PRIO	TMC1
C20ORF39	GAK10	MPP6	PRKCA	TMC2
C20ORF46	GAK19	MPPE1	PRKCB	TMC4
C20ORF54	GAK21	MPRA	PRKCD	TMC5
C20ORF59	GAK24	MPRB	PRKCDBP	TMC6
C21ORF63	GAK5	MPV17L2	PRKCG	TMC7
C22ORF32	GAK6	MPZ	PRKCH	TMC8
C2CD5	GAK7	MPZL1	PRKCI	TMCC1
C2ORF18	GAK8	MPZL2	PRKCQ	TMCC2
C2ORF28	GAK9	MPZL3	PRKCSH	TMCC3
C2ORF72	GALNAC4S-6ST	MR1	PRKCZ	TMCO1
C3	GALNS	MRAP	PRKD1	TMCO2
C3AR	GALNT1	MRAP2	PRKD2	TMCO3
C3AR1	GALNT10	MRAS	PRKD3	TMCO4

C3ORF14	GALNT11	MRC1	PRKG2	TMCO5A
C3ORF17	GALNT15	MRC1L1	PRLHR	TMED1
C3ORF18	GALNT5	MRC2	PRLR	TMED10
C3ORF35	GALNT8	MRCKB	PRND	TMED2
C3ORF38	GALR1	MREG	PRNP	TMED3
C3ORF45	GALR2	MRGPRD	PROCR	TMED4
C3ORF57	GALR3	MRGPRE	PROKR1	TMED7
C42S1	GAP43	MRGPRF	PROKR2	TMED9
C42S2	GAPDH	MRGPRX1	PROM1	TMEDA
C4A	GAPDHS	MRGPRX2	PROM2	TMEFF1
C4BPA	GAPT	MRGPRX3	PRPH2	TMEFF2
C4BPB	GAS1	MRGPRX4	PRR7	TMEM100
C4ORF22	GAS2	MRGRD	PRRG1	TMEM101
C4ORF32	GAS2L1	MRGRE	PRRG2	TMEM104
C4ORF34	GAS7	MRGRF	PRRG3	TMEM105
C5AR1	GAS8	MRGRG	PRRG4	TMEM106A
C5AR2	GASR	MRGX1	PRRT1	TMEM106B
C5ORF15	GBA	MRGX2	PRRT2	TMEM106C
C5ORF26	GBAS	MRGX3	PRRT3	TMEM107
C5ORF28	GBG1	MRGX4	PRRT4	TMEM109
C5ORF4	GBG10	MRP	PRS41	TMEM11
C5ORF40	GBG11	MRP1	PRSS21	TMEM110
C5ORF45	GBG12	MRP2	PRSS23	TMEM111

C6ORF105	GBG13	MRP6	PRSS35	TMEM115
C6ORF120	GBG2	MRP7	PRSS8	TMEM116
C6ORF129	GBG3	MRPL42	PRTG	TMEM117
C6ORF138	GBG4	MRPL44	PRTN3	TMEM119
C6ORF162	GBG5	MRS2	PRX	TMEM120A
C6ORF174	GBG7	MRVI1	PSAP	TMEM121
C6ORF191	GBG8	MS4A1	PSCA	TMEM123
C6ORF192	GBGT2	MS4A10	PSD	TMEM125
C6ORF25	GBP1	MS4A12	PSD1	TMEM126A
C6ORF35	GBRA1	MS4A13	PSD3	TMEM126B
C6ORF64	GBRA2	MS4A14	PSD4	TMEM127
C6ORF70	GBRA3	MS4A15	PSEN1	TMEM128
C6ORF72	GBRA4	MS4A2	PSEN2	TMEM129
C6ORF89	GBRA5	MS4A3	PSENEN	TMEM130
C7ORF23	GBRA6	MS4A4A	PSG1	TMEM131
C7ORF38	GBRB1	MS4A5	PSG2	TMEM132A
C7ORF42	GBRB2	MS4A6A	PSMC5	TMEM132B
C7ORF43	GBRB3	MS4A6E	PSMC6	TMEM132D
C7ORF45	GBRD	MS4A7	PSMD9	TMEM132E
C7ORF53	GBRE	MS4A8B	PSMG1	TMEM133
C8ORF37	GBRG1	MSHR	PSN1	TMEM134
C8ORF4	GBRG2	MSLN	PSTPIP1	TMEM135
C8ORF40	GBRG3	MSN	PSTPIP2	TMEM136

C99L2	GBRP	MSR1	PSYR	TMEM138
C9ORF11	GBRR1	MST1R	PTAFR	TMEM139
C9ORF123	GBRR2	MTCL1	PTBP1	TMEM140
C9ORF125	GBRR3	MTCP1	PTCH1	TMEM141
C9ORF127	GBRT	MTDH	PTCH2	TMEM143
C9ORF4	GC	MTLR	PTCHD1	TMEM144
C9ORF46	GCA	MTM1	PTCHD3	TMEM145
C9ORF5	GCC1	MTMR1	PTCHD4	TMEM146
C9ORF61	GCC2	MTMR14	PTCRA	TMEM147
C9ORF7	GCGR	MTMR3	PTDSS1	TMEM149
C9ORF71	GCLM	MTND	PTDSS2	TMEM14A
C9ORF89	GCSAM	MTNR1A	PTEN	TMEM14B
C9ORF91	GDA	MTNR1B	PTGDR	TMEM14C
CA12	GDAP1	MTOR	PTGER1	TMEM150
CA14	GDAP1L1	MTPN	PTGER2	TMEM150A
CA2	GDE1	MTR1A	PTGER3	TMEM154
CA210	GDF15	MTR1B	PTGER4	TMEM156
CA4	GDF5	MTR1L	PTGES	TMEM158
CA9	GDI2	MTUS1	PTGFR	TMEM159
CAB45	GDPD1	MUC1	PTGFRN	TMEM160
CABP1	GDPD2	MUC13	PTGIR	TMEM161A
CABP2	GDPD3	MUC15	PTGIS	TMEM161B
CABP7	GDPD4	MUC16	PTGS2	TMEM162

CABP8	GDPD5	MUC17	PTH1R	TMEM163
CAC1C	GEM	MUC20	PTH2R	TMEM164
CACB1	GEPH	MUC21	PTHB1	TMEM165
CACB2	GFOD1	MUC24	PTHD1	TMEM167A
CACHD1	GFRA1	MUC4	PTHR1	TMEM167B
CACNA1B	GFRA2	MUCEN	PTK2B	TMEM168
CACNA1C	GFRA3	MUL1	PTK6	TMEM169
CACNA1D	GFRA4	MUPCDH	PTK7	TMEM17
CACNA1E	GFRAL	MUSK	PTN2	TMEM171
CACNA1F	GGCX	MXRA5	PTN3	TMEM174
CACNA1G	GGH	MXRA8	PTN4	TMEM175
CACNA1H	GGT1	MYADM	PTOV1	TMEM176A
CACNA1I	GGT3P	MYCT1	PTP4A2	TMEM176B
CACNA1S	GGT5	MYD88	PTP4A3	TMEM178
CACNA2D1	GGT7	MYH11	PTPLA	TMEM179
CACNA2D2	GGTLC2	MYH9	PTPLAD1	TMEM179B
CACNA2D3	GHITM	MYL2	PTPLAD2	TMEM18
CACNA2D4	GHR	MYL4	PTPLB	TMEM180
CACNB1	GHRHR	MYL6B	PTPN1	TMEM181
CACNB2	GHSR	MYLIP	PTPN11	TMEM182
CACNB4	GIMAP2	MYLK	PTPN12	TMEM183A
CACNG1	GINM1	MYO10	PTPN13	TMEM183B
CACNG2	GIPC1	MYO15A	PTPN2	TMEM184B

CACNG3	GIPR	MYO1B	PTPN20A	TMEM184C
CACNG4	GIT2	MYO1C	PTPN22	TMEM185A
CACNG5	GJA1	MYO1G	PTPN3	TMEM186
CACNG6	GJA10	MYO3A	PTPN4	TMEM187
CACNG7	GJA3	MYO5B	PTPN5	TMEM188
CACNG8	GJA4	MYO5C	PTPN7	TMEM19
CAD10	GJA5	MYO6	PTPRA	TMEM190
CAD11	GJA8	MYO7A	PTPRB	TMEM194A
CAD12	GJA9	MYO9A	PTPRC	TMEM195
CAD13	GJB1	MYOF	PTPRCAP	TMEM196
CAD15	GJB2	MYOT	PTPRD	TMEM198
CAD16	GJB3	MYOTI	PTPRE	TMEM199
CAD17	GJB4	MYP0	PTPRF	TMEM2
CAD18	GJB5	MYPR	PTPRG	TMEM20
CAD19	GJB6	MYSM1	PTPRH	TMEM200A
CAD20	GJB7	MYZAP	PTPRJ	TMEM200B
CAD22	GJC1	N2DL1	PTPRK	TMEM201
CAD23	GJC2	N2DL2	PTPRM	TMEM203
CAD24	GJC3	N2DL3	PTPRN	TMEM204
CAD26	GJD2	N4BP2	PTPRN2	TMEM205
CADH1	GJD3	N6AMT2	PTPRO	TMEM206
CADH2	GJD4	NAAA	PTPRQ	TMEM207
CADH3	GLA	NAALAD2	PTPRR	TMEM209

CADH4	GLB1	NAALADL1	PTPRS	TMEM211
CADH5	GLCE	NAALADL2	PTPRT	TMEM215
CADH6	GLDN	NAB1	PTPRU	TMEM216
CADH7	GLG1	NAC1	PTPRZ	TMEM217
CADH8	GLIPR1	NAC2	PTPRZ1	TMEM218
CADH9	GLIS2	NAC3	PTRF	TMEM219
CADM1	GLO1	NAGA	PTTG1IP	TMEM22
CADM2	GLP1R	NAGLT1	PUNC	TMEM220
CADM3	GLP2R	NAGLU	PURG	TMEM222
CADM4	GLPA	NAGPA	PVR	TMEM229A
CADPS	GLPB	NAIF1	PVRIG	TMEM245
CADPS2	GLPC	NAL10	PVRL1	TMEM25
CAH2	GLR	NALCN	PVRL2	TMEM254
CAH4	GLRA1	NALD2	PVRL3	TMEM258
CAH9	GLRA2	NALDL	PVRL4	TMEM26
CAHM1	GLRA3	NAP1L2	PXDC1	TMEM266
CALB1	GLRA4	NAPG	PXDN	TMEM27
CALCR	GLRB	NAR3	PXK	TMEM30A
CALCRL	GLRX	NAR4	PXN	TMEM30B
CALD1	GLT25D1	NARFL	PYCARD	TMEM31
CALHM1	GLT25D2	NAT2	PYGL	TMEM32
CALHM2	GLT8D1	NAT8	QPCTL	TMEM33
CALM1	GLT8D2	NAT8B	QRFPR	TMEM35

CALM2	GLTPD2	NAV2	QSOX1	TMEM37
CALM3	GMFB	NBEA	QSOX2	TMEM38A
CALML3	GMIP	NCAM1	R144A	TMEM38B
CALML5	GML	NCAM2	R4RL1	TMEM39A
CALN1	GMPPB	NCEH1	R4RL2	TMEM39B
CALR	GMPS	NCF2	R7BP	TMEM40
CALRL	GNA11	NCF4	RAB11B	TMEM41A
CALU	GNA12	NCHL1	RAB11FIP1	TMEM41B
CALY	GNA13	NCK1	RAB11FIP2	TMEM42
CAMK1G	GNA14	NCKAP1L	RAB11FIP3	TMEM43
CAMK2A	GNA15	NCKAP5	RAB11FIP5	TMEM44
CAMK2D	GNAI1	NCKIPSD	RAB13	TMEM45A
CAMK2G	GNAI2	NCKP1	RAB15	TMEM45B
CAMKV	GNAI3	NCKPL	RAB17	TMEM47
CAN1	GNAL	NCL	RAB18	TMEM48
CAN2	GNAO	NCLN	RAB19	TMEM49
CANB1	GNAO1	NCMAP	RAB20	TMEM50A
CANT1	GNAQ	NCOA1	RAB23	TMEM50B
CANX	GNAS	NCR1	RAB25	TMEM51
CAP1	GNAS1	NCR2	RAB27A	TMEM52
CAP2	GNAS2	NCR3	RAB27B	TMEM53
CAPN1	GNAT1	NCR3LG1	RAB28	TMEM54
CAPN10	GNAT2	NCS1	RAB2B	TMEM55A

CAPN2	GNAZ	NCSTN	RAB31	TMEM55B
CAPNS1	GNB1	NCTR1	RAB32	TMEM56
CAPRIN1	GNB2	NCTR2	RAB33A	TMEM57
CAPS	GNB2L1	NCTR3	RAB34	TMEM59
CARD10	GNB3	NDC1	RAB35	TMEM59L
CARD11	GNB4	NDFIP1	RAB38	TMEM60
CARD14	GNG10	NDFIP2	RAB3A	TMEM61
CARL1	GNG7	NDRG1	RAB3B	TMEM62
CARL2	GNG8	NEB2	RAB3C	TMEM63A
CARL3	GNGT1	NECP1	RAB3D	TMEM63B
CASC4	GNGT2	NECP2	RAB3GAP2	TMEM63C
CASD1	GNPTAB	NEDD4	RAB43	TMEM64
CASK	GNPTG	NEDD4L	RAB44	TMEM65
CASP4	GNRHR	NEDD9	RAB4A	TMEM66
CASP7	GNRR2	NEGR1	RAB4B	TMEM67
CASQ2	GNS	NELF	RAB5A	TMEM68
CASR	GOLGA5	NELL1	RAB5B	TMEM69
CATIP	GOLIM4	NEO1	RAB5C	TMEM70
CATSPER1	GOLM1	NEP	RAB6A	TMEM71
CATSPER2	GOLP3	NETO1	RAB7L	TMEM74
CATSPER3	GOLPH3L	NETO2	RAB8A	TMEM75
CATSPER4	GOLT1A	NEU1	RAB8B	TMEM77
CATSPERB	GOLT1B	NEU3	RAB9A	TMEM79

CAV1	GOPC	NEUL1	RAB9B	TMEM8
CAV2	GP101	NEUM	RABAC1	TMEM80
CAV3	GP119	NEUR1	RABL2A	TMEM82
CBARP	GP126	NEUR3	RAC1	TMEM85
CBL	GP132	NEURL	RAC2	TMEM86A
CBLN1	GP135	NEXN	RAC3	TMEM86B
CBPD	GP139	NF2	RACK1	TMEM87A
CBPM	GP141	NFAM1	RAD	TMEM87B
CBWD1	GP142	NFASC	RAD52	TMEM88
CBWD2	GP143	NGEF	RADI	TMEM89
CBWD3	GP146	NGFR	RAET1E	TMEM8A
CBWD5	GP148	NHEDC1	RAET1G	TMEM8B
CBWD7	GP149	NHEDC2	RAET1L	TMEM9
CC2D1A	GP150	NHRF1	RAF1	TMEM92
CC50A	GP151	NHRF2	RAG1AP1	TMEM93
CC50B	GP152	NHRF3	RAGE	TMEM95
CCBE1	GP153	NHRF4	RAI2	TMEM97
CCBP2	GP156	NHS	RAI3	TMEM99
CCD78	GP157	NID2	RALA	TMEM9B
CCDC107	GP158	NIN	RALB	TMHS
CCDC108	GP160	NINJ1	RALBP1	TMIE
CCDC109A	GP161	NINJ2	RALGAPA2	TMIG2
CCDC109B	GP162	NIPA1	RALGAPB	TMIGD1

CCDC112	GP171	NIPA2	RALGDS	TMIGD2
CCDC115	GP173	NIPAL4	RALGPS2	TMM17
CCDC124	GP174	NIPSNAP3A	RALY	TMM25
CCDC126	GP176	NISCH	RAMP1	TMM59
CCDC136	GP179	NK1R	RAMP2	TMM65
CCDC144NL	GP182	NK2R	RAMP3	TMM88
CCDC15	GP183	NK3R	RANB9	TMM8B
CCDC154	GP1BA	NKAI1	RAP1A	TMM8C
CCDC24	GP1BB	NKAI2	RAP1B	TMM97
CCDC51	GP2	NKAI3	RAP1GAP	TMPRSS11D
CCDC56	GP5	NKAI4	RAP2A	TMPRSS11F
CCDC62	GP6	NKAIN1	RAP2B	TMPRSS12
CCDC69	GP9	NKAIN2	RAPGEF2	TMPRSS13
CCDC8	GPA33	NKAIN3	RAPGEF6	TMPRSS2
CCDC80	GPAA1	NKAIN4	RAPH1	TMPRSS3
CCDC88A	GPBAR	NKD1	RAPSN	TMPRSS4
CCDC90A	GPBAR1	NKD2	RARRES3	TMPRSS5
CCDC90B	GPBP1	NKG2D	RAS4B	TMPRSS6
CCDC93	GPBP1L1	NKG7	RASA1	TMPRSS9
CCDC97	GPC1	NKPD1	RASA2	TMPS2
CCG4	GPC2	NKTR	RASA3	TMPS5
CCG5	GPC3	NKX2-8	RASA4	TMPS6
CCG8	GPC4	NLGN1	RASAL1	TMPS7

CCKAR	GPC5	NLGN2	RASAL2	TMPS9
CCKBR	GPC5B	NLGN3	RASD1	TMTC1
CCL26	GPC5C	NLGN4X	RASD2	TMTC2
CCNB1	GPC5D	NLGN4Y	RASE	TMTC3
CCND3	GPC6	NLGNX	RASGRF1	TMTC4
CCNT2	GPC6A	NLGNY	RASGRF2	TMUB1
CCNY	GPER	NLN	RASGRP1	TMUB2
CCNYL1	GPER1	NLRP6	RASGRP2	TMX1
CCPG1	GPI	NLRX1	RASGRP3	TMX3
CCR1	GPM6A	NLS1	RASGRP4	TMX4
CCR10	GPM6B	NMB	RASH	TN13B
CCR2	GPNMB	NMBR	RASK	TNC
CCR3	GPR1	NMD3A	RASL2	TNF
CCR4	GPR101	NMD3B	RASM	TNF11
CCR5	GPR103	NMDE1	RASN	TNF12
CCR6	GPR107	NMDE2	RASSF3	TNF14
CCR7	GPR109A	NMDE3	RASSF5	TNF18
CCR8	GPR109B	NMDE4	RB11A	TNFA
CCR9	GPR110	NMDZ1	RB22A	TNFL6
CCRL1	GPR111	NME2	RB33A	TNFRSF10A
CCRL2	GPR112	NMT1	RB39A	TNFRSF10B
CCRN4L	GPR113	NMT2	RB39B	TNFRSF10C
CCT3	GPR114	NMUR1	RB40A	TNFRSF10D

CCT4	GPR115	NMUR2	RB40B	TNFRSF11A
CD101	GPR116	NOD1	RB40C	TNFRSF11B
CD109	GPR119	NOD2	RBNS5	TNFRSF12A
CD14	GPR12	NOMO3	RC3H1	TNFRSF13B
CD151	GPR120	NOP56	RC3H2	TNFRSF13C
CD160	GPR123	NOPE	RCCD1	TNFRSF14
CD163	GPR124	NOS1	RCE1	TNFRSF17
CD163L1	GPR125	NOS2	RCN1	TNFRSF18
CD164	GPR126	NOS3	RCN3	TNFRSF19
CD164L2	GPR128	NOSTN	RDH10	TNFRSF1A
CD166	GPR132	NOSTRIN	RDH12	TNFRSF1B
CD177	GPR133	NOTC1	RDH8	TNFRSF21
CD180	GPR135	NOTC2	RDX	TNFRSF25
CD19	GPR137	NOTC3	RECK	TNFRSF4
CD1A	GPR137B	NOTC4	REEP1	TNFRSF8
CD1B	GPR139	NOTCH1	REEP2	TNFRSF9
CD1C	GPR141	NOTCH2	REEP3	TNFSF10
CD1D	GPR142	NOTCH3	REEP4	TNFSF11
CD1E	GPR143	NOTCH4	REEP5	TNFSF12
CD2	GPR144	NOX1	REEP6	TNFSF12-TNFSF13
CD20	GPR146	NOX3	RELL1	TNFSF13
CD200	GPR148	NOX4	RELL2	TNFSF13B

CD200R1	GPR149	NOX5	RELN	TNFSF14
CD200R2	GPR15	NOXA1	RELT	TNFSF15
CD207	GPR150	NOXO1	REM2	TNFSF18
CD209	GPR151	NPAL1	REPS1	TNFSF4
CD22	GPR152	NPAL2	RERG	TNFSF8
CD226	GPR153	NPAL3	RET	TNFSF9
CD24	GPR155	NPBW1	RET1G	TNMD
CD244	GPR156	NPBW2	RET1L	TNR11
CD247	GPR157	NPBWR1	RFFL	TNR17
CD248	GPR158	NPBWR2	RFIP2	TNR18
CD27	GPR160	NPC1	RFNG	TNR1A
CD274	GPR161	NPC1L1	RFPL1	TNR1B
CD276	GPR162	NPCL1	RFPL2	TNR21
CD28	GPR17	NPDC1	RFPL3	TNR25
CD2AP	GPR171	NPFF	RFT1	TNR5
CD300A	GPR172A	NPFF1	RFTN1	TNR6
CD300C	GPR172B	NPFF2	RFTN2	TNR8
CD300E	GPR173	NPFFR1	RGAP1	TNS1
CD300LB	GPR174	NPFFR2	RGL4	TNS2
CD300LF	GPR175	NPHN	RGMA	TNS4
CD300LG	GPR176	NPHP1	RGMB	TNXB
CD302	GPR177	NPHS1	RGMC	TOB1
CD320	GPR179	NPHS2	RGP1	TOLLIP

CD33	GPR18	NPL	RGPS1	TOM1
CD34	GPR180	NPR1	RGPS2	TOR1AIP1
CD36	GPR182	NPR2	RGR	TOR1AIP2
CD37	GPR183	NPR3	RGRF2	TOR1B
CD38	GPR19	NPSR1	RGS1	TOR3A
CD3D	GPR20	NPT4	RGS10	TP4A1
CD3E	GPR21	NPTN	RGS13	TP4A2
CD3G	GPR22	NPTX1	RGS14	TP4A3
CD4	GPR25	NPTX2	RGS16	TP53I11
CD40	GPR26	NPTXR	RGS17	TP53I13
CD40L	GPR27	NPY1R	RGS19	TP8L3
CD40LG	GPR3	NPY2R	RGS2	TPBG
CD44	GPR31	NPY4R	RGS20	TPCN1
CD46	GPR32	NPY5R	RGS3	TPCN2
CD47	GPR33	NQO1	RGS4	TPD52L1
CD48	GPR34	NR3C1	RGS5	TPI1
CD5	GPR35	NR3L1	RGS6	TPO
CD52	GPR37	NRAC	RGS7	TPOR
CD53	GPR37L1	NRAM2	RGS8	TPP1
CD55	GPR39	NRAS	RGS9	TPP2
CD58	GPR4	NRBP1	RGS9BP	TPRN
CD59	GPR42	NRCAM	RGSL1	TPSG1
CD5R1	GPR42P	NRD1	RHAG	TPSNR

CD5R2	GPR44	NRG1	RHBDD1	TPT1
CD6	GPR45	NRG2	RHBDD2	TPTE
CD63	GPR50	NRG3	RHBDD3	TPTE2
CD68	GPR52	NRG4	RHBDF1	TR10C
CD69	GPR55	NRM	RHBDF2	TR19L
CD7	GPR56	NRN1	RHBDL1	TRAC
CD70	GPR6	NRN1L	RHBDL2	TRADD
CD72	GPR61	NRP1	RHBDL3	TRAF2
CD74	GPR62	NRP2	RHBG	TRAF3
CD79A	GPR63	NRROS	RHBL2	TRAF3IP3
CD79B	GPR64	NRSN1	RHCE	TRAF4
CD80	GPR65	NRSN2	RHCG	TRAF6
CD81	GPR68	NRX1A	RHD	TRAF7
CD82	GPR75	NRX1B	RHDF2	TRAIP
CD83	GPR77	NRXN1	RHEB	TRAK2
CD84	GPR78	NRXN2	RHES	TRAM1L1
CD86	GPR81	NRXN3	RHG10	TRAM2
CD8A	GPR82	NSD1	RHG32	TRAP1
CD8B	GPR83	NSDHL	RHO	TRAPPC13
CD8BL	GPR84	NSF	RHOA	TRAPPC4
CD9	GPR85	NSFL1C	RHOB	TRAT1
CD93	GPR87	NSMA2	RHOBTB2	TRBC1
CD96	GPR88	NSMAF	RHOC	TRBV12-3

CD97	GPR89B	NSMF	RHOD	TRDN
CD99	GPR97	NT5E	RHOF	TREA
CD99L2	GPR98	NTAL	RHOG	TREH
CDAN1	GPRC5A	NTCP4	RHOH	TREM1
CDC42	GPRC5B	NTM	RHOJ	TREM2
CDC42EP1	GPRC5C	NTN3	RHOQ	TREML1
CDC42EP2	GPRC5D	NTNG1	RHOT1	TREML2
CDC42EP4	GPRC6A	NTNG2	RHOT2	TRFM
CDC42EP5	GPRIN1	NTR1	RHOU	TRFR
CDC42SE1	GPSM1	NTR2	RHOV	TRGV9
CDC42SE2	GPSN2	NTRI	RIC3	TRHDE
CDCP1	GPVI	NTRK1	RIC8A	TRHR
CDH1	GPX5	NTRK2	RIC8B	TRI72
CDH10	GRAMD1A	NTRK3	RILPL1	TRIB1
CDH11	GRAMD1C	NTSR1	RIMB2	TRIM13
CDH12	GRAMD2	NTSR2	RIMS1	TRIM16
CDH13	GRAMD3	NUCB2	RIMS2	TRIM16L
CDH15	GRAMD4	NUDCD3	RIN1	TRIM21
CDH16	GRASP	NUDT16L1	RIOK2	TRIM38
CDH17	GRB14	NUMB	RIOK3	TRIM4
CDH18	GRB2	NUP210	RIPK1	TRIM44
CDH19	GRB7	NUP35	RIPK3	TRIM71
CDH2	GRD2I	NUP58	RIT1	TRIML2

CDH20	GRDN	NUS1	RIT2	TRIP10
CDH22	GREM1	NXF2	RL3R1	TRIP6
CDH23	GRIA1	NXPH4	RL3R2	TRML1
CDH24	GRIA2	NYX	RNASE2	TRML2
CDH26	GRIA3	O10A2	RNASEK	TRMT1
CDH3	GRIA4	O10A3	RND1	TRMT10B
CDH4	GRID1	O10A4	RND3	TRMT13
CDH5	GRID2	O10A5	RNF103	TRO
CDH6	GRIK1	O10A6	RNF112	TROVE2
CDH7	GRIK2	O10A7	RNF114	TRPA1
CDH8	GRIK3	O10AC	RNF121	TRPC1
CDH9	GRIK4	O10AD	RNF122	TRPC3
CDHR1	GRIK5	O10AG	RNF128	TRPC4
CDHR2	GRIN1	O10C1	RNF13	TRPC5
CDHR3	GRIN2A	O10D3	RNF130	TRPC6
CDHR5	GRIN2B	O10D4	RNF133	TRPC7
CDIPT	GRIN2C	O10G2	RNF139	TRPM1
CDK14	GRIN2D	O10G3	RNF144B	TRPM2
CDK16	GRIN3A	O10G4	RNF145	TRPM3
CDK5	GRIN3B	O10G6	RNF146	TRPM4
CDK5R1	GRINA	O10G7	RNF148	TRPM5
CDK5R2	GRIP1	O10G8	RNF149	TRPM6
CDKAL1	GRK5	O10G9	RNF150	TRPM7

CDKN1B	GRK6	O10H1	RNF152	TRPM8
CDKN2A	GRK7	O10H2	RNF165	TRPV1
CDON	GRM1	O10H3	RNF166	TRPV2
CDR1	GRM2	O10H4	RNF167	TRPV3
CDSN	GRM3	O10H5	RNF170	TRPV4
CDV3	GRM4	O10J1	RNF175	TRPV5
CEACAM1	GRM5	O10J3	RNF182	TRPV6
CEACAM16	GRM6	O10J4	RNF183	TS1R1
CEACAM19	GRM7	O10J5	RNF185	TS1R2
CEACAM21	GRM8	O10J6	RNF186	TS1R3
CEACAM3	GRN	O10K1	RNF19A	TSC2
CEACAM4	GRP1	O10K2	RNF19B	TSCOT
CEACAM5	GRP2	O10P1	RNF217	TSG101
CEACAM6	GRP4	O10Q1	RNF26	TSHR
CEACAM7	GRPR	O10R2	RNF34	TSN12
CEACAM8	GSG1	O10S1	RNF43	TSNARE1
CEAM1	GSG1L	O10T2	RNF5	TSPAN1
CEAM5	GSK3B	O10V1	RNFT1	TSPAN10
CEAM6	GSR	O10W1	RNFT2	TSPAN12
CEAM7	GSTO1	O10X1	RNPEP	TSPAN13
CEAM8	GTR1	O10Z1	ROBO1	TSPAN14
CECR6	GTR11	O11A1	ROBO2	TSPAN15
CELR1	GTR3	O11G2	ROBO3	TSPAN16

CELR2	GTR4	O11H1	ROBO4	TSPAN17
CELR3	GTR5	O11H2	ROCK1	TSPAN18
CELSR1	GTR6	O11H4	ROCK2	TSPAN2
CELSR2	GTR8	O11H6	ROM1	TSPAN3
CELSR3	GTR9	O11H7	ROR1	TSPAN31
CEMIP	GTSE1	O11HC	ROR2	TSPAN32
CEND1	GUC1B	O11L1	ROS1	TSPAN33
CEP162	GUC2C	O12D1	RP11-192H23.4	TSPAN4
CEP295	GUCY1B2	O12D2	RP11-298I3.5	TSPAN5
CEP55	GUCY2C	O12D3	RP11-385D13.1	TSPAN6
CEP57L1	GUCY2D	O13A1	RP1BL	TSPAN7
CEP89	GUCY2F	O13C2	RP2	TSPAN8
CEPT1	GUSB	O13C3	RP5-862P8.2	TSPAN9
CERCAM	GXYLT1	O13C4	RP5-877J2.1	TTC17
CERK	GYPA	O13C5	RPC9	TTC23
CES1	GYPB	O13C6	RPE65	TTC7A
CFAP157	GYPC	O13C7	RPGF2	TTC7B
CFAP97	GYPE	O13C8	RPGF6	TTC8
CFC1	GZMH	O13C9	RPIA	TTLL12
CFD	HAPLN3	O13D1	RPL17	TTLL5
CFH	HARBI1	O13F1	RPL18	TTMA

CFL1	HAS1	O13G1	RPL19	TTYH1
CFL2	HAS2	O13H1	RPL21	TTYH2
CFLAR	HAS3	O13J1	RPL27A	TTYH3
CFTR	HAUS1	O14A2	RPL28	TUB
CGN	HAUS4	O14AG	RPL29	TUBA1B
CH25H	HAVCR1	O14CZ	RPL3	TUBA4A
CHD9	HAVCR2	O14I1	RPL32	TUBB
CHGA	HBEGF	O14J1	RPL35	TUBB2C
CHIC1	HCAR1	O14K1	RPL35A	TUBB4
CHIC2	HCAR2	O14L1	RPL36A	TUBB6
CHL1	HCAR3	O2A12	RPL36A-HNRNPH2	TULP1
CHMP2A	HCFC2	O2A14	RPL36AL	TULP3
CHN1	HCK	O2A25	RPL37A	TUSC3
CHN2	HCLS1	O2AE1	RPL9	TUSC5
CHODL	HCN1	O2AG1	RPLP2	TUTLA
CHP1	HCN2	O2AG2	RPN1	TUTLB
CHP2	HCN3	O2AJ1	RPN2	TWF2
CHP3	HCN4	O2AK2	RPRM	TX101
CHPF2	HCRTR1	O2AP1	RPRML	TX1B3
CHPT1	HCRTR2	O2AT4	RPS15A	TXK
CHRFAM7A	HCST	O2T10	RPS23	TXLNA
CHRM1	HDAC8	O2T11	RPS6KB1	TXLNB

CHRM2	HDBP1	O2T12	RPSA	TXNDC10
CHRM3	HDLBP	O2T27	RRAD	TXNDC11
CHRM4	HEG1	O2T29	RRAGA	TXNDC13
CHRM5	HELZ	O2T33	RRAS	TXNDC15
CHRNA1	HENMT1	O2T34	RRAS2	TXNDC5
CHRNA10	HEPACAM	O2T35	RRH	TXNIP
CHRNA2	HEPACAM2	O4A15	RRM2B	TXNRD1
CHRNA3	HEPH	O4A16	RSAD1	TYR
CHRNA4	HEPS	O4A47	RSBN1L	TYRO3
CHRNA5	HERPUD2	O4C45	RSC1A1	TYROBP
CHRNA6	HERV-FRD	O4C46	RSCA1	TYRP1
CHRNA7	HEXA	O4F15	RSL1D1	TYW3
CHRNA9	HEXB	O4F17	RSLAA	UAP1
CHRNB1	HFE	O4F21	RSLAB	UB2D3
CHRNB2	HFE2	O51A2	RSPH3	UBA1
CHRNB3	HG2A	O51A4	RSSA	UBA3
CHRNB4	HGF	O51A7	RTN4	UBA5
CHRND	HGSNAT	O51B2	RTN4R	UBA52
CHRNE	HHAT	O51B4	RTN4RL1	UBAC1
CHRNG	HHATL	O51B5	RTN4RL2	UBAC2
CHST11	HHIP	O51B6	RTP1	UBAP1
CHST12	HHIPL1	O51D1	RTP2	UBASH3B
CHST3	HHLA2	O51E1	RTP3	UBE2A

CHST6	HIAT1	O51E2	RTP4	UBE2B
CHST7	HIATL1	O51F1	RUFY1	UBE2C
CHSY3	HIF1A	O51F2	RUSC1	UBE2D2
CHUK	HIF3A	O51G1	RUVBL1	UBE2D3
CIB1	HIGD1A	O51G2	RXFP1	UBE2D4
CIB2	HIGD1B	O51H1	RXFP2	UBE2N
CIP4	HIGD2A	O51I1	RXFP3	UBE2S
CISD2	HINT1	O51I2	RXFP4	UBE3C
CK024	HIP1R	O51J1	RYK	UBIAD1
CK2N1	HK1	O51L1	RYR1	UBL3
CKAP4	HLA-A	O51M1	RYR2	UBP14
CKAP5	HLA-B	O51Q1	RYR3	UBP6
CKLF	HLA-C	O51S1	S100A10	UBP8
CL12A	HLA-DMA	O51T1	S100A11	UBQL1
CL12B	HLA-DMB	O51V1	S100A14	UBQLN2
CLAP2	HLA-DOA	O52A1	S100A16	UBR2
CLC	HLA-DOB	O52A4	S100A2	UBTD1
CLC2A	HLA-DPA1	O52A5	S100A4	UBXN11
CLC2D	HLA-DPB1	O52B2	S100A7	UBXN8
CLC4G	HLA-DQA1	O52B4	S100A8	UCP3
CLC4M	HLA-DQA2	O52B6	S100A9	UFO
CLC5A	HLA-DQB1	O52D1	S100P	UGGT1
CLC7A	HLA-DQB2	O52E1	S10A6	UGGT2

CLCA1	HLA-DRA	O52E2	S10A8	UGT1A10
CLCA2	HLA-DRB1	O52E4	S10A9	UGT1A9
CLCA3	HLA-DRB3	O52E5	S10AC	UGT2A3
CLCA4	HLA-DRB4	O52E6	S12A3	UGT3A1
CLCC1	HLA-DRB5	O52E8	S12A6	UGT3A2
CLCKB	HLA-E	O52H1	S12A7	UGT8
CLCL1	HLA-F	O52I1	S12A9	UHRF1BP1L
CLCN1	HLA-G	O52I2	S13A3	ULA1
CLCN2	HLAH	O52J3	S17A4	ULBP1
CLCN3	HM13	O52K1	S17A5	ULBP2
CLCN4	HMCN1	O52K2	S1PR1	ULBP3
CLCN5	HMGB1	O52L1	S1PR2	UMOD
CLCN6	HMGCS1	O52L2	S1PR3	UMODL1
CLCN7	HMHA1	O52M1	S1PR4	UN13A
CLCNKA	HMMR	O52N1	S1PR5	UN13B
CLCNKB	HMR1	O52N2	S20A2	UN13C
CLD1	HN1L	O52N4	S22A1	UN93A
CLD10	HNRNPK	O52N5	S22A6	UNC119
CLD11	HNRNPM	O52P1	S22A7	UNC13A
CLD12	HNRNPU	O52R1	S22A8	UNC13B
CLD14	HOME1	O52W1	S22A9	UNC13D
CLD15	HOME2	O52Z1	S22AB	UNC50
CLD16	HOME3	O56A1	S22AC	UNC5A

CLD17	HOMER3	O56A3	S22AG	UNC5B
CLD18	HP	O56A4	S22AH	UNC5C
CLD19	HPCAL1	O56A5	S22AI	UNC5D
CLD2	HPN	O56B1	S23A1	UNC84A
CLD20	HPS3	O56B2	S23A2	UNC84B
CLD22	HPSE2	O56B4	S2611	UNC93A
CLD23	HPX	O5AC1	S26A3	UNC93B1
CLD24	HRAS	O5AC2	S26A4	UNQ830
CLD25	HRASLS2	O5AK2	S26A5	UNQ870/PRO1886
CLD3	HRG	O5AK3	S26A6	UPAR
CLD34	HRH1	O5AL1	S27A1	UPK1A
CLD4	HRH2	O5AN1	S27A6	UPK1B
CLD5	HRH3	O5AP2	S28A1	UPK2
CLD6	HRH4	O5AR1	S28A3	UPK3A
CLD7	HRK	O5AS1	S29A1	UPK3B
CLD8	HRNR	O5AU1	S29A2	UQCC1
CLD9	HRSP12	O5H14	S29A4	UQCRC2
CLDN1	HS2ST1	O5H15	S35G1	UR2R
CLDN10	HS3ST2	O6C65	S35G2	UROL1
CLDN11	HS3ST3A1	O6C68	S36A1	UROM
CLDN12	HS3ST3B1	O6C70	S36A2	USE1
CLDN14	HS3ST5	O6C74	S38A1	USH1G
CLDN15	HS6ST1	O6C75	S38A2	USH2A

CLDN16	HS6ST2	O6C76	S38A3	USP19
CLDN17	HS90A	O7E24	S38A4	USP2
CLDN18	HSD11B1	OAS2	S38A5	USP20
CLDN19	HSD17B10	OAS3	S38A6	USP21
CLDN2	HSD17B12	OBSCN	S39A1	USP4
CLDN20	HSD17B2	OCA2	S39A2	USP48
CLDN23	HSD17B3	OCLM	S39A4	USP6
CLDN3	HSH2D	OCLN	S39A5	UST
CLDN4	HSP7C	ODC1	S39A6	UT1
CLDN5	HSP90AA1	ODF4	S39AB	UT2
CLDN6	HSP90AB1	ODZ1	S39AE	UTP20
CLDN7	HSP90B1	ODZ2	S40A1	UTRN
CLDN8	HSPA13	ODZ3	S41A1	UTRO
CLDN9	HSPA4	ODZ4	S41A2	UTS2
CLDND1	HSPA5	OGFOD1	S41A3	UTS2R
CLDND2	HSPA8	OGN	S47A1	V1AR
CLEC10A	HSPA9	OGR1	S47A2	V1BR
CLEC12A	HSPB1	OGT	S4A10	V2R
CLEC12B	HSPB8	OGT1	S4A11	VAMP1
CLEC14A	HSPC105	OIP5	S4A4	VAMP2
CLEC17A	HSPD1	OLFM1	S4A5	VAMP3
CLEC1A	HSPG2	OLFML2A	S4A7	VAMP5
CLEC1B	HTR1A	OLFML2B	S52A1	VAMP7

CLEC2B	HTR1B	OLFML3	S52A2	VAMP8
CLEC2D	HTR1D	OLR1	S52A3	VANG1
CLEC4A	HTR1E	OMA1	S6A13	VANG2
CLEC4D	HTR1F	OMG	S6A20	VANGL1
CLEC4E	HTR2A	OMGP	SAC	VANGL2
CLEC4F	HTR2B	OPALI	SACA4	VAPA
CLEC4G	HTR2C	OPALIN	SACM1L	VAPB
CLEC4M	HTR3A	OPCM	SAHH2	VASN
CLEC5A	HTR3B	OPCML	SAMD8	VASP
CLEC6A	HTR3C	OPHN1	SAMHD1	VAV1
CLEC7A	HTR3D	OPN1LW	SC4MOL	VCAM1
CLEC9A	HTR3E	OPN1MW	SC5A8	VCAN
CLIC1	HTR4	OPN1MW2	SC5AA	VCL
CLIC2	HTR5A	OPN1SW	SC5AC	VCPIP1
CLIC4	HTR6	OPN3	SC5D	VDAC1
CLIC5	HTR7	OPN4	SC65	VDAC2
CLIC6	HTRA1	OPN5	SC6A1	VEGFC
CLIP3	HTT	OPRD	SC6A2	VEZA
CLM1	HUNK	OPRD1	SC6A3	VEZT
CLM2	HVCN1	OPRK	SC6A4	VGFR1
CLM4	HYAL2	OPRK1	SC6A5	VGFR2
CLM6	HYALP	OPRL1	SC6A6	VGFR3
CLM7	HYCCI	OPRM	SCAM5	VIL1

CLM8	HYDIN	OPRM1	SCAMP1	VIM
CLM9	HYI	OPRS1	SCAMP2	VINC
CLMN	HYLS1	OPRX	SCAMP3	VIP1
CLMP	HYOU1	OPSG	SCAMP4	VIPR1
CLN3	I17RA	OPSG2	SCAMP5	VIPR2
CLN5	I17RB	OPSG3	SCAP	VISA
CLN6	I17RC	OR10A2	SCAR5	VISTA
CLN8	I17RD	OR10A3	SCARA3	VKORC1L1
CLNS1A	I17RE	OR10A4	SCARA5	VLDLR
CLPTM1	IBTK	OR10A5	SCARB1	VMP1
CLPTM1L	ICAM1	OR10A6	SCARB2	VN1R1
CLRN1	ICAM2	OR10A7	SCARF1	VN1R2
CLRN3	ICAM3	OR10AD1	SCARF2	VN1R3
CLSTN1	ICAM4	OR10AG1	SCCPDH	VN1R4
CLSTN2	ICAM5	OR10C1	SCEL	VN1R5
CLSTN3	ICEF1	OR10G2	SCF	VNN1
CLTA	ICOS	OR10G3	SCFD1	VNN2
CLTB	ICOSLG	OR10G4	SCGN	VNN3
CLTC	IDE	OR10G7	SCIN	VNRL4
CLTR1	IDH2	OR10G8	SCN10A	VPP2
CLTR2	IER3IP1	OR10G9	SCN11A	VPP4
CLU	IFFO2	OR10H1	SCN1A	VPREB1
CMKLR1	IFI27	OR10H2	SCN1B	VPREB3

CML1	IFI30	OR10H3	SCN2A	VPS11
CMTM1	IFI6	OR10H4	SCN2B	VPS13B
CMTM2	IFIT5	OR10H5	SCN3A	VPS26B
CMTM3	IFITM1	OR10J1	SCN3B	VPS28
CMTM4	IFITM2	OR10J3	SCN4A	VPS35
CMTM5	IFITM3	OR10J5	SCN4B	VPS36
CMTM6	IFITM5	OR10K1	SCN5A	VPS37C
CMTM7	IFM1	OR10K2	SCN7A	VPS41
CMTM8	IFM10	OR10P1	SCN8A	VPS8
CNGA1	IFM2	OR10Q1	SCN9A	VRK2
CNGA2	IFM3	OR10R2	SCNN1A	VSIG1
CNGA3	IFM5	OR10S1	SCNN1B	VSIG10
CNGA4	IFNAR1	OR10T2	SCNN1D	VSIG2
CNGB3	IFNAR2	OR10V1	SCNN1G	VSIG4
CNIH	IFNGR1	OR10W1	SCNNA	VSIG8
CNIH2	IFNGR2	OR10X1	SCNNB	VSNL1
CNIH3	IFRD1	OR10Z1	SCNND	VSTM1
CNIH4	IGDC4	OR11A1	SCNNG	VSTM3
CNKSR2	IGDCC3	OR11G2	SCOC	VSTM4
CNNM1	IGDCC4	OR11H1	SCP2	VTCN1
CNNM2	IGF1R	OR11H12	SCRB1	VTI1A
CNNM3	IGF2R	OR11H4	SCRIB	VTI1B
CNNM4	IGFBP3	OR11H6	SCTM1	VTN

CNOT1	IGFBP7	OR11L1	SCTR	VWA5B1
CNP	IGFL2	OR12D2	SCUB1	VWA7
CNR1	IGFR1	OR12D3	SCUBE1	VWF
CNR2	IGHA1	OR13A1	SCUBE3	WARS2
CNST	IGHD	OR13C2	SCYL3	WAS
CNTFR	IGHG1	OR13C3	SDC1	WASF1
CNTN1	IGHG2	OR13C4	SDC2	WASF2
CNTN2	IGHG4	OR13C5	SDC3	WBP1
CNTN3	IGHM	OR13C8	SDC4	WBSCR17
CNTN4	IGLON5	OR13C9	SDCB1	WBSCR28
CNTN5	IGS11	OR13D1	SDCBP	WDR12
CNTN6	IGSF1	OR13F1	SDE2	WDR13
CNTNAP1	IGSF11	OR13G1	SDK1	WDR53
CNTNAP2	IGSF2	OR13H1	SDK2	WDR6
CNTNAP3	IGSF3	OR13J1	SDPR	WDR81
CNTNAP4	IGSF5	OR14A16	SE6L2	WDR82
CNTNAP5	IGSF6	OR14C36	SEC16B	WFS1
CNTP3	IGSF8	OR14I1	SEC22B	WIPF1
CNTP4	IGSF9	OR14J1	SECTM1	WLS
CO8A	IHH	OR14K1	SEL1L	WNK2
CO9	IKBKB	OR1A1	SEL1L3	WNT11
COBL	IKBKG	OR1A2	SELE	WNT5A
COBLL1	IKIP	OR1B1	SELENBP1	WNT5B

CODA1	IKZF5	OR1C1	SELI	WNT8A
COF1	IL10RA	OR1D2	SELK	WTIP
COL11A1	IL10RB	OR1D4	SELL	WWP1
COL12A1	IL11RA	OR1D5	SELP	X3CL1
COL13A1	IL12RB1	OR1E1	SELPLG	XCR1
COL14A1	IL12RB2	OR1E2	SEM4A	XG
COL16A1	IL13	OR1E3	SEM4C	XIAP
COL17A1	IL13RA1	OR1F1	SEM4D	XIRP1
COL1A1	IL13RA2	OR1F2	SEM4G	XK
COL1A2	IL16	OR1FC	SEM6A	XKR2
COL23A1	IL17D	OR1G1	SEM6B	XKR3
COL25A1	IL17RA	OR1I1	SEM6C	XKR4
COL29A1	IL17RB	OR1J1	SEM6D	XKR5
COL2A1	IL17RC	OR1J2	SEM7A	XKR6
COL3A1	IL17RD	OR1J4	SEMA3A	XKR7
COL4A3BP	IL17RE	OR1K1	SEMA3C	XKR8
COL5A1	IL18R1	OR1L1	SEMA3D	XKR9
COL5A2	IL18RAP	OR1L3	SEMA3F	XKRX
COL5A3	IL1A	OR1L4	SEMA4A	XKRY
COL6A1	IL1AP	OR1L6	SEMA4B	XKRY2
COL6A2	IL1B	OR1L8	SEMA4C	XPC
COL6A3	IL1R1	OR1M1	SEMA4D	XPNPEP2
COL6A6	IL1R2	OR1N1	SEMA4F	XPO6

COL7A1	IL1RAP	OR1N2	SEMA4G	XPP2
COLEC12	IL1RAPL1	OR1P1	SEMA5A	XPR1
COLQ	IL1RAPL2	OR1Q1	SEMA5B	XRCC6
COMMD4	IL1RL1	OR1S1	SEMA6A	XRCC6BP1
COMT	IL1RL2	OR1S2	SEMA6B	XRN1
COMTD1	IL20RA	OR2A1	SEMA6C	XRP2
CONA1	IL20RB	OR2A12	SEMA6D	XYLT2
COPB	IL21R	OR2A14	SEMA7A	YAP1
COPRS	IL22RA1	OR2A2	SEPR	YBX1
COPS5	IL23R	OR2A25	SEPT3	YES
COPT1	IL26	OR2A4	SEPT8	YES1
COQ2	IL27RA	OR2A42	SERC3	YIF1A
COQ7	IL28RA	OR2A5	SERC5	YIF1B
COR1C	IL2RA	OR2A7	SERF1A	YIPF1
CORIN	IL2RB	OR2AE1	SERF1B	YIPF2
CORO1B	IL2RG	OR2AG1	SERINC1	YIPF3
CORO2A	IL31R	OR2AG2	SERINC2	YIPF4
CORO2B	IL31RA	OR2AK2	SERINC3	YIPF5
COX6B2	IL3RA	OR2AT4	SERINC4	YIPF6
COX8C	IL4R	OR2B11	SERINC5	YLAT1
CP	IL4RA	OR2B2	SERP2	YLAT2
CP2W1	IL5RA	OR2B3	SERPINA1	YOD1
CPD	IL6R	OR2B3P	SERPINA3	YTHDC1

CPE	IL6RA	OR2B6	SERPINA5	YWHAQ
CPEB1	IL6RB	OR2B8	SERPINB2	YWHAZ
CPLX1	IL6ST	OR2BB	SERPINB3	ZACN
CPLX2	IL7R	OR2C1	SERPINB5	ZAN
CPM	IL7RA	OR2C3	SERPINE2	ZAP70
CPMD8	IL8RA	OR2D2	SERPINF1	ZBED3
CPNE1	IL8RB	OR2D3	SERPING1	ZBTB33
CPNE2	IL9R	OR2F1	SERPINH1	ZBTB8OS
CPNE3	ILDR1	OR2F2	SERPINI1	ZC3HAV1
CPNE6	ILK	OR2G2	SET	ZC4H2
CPNE7	ILRL1	OR2G3	SEZ6	ZCCHC16
CPNS1	IMP5	OR2G6	SEZ6L	ZCCHC5
CPNS2	IMPAD1	OR2H1	SEZ6L2	ZDHC5
CPPED1	IMPG1	OR2H2	SFRP4	ZDHHC1
CPTP	IMPG2	OR2I1	SFRS17A	ZDHHC12
CPVL	INADL	OR2J1	SFRS3	ZDHHC13
CPZ	INAR1	OR2J2	SFSWAP	ZDHHC14
CR1	INAR2	OR2J3	SFT2D1	ZDHHC15
CR2	INGR1	OR2K2	SFT2D2	ZDHHC16
CRB1	INGR2	OR2L13	SFT2D3	ZDHHC18
CRB2	INHBE	OR2L2	SFTPC	ZDHHC19
CRB3	INO80B	OR2L3	SFXN1	ZDHHC2
CRCM	INP5E	OR2L5	SFXN3	ZDHHC20

CRCM1	INPP4A	OR2L8	SFXN4	ZDHHC21
CREB5	INPP5A	OR2LD	SGCA	ZDHHC22
CREG1	INPP5B	OR2M2	SGCB	ZDHHC23
CRELD1	INPP5J	OR2M3	SGCD	ZDHHC24
CRELD2	INPP5K	OR2M4	SGCE	ZDHHC4
CRFR1	INPPL1	OR2M5	SGCG	ZDHHC5
CRFR2	INSIG1	OR2M7	SGCZ	ZDHHC6
CRHBP	INSIG2	OR2S1	SGK1	ZDHHC7
CRHR1	INSR	OR2S2	SGK196	ZDHHC8
CRHR2	INSRR	OR2T1	SGK494	ZDHHC9
CRIM1	INTS12	OR2T10	SGMR1	ZFP90
CRIP2	INVS	OR2T11	SGMS1	ZFPL1
CRK	IOD3	OR2T12	SGMS2	ZFY27
CRLF1	IPCEF1	OR2T2	SGPL1	ZFYVE27
CRLF2	IPRI	OR2T27	SGPP1	ZFYVE28
CRLF3	IQCF1	OR2T29	SGPP2	ZGPAT
CRLS1	IQGA1	OR2T3	SGSM3	ZMY19
CRNN	IQGAP1	OR2T33	SH2B2	ZMYND10
CROCC	IQGAP2	OR2T34	SH2B3	ZMYND19
CRTAM	IRAK4	OR2T35	SH2D1A	ZNF10
CRTAP	IRGM	OR2T4	SH2D5	ZNF18
CRTC1	IRPL1	OR2T5	SH3BGR	ZNF185
CRUM1	IRS4	OR2T6	SH3BGRL3	ZNF2

CRUM2	ISLR	OR2T7	SH3D19	ZNF20
CRUM3	ISLR2	OR2T8	SH3D21	ZNF227
CRYAB	ITA3	OR2V1	SH3PXD2B	ZNF229
CRYL1	ITA6	OR2V2	SHAN1	ZNF276
CSEN	ITA8	OR2W1	SHAN2	ZNF277
CSF1	ITB1	OR2W3	SHAN3	ZNF337
CSF1R	ITB3	OR2W5	SHANK1	ZNF397
CSF2R	ITB4	OR2W6	SHARPIN	ZNF407
CSF2RA	ITBP1	OR2Y1	SHB	ZNF503
CSF2RB	ITCH	OR2Z1	SHC1	ZNF510
CSF3R	ITFG1	OR3A1	SHC4	ZNF571
CSGALNACT2	ITFG3	OR3A2	SHH	ZNF598
CSK	ITGA1	OR3A3	SHIP1	ZNF599
CSKP	ITGA10	OR3A4	SHISA2	ZNF620
CSMD1	ITGA11	OR4A15	SHISA4	ZNF646
CSMD2	ITGA2	OR4A16	SHRM2	ZNF655
CSMD3	ITGA2B	OR4A4	SHROOM2	ZNF688
CSNK1D	ITGA3	OR4A47	SHROOM3	ZNF706
CSNK1G2	ITGA4	OR4A5	SHSA9	ZNF717
CSNK2A1	ITGA5	OR4A8	SHTN1	ZNF783
CSNK2A2	ITGA6	OR4B1	SI	ZNF785
CSPG4	ITGA7	OR4C11	SI1L1	ZNF79
CSPG5	ITGA8	OR4C12	SI1L3	ZNF839

CST6	ITGA9	OR4C13	SIDT1	ZNF844
CSTN1	ITGAD	OR4C15	SIDT2	ZNRF2
CSTN2	ITGAE	OR4C16	SIG10	ZNRF3
CSTN3	ITGAL	OR4C3	SIG14	ZNRF4
CTAG2	ITGAM	OR4C45	SIGIRR	ZNT1
CTAGE5	ITGAV	OR4C46	SIGL6	ZNT10
CTL1	ITGAX	OR4C5	SIGLEC1	ZNT8
CTLA4	ITGB1	OR4C6	SIGLEC10	ZO1
CTNA1	ITGB1BP2	OR4CB	SIGLEC11	ZO2
CTNA2	ITGB2	OR4CC	SIGLEC12	ZO3
CTNB1	ITGB3	OR4CD	SIGLEC15	ZP1
CTND1	ITGB4	OR4CF	SIGLEC5	ZP2
CTNL1	ITGB5	OR4CG	SIGLEC6	ZP3
CTNNA1	ITGB6	OR4D1	SIGLEC7	ZP4
CTNNA3	ITGB7	OR4D10	SIGLEC8	ZPLD1
CTNNB1	ITGB8	OR4D11	SIGLEC9	ZSCAN22
CTNNBIP1	ITIH2	OR4D2	SIGMAR1	ZSCAN29
CTNND1	ITK	OR4D5	SIL1	ZSWIM1
CTNND2	ITLN1	OR4D6	SILV	ZSWIM8
CTNS	ITM2A	OR4D9	SIM23	ZYX
CTR1	ITM2B	OR4DA	SIN1	-