AN INVESTIGATION OF PARAMETERS THAT INFLUENCE NON-HEPATOCYTE RNA DELIVERY IN VIVO

A Dissertation
Presented to
The Academic Faculty

by

Kalina Paunovska

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Biomedical Engineering in the
Wallace H. Coulter Department of Biomedical Engineering

Georgia Institute of Technology and Emory University
August 2020

AN INVESTIGATION OF PARAMETERS THAT INFLUENCE NON-HEPATOCYTE RNA DELIVERY IN VIVO

Approved by:

Dr. James E. Dahlman, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Andres J. Garcia
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Philip J. Santangelo
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Julie A. Champion
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Edward A. Botchwey
Department of Biomedical Engineering
Georgia Institute of Technology

Date Approved: May 6th, 2020

This thesis is dedicated to my grandmother, Katja Paunovska, who was a constant source of inspiration, knowledge, and support.

ACKNOWLEDGEMENTS

I attribute my scientific curiosity at a young age to my parents, to whom I owe the most. You taught me that the only real limits I would face were the ones I placed on myself and you never had any doubt in the things I wanted to pursue and the challenges I wanted to conquer, you only ever asked me how you could help. Thank you for your constant source of unfaltering support, love, and strength.

One of the best parts of working in science is the people that you get to meet and work with. I have been lucky to be mentored Dr. James Dahlman, who took a chance on me early on. Thank you for your support, positive attitude, and guidance throughout my PhD. I would also like to thank my committee members for their mentorship and advisement during my PhD. That being said, I owe the day-to-day conversations, excitement, and curiosity to the people that I was privileged to work with closely. Specifically, I want to thank Alejandro Da Silva Sanchez, Melissa Lokugamage, Marine Hatit, and Curtis Dobrowolski for all of their help with the science, but most importantly for their friendship over the last few years of my PhD.

Finally, I would like to thank Anna Bartoli, Abby Paulson, and Chris Douglas who have stayed by my side throughout this endeavor. Anna – thank you for your friendship and for always being there to listen to my stories. Abby – thank you for your encouragement and optimism. And last, but certainly not least, thank you, Chris for being the most patient person that I know. You have been a constant source of happiness, conversation, and reason throughout my PhD and for that I will always be grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AAV	Adeno-associated virus
AGO2	Argonaute 2
ASO	Antisense oligonucleotide
BSA	Bovine serum albumin
DC	Dendritic cell
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DOPE	1,2-dioleoyl-*sn*-glycero-3-phosphoethanoloamine
dsRNA	Double stranded RNA
EC	Endothelial cell
ECGS	Endothelial cell growth supplement
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FIND	Fast Identification of Nanoparticle Delivery
GFP	Green fluorescent protein
GO	Gene ontology
HAEC	Human aortic endothelial cell
HAVEC	Human aortic vein endothelial cell
HDL	High density lipoprotein
HEK	Human embryonic kidney
HUVEC	Human umbilical vein endothelial cell
IACUC	Institutional Animal Care and Use Committee

ICAM2	Intracellular adhesion molecule 2
iMAEC	Immortalized mouse aortic endothelial cells
Indel	Insertion/deletion
JORDAN	Joint Rapid DNA Analysis of Nanoparticles
KEGG	Kyoto encyclopedia of genes and genomes
L2K	Lipofectamine 2000
LDL	Low density lipoprotein
LDLR	Low density lipoprotein receptor
LNP	Lipid Nanoparticle
LRP-1	Lipoprotein-related protein-1
LSL	Lox-stop-lox
MAPK	Mitogen-activated protein kinase
MEMNEAA	MEM non-essential amino acid solution
MFI	Mean fluorescent intensity
miRNA	Micro-RNA
mRNA	Messenger RNA
MW	Molecular weight
NGS	Next generation sequencing
NHP	Non-human primate
Ox-LDL	Oxidized low density lipoprotein
OxPhos	Oxidative phosphorylation
PCA	Principal component analysis
PCR	Polymerase chain reaction
PEG	Poly(ethylene glycol)
PIP3	Phosphatidylinositol (3,4,5)-triphosphate

Rheb	GTPase ras homolog enriched in brain
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
SAM	Significance analysis of microarrays
scRNA	Single cell RNA
sgRNA	Single guide RNA
siRNA	Small interfering RNA
TCA	Tricarboxylic acid
Thbs2	Thrombospondin-2
TIDE	Tracking indels with decomposition
TLC	Thin layer chromatography
TSC	Tuberous sclerosis complex
VLDL	Very low-density lipoprotein
VLDLR	Very low-density lipoprotein receptor
WT	Wildtype
NPC	Nuclear pore complex

SUMMARY

Nanoparticle-mediated nucleic acid delivery can regulate the expression of any gene, making it a promising way to treat disease. However, clinically relevant[1, 2] delivery of RNA therapies to non-hepatocytes *in vivo* remains challenging[3]. It is known that the efficiency with which nanoparticles deliver RNA depends on (i) the route of administration[4] (ii) the chemical composition of a nanoparticle, (iii) the protein corona that adsorbs onto a nanoparticle[5], and (iv) interactions between a nanoparticle and a cell[6]. Despite these differences, LNPs are typically screened *in vitro* before a few finalists are confirmed *in vivo*. Since particles that tend to work *in vitro* do not necessarily work *in vivo*, this leads to a small number of viable candidates.

This work will describe the development and use of novel nanoparticle DNA barcoding systems[7-9] which increase the number of LNPs that can be concurrently studied *in vivo*. Specifically, Chapter 2 discusses the inefficiency of using *in vitro* static cell culture to find LNPs that have potency and specificity *in vivo* by testing whether *in vitro* delivery is predictive of *in vivo* delivery, and understanding how different cell types within a tissue microenvironment are targeted. Importantly, this thesis shows that *in vitro* LNP delivery is not predictive of *in vivo* LNP delivery – evidence that prompted the development of newer generation *in vivo* screening models.

Chapter 3 incorporates a diverse set of chemically modified cholesterols into LNPs in an attempt to shift LNP tropism. Using high-throughput DNA barcoding technologies, the impact of modifications to both the ring and tail of cholesterol on LNP delivery as well as structure-function relationships are assessed. This work leads to discussion on the impact

of cholesterol trafficking on LNP delivery, specifically looking at the impact of LDLR and VLDLR.

In Chapter 4, an additional set of chemically modified cholesterols is screened using FIND, a high-throughput functional LNP delivery screening system[9]. The addition of a hydroxyl group on the cholesterol tail leads to potent, low dose delivery while addition of a hydroxyl group to different positions on the cholesterol rings does not. Finally, this thesis shows that novel cholesterol variants can be used in place of regular cholesterol to generate potent, and selective delivery vehicles for siRNA, sgRNA, or mRNA.

Chapter 5 investigates the ability of a small molecule to manipulate LNP mRNA delivery. Upregulation of the mTOR/PI3K pathway through exogenous administration of the small molecule PIP3 leads to a decrease in exogenous LNP-delivered mRNA translation both *in vitro* and *in vivo*. The mechanism of action that leads to decrease in mRNA translation is investigated using cell transcriptomics and metabolomics.

Finally, Chapter 6 addresses the potential for related works and proposes new directions worthy of pursuit within the field of nucleic acid drug delivery. Specifically, highlighted is the need for new screening systems that are agnostic to both cell type and mouse model, as well as potential new directions that arise by coupling nucleic acid drug delivery with transcriptomics and genomics.

CHAPTER 1. INTRODUCTION

1.1 An overview of nucleic acid-based therapies

1.1.1 Nucleic acids target nodes in the DNA-to-protein information pipeline

The discovery of DNA as the source of hereditary material[10], the subsequent identification of DNA's helical structure a decade later[11], and the proposition that transfer of genetic information flows unidirectionally from nucleic acid to protein[12] were groundbreaking works that paved the way for future studies on nucleic acids. It is now known that DNA can be copied via replication and transcribed into RNA, and that some RNAs – such as fully processed mRNAs – can be translated into protein. While RNA replication and RNA-DNA reverse transcription does occur, these processes are limited to viral and bacterial species, as opposed to eukaryotes. Nucleic acid-based therapeutics, like many other therapeutic strategies, hijack, modify, or complement a missing part of the information pipeline. Examples of this include gene editing (e.g. CRISPR-Cas9), gene silencing (e.g. siRNA, miRNA), gene replacement (e.g. mRNA), and viral delivery (e.g. AAV, lentivirus) strategies.

In order to modulate RNA translation and suppress excess gene transcripts, cells rely on an RNAi-based method that utilizes two non-coding RNAs known as miRNA and siRNA. miRNAs and siRNAs can be delivered exogenously or can be transcribed endogenously. Naturally, miRNA genes are transcribed into hairpin laden primary miRNA (pri-miRNA), which is processed by the enzyme Drosha in the nucleus before being exported to the cytoplasm as pre-miRNA. In the cytoplasm, these pre-miRNAs are cleaved by a ribonuclease called Dicer to form miRNA (~19-22 nt). miRNAs can then interact with

RISC for further processing – the endonuclease AGO2 unbinds and discards the passenger strand (sense strand) but, in most cases, does not cleave it due to imperfect complementarity between the sense and antisense strand[13]. This imperfect complementarity allows the remaining miRNA guide strand (antisense strand) within mature RISC to act upon its target in a multitude of ways. miRNAs can inhibit translation initiation, cause mRNA degradation or, in cases where perfect complementarity between the guide strand and the target mRNA exists, lead to mRNA cleavage[13].

Interestingly, miRNA binding specificity and method of action is partially determined, among other factors, by a 6-8 nucleotide (nt) seed region on its 5' end[14]. Furthermore, over time, cells have evolved to have multiple miRNA binding sites; this ensures that full silencing of a target mRNA requires binding of multiple miRNAs at independent miRNA-binding sites. These binding sites have been found to act synergistically[15, 16]. Similarly, one miRNA can bind to repress many different mRNAs by binding to the 3' or 5'UTR, making this approach a non-specific but potent endogenous RNAi mechanism[17]. This non-specificity has been explored as a method to replace or replenish low levels of miR-34s, which act on a number of genes, are closely coupled to regulation of p53, and are dysregulated in cancers[18].

While both RNAi types are powerful endogenous control mechanisms, siRNA-based gene silencing has been the dominant approach clinically[1, 2, 19]. In the cytoplasm, long dsRNAs – siRNA precursors –are also cleaved by Dicer to form short dsRNA – siRNA (~20-25 nt). Within RISC, AGO2 cleaves the passenger strand (sense strand) and the remaining guide strand guides RISC to the target complementary mRNA and cleaves it[20]. While inside RISC, the protected guide strand can repeatedly target complementary

mRNA, making this approach specific, potent, long-lasting, and clinically viable. Current clinically viable candidate siRNAs and miRNAs have been modified for stability and shortened to bypass Dicer[21].

Prior to the commercial approval of siRNA-based gene therapies in 2018, ASOs dominated the field of RNA therapy. Different definitions of ASOs encompass RNAi under the same umbrella, but for the purposes of this introduction, they will be kept separate. ASOs have the same principle function as RNAi therapies – to facilitate translational repression and modify RNA splicing. They are short, single stranded sequences between 8 – 50 nt in length that interact with a complementary mRNA sequence through Watson-Crick base pairing, forming DNA-RNA or RNA-RNA hybrids. These hybrids can either mediate RNA cleavage by acting as a substrate for RNase H, an enzyme that cleaves dsRNA, or facilitate binding mechanisms that interfere with the target RNA function without degradation. Examples of ASOs include antagomirs, splice-switching oligonucleotides (SSOs), gapmers, steric block oligonucleotides, and others[21-23]. At present, there are a handful of ASOs in the clinic and others making their way there[24]. A key example is the recently clinically-approved ASO Spinraza, which works by binding to an intronic sequence on SMN2 pre-mRNA in order to control mRNA splicing, increase the proportion of mature mRNAs that contain a specific exon, and, in turn, modulate the production of a specific protein isoform[25].

RNA is traditionally thought of as a middleman between DNA and protein, but RNA can also act as a catalyst and facilitate biological processes such as RNA processing, and RNA cleavage in addition to protein synthesis[26]. These catalytically-active RNAs, known as ribozymes, use their secondary structure to cleave complimentary RNAs on their own.

Therapeutic ribozymes have been developed to target and degrade RNA-based viruses[27, 28]. Yet, other groups of RNAs (e.g. aptamers) have significant specificity for proteins due to their unique secondary structures. Aptamers have been used as a potential replacement for antibodies in diagnostics and as biosensors to detect disease biomarkers or environmental contaminants[29].

While RNAi has become a staple of exogenous gene control at the RNA level, recent advances in molecular biology have yielded control mechanisms that act at the DNA level. These mechanisms make up the field of genome editing and typically utilize the actions of a nuclease that is directed to make double stranded breaks (DSB) at specific locations in the genome. The cells endogenous repair machinery will then use non-homologous end joining (NHEJ) or homology directed repair (HDR) to fix the DSB. NHEJ typically results in indels, potentially causing the transcribed gene to be out-of-frame. HDR relies on a template strand to precisely repair the DSB; this template strand can be exogenously delivered in order to induce precise gene insertions. Gene editing approaches include the use of meganucleases, zinc finger nucleases, transcription activator-like effector nucleases (TALENs), and the clustered regularly interspaced short palindromic repeats (CRISPR/Cas9) system. Of these, CRISPR/Cas9, adapted for mammalian cell use from a bacterial defense mechanism, is regularly seen as the most feasible editing mechanism[30, 31]. This feasibility arises from the ease at which different parts of the genome can be specifically targeted. Briefly, an sgRNA is designed to bind complementary target DNA and induce the endonuclease activity of Cas9[32]. This system has been used with catalytically inactive Cas9, termed ‘dead-Cas9’ (dCas9)[33] or short sgRNAs, termed ‘dead-guides’ to induce site-specific transcriptional activation without the formation of a DSB[34].

Analogs of the CRISPR/Cas9 system have since been discovered with applications in diagnostics, imaging, RNA cleavage, base editing, and deamination[35-38]. Despite the fact that most RNA-targeted therapeutics result in gene silencing or inactivation, the CRISPR/Cas9 system can be used to both curtail or induce the expression of a gene, as well as to induce the expression of a functionally inserted gene[32]. These attributes have made CRISPR/Cas9-based systems a viable option for diseases with single-gene disruptions[39-43]. Like any system, the CRISPR/Cas9 system has flaws that will need to be addressed before it becomes a clinically viable platform. Some of these caveats include the need to optimize the CRISPR/Cas9 system to avoid off-target editing[44] and to accommodate the presence of pre-existing immunity in a large percentage of the population[45, 46]. A commonly used version of Cas9 (spCas9) comes from the bacterium *Streptococcus pneumoniae*, the most common cause of community acquired pneumonia and meningitis. Similarly, researchers will need to understand the pharmacokinetics and temporal relationship between Cas9 mRNA translation and guide RNA delivery in order to optimize the interaction between the two components.

Most of the nucleic acid payloads discussed previously aim at reducing the presence of a target protein by interfering during transcription or translation. However, in the early 1990s, researchers found that by injecting naked mRNA into mice they could correct disease phenotypes through transient protein production[47]. Since then, synthetic mRNAs have been developed for many applications including immunotherapies, infectious disease vaccines, transient protein expression, and nuclease delivery for editing systems. mRNA vaccines have been used to deliver the coding sequence of an antigen in order to immunize against influenza A[48], HIV[49, 50], and ZIKV[51] in large animal models with some vaccine

candidates having begun testing in humans. Similarly, mRNA coding for nucleases such as Cas9, Cpf1, ZFNs, and TALENs have been successfully used for genome editing purposes. To this end, a number of companies have been formed to find successful delivery vehicles for co-delivery of Cas9 mRNA and sgRNA[52]. Nucleases have traditionally been delivered in the form of a protein or as DNA in a viral vector. Protein nucleases can be difficult to deliver *in vivo* because of their size and stability, while viral vectors ensure long-term nuclease production, which can result in non-specific editing over time[53].

mRNA therapeutics are convenient for a few reasons – namely, because they are transient and do not require genome integration. The pharmacokinetics of an mRNA drug can be optimized on a case-by-case basis through structural RNA modifications and sequence optimizations. These changes are made to increase the stability and duration of the RNA, protect it from exonucleases and endonucleases, and to improve mRNA translatability, respectively. Codon optimization is a useful technique to increase the translatability and stability of an mRNA therapeutic. By modifying the mRNA sequence, rare codons that are not commonly utilized by the target species' cell machinery can be removed and frequently used codons can be included[54]. This allows for changing of the RNA sequence, but not the resulting amino acid sequence of the therapeutic protein. The diversity of available payloads has made the field of nucleic acid therapeutics wide and promising. However, despite exciting advances in the design, characterization, and utility of these payloads, their feasibility as therapeutic agents depends heavily on their stability in an *in vivo* environment, and their ability to be delivered exogenously to their intended site of action. Nucleic acid stability will be addressed in section 1.1.2 and delivery approaches in section 1.2.

1.1.2 Nucleic acid stability and the role of chemical modifications

Oligonucleotide therapeutics obtain their potency from their chemical structure. Early nucleic acid-based drugs were plagued by their inability to function *in vivo* at clinically relevant doses. At that time, systemic administration of minimally modified nucleic acid drugs was done at relatively high doses in order to elicit effective translational silencing[55]. These high-dose, unmodified RNA administrations would often trigger the innate immune system by activating TLRs[56]. Whereas target specificity is conferred by the sequence of the oligonucleotide, in order to improve the potency and stability of oligonucleotides, modifications need to be made to the sugar, backbone, nucleobase, or 3' or 5' end of the nucleic acid[57]. Modifications are typically included in order to curtail degradation via endonucleases, exonucleases, and hydrolysis, and in turn increase stability and durability, or to improve target RNA affinity. The most common modification is the use of a phosphorothioate backbone to block digestion by exonuclease III and some restriction enzymes[58, 59]. However, the addition of a phosphorothioate backbone also reduces target RNA binding affinity[57, 58], typically offset by the addition of sugar modifications that improve target affinity and stability. The most common of these sugar modifications occurs at the 2' position of RNA and DNA – specifically, 2'-O-methyl (2'-OMe), 2'-O-methoxyethyl (2'-MOE), and 2'-deoxy-2'-fluoro (2'-F) modifications are commonly used[21, 57].

Antisense and RNAi technologies have evolved to both positionally and structurally optimize the use of chemical modifications. Depending on the delivery vehicle and route of administration, different levels and types of chemical modification may be required[60-62]. For instance, only half of the bases on an siRNA need to be modified if the siRNA is getting

delivered with a LNP, whereas heavier modification patterns are required if the siRNA is getting delivered with a conjugate. While many biomolecules have been assessed as oligonucleotide drug conjugates, successful systemic conjugate-mediated delivery has been limited to the liver by utilizing GalNAc, a ligand for the rapidly recycled asialoglycoprotein receptor (ASGPR) on hepatocytes. GalNAc-conjugated oligos require modification of every nucleotide in order to function potently – however the type of modifications used may contribute to off-target delivery[63].

1.2 The utility of nucleic acid drug delivery vehicles

1.2.1 Oligonucleotide delivery vehicles

Currently, successful nucleic acid therapeutics are either locally administered to easily accessible sites using intraocular, intranasal, intramuscular, or intrathecal delivery methods or target hepatocytes in the liver using a conjugate ligand (e.g. GalNAc), a LNP, or a viral vector. Despite the potency of the payload, there is an imminent need for technologies that allow for the identification of nucleic acid delivery vehicles outside of the liver using systemic administration. With developments and improvements in oligonucleotide potency, the biggest barrier to successful delivery is reaching the intended target site.

Lipid-based delivery vehicles have been used since the mid-2000s to deliver RNA therapeutics[64-66] and early papers showed the potential of lipid delivery systems. Researchers were able to systemically deliver an siRNA, using a lipid-based delivery vehicle, to NHPs. The siRNA, targeting the APOB gene, was encapsulated in a stable nucleic acid lipid particle (SNALP) and led to protein knockdown out to six days at doses

as low as 1.0 mg/kg. These particles were similar in composition –they contained a PEG-lipid, phospholipid, cholesterol, and a cationic lipid – and size (70 – 90 nm) to present day LNPs and laid the foundation for future LNP optimization and discovery papers[66].

Despite prior work to find potent lipid-based vehicles, discovery and synthesis of novel materials required multistep, time-intensive synthetic routes. To address this bottleneck, researchers developed a chemical method to rapidly develop new lipid-like molecules[64]. This method relied on chemistry that conjugated alkyl-acrylates or alkyl-acrylamides to primary or secondary amines using Michael addition. This led to the production of 700 lipidoids that were then screened in HeLa cells one-by-one at multiple doses. After *in vitro* screening, researchers showed that they could take the best performing LNPs *in vitro* and use them to silence Factor VII, a blood clotting protein produced by hepatocytes, *in vivo*. This study was followed by multiple others that in total screened over 1000 nanoparticles for potent, low dose, *in vivo* siRNA delivery[67, 68]. These studies analyzed the impact of the lipomer:RNA ratio, PEG chain length, and particle size on LNP pharmacokinetics and biodistribution[69, 70]. Combinatorial approaches to making new compounds for siRNA delivery continued, with the synthesis of cKK-E12, a potent ionizable lipid-like material[71]. This ionizable lipid had an ED_{50} of 0.002 mg/kg when delivering Factor VII siRNA in mice and is currently licensed for clinical development. cKK-E12 set the benchmark for potent and specific hepatocyte delivery vehicles and has been used in multiple formulations since its original publication[72-74].

Other approaches to finding novel chemical structures used pre-existing cationic lipids (e.g. DLinDMA, DOTAP, DOTMA) to inform the creation of novel cationic lipids[65]. These amphiphilic molecules have an amine based headgroup and a hydrocarbon chain

with a linker between the two. Researchers varied the linker between the headgroup and hydrocarbon chain in order to make novel cationic lipids. They then took the best preforming linker modification and modified the headgroup to additional lipids. Of these, the best performing lipid (DLin-KC2-DMA) had an ED_{50} of 0.3 mg/kg and worked as low as 0.1 mg/kg when delivered to NHPs[75]. Similar advances have been made for the identification of non-liver delivery vehicles. By screening epoxide-modified lipid-polymer hybrids at different lipid:RNA ratios *in vitro*, researchers were able to identify a nanoparticle containing a new lipid-compound termed 7C1 that potently delivered siRNA *in vivo*[76]. Specifically – 7C1 potently delivers siRNA to endothelial cells in the lung (ED_{50} ~ 0.01 mg/kg) and heart (ED_{50} ~ 0.05 mg/kg) after a single systemic injection.

More recently, utility of lipid nanoparticles has expanded to include mRNA and DNA delivery. The biggest challenges that mRNA therapeutics face is their ability to be effectively delivered to their target cell type. A number of therapeutics have been developed that utilize LNPs, particularly focusing around the use of ionizable lipids. This shift in payload comes with its own challenges – siRNAs and miRNAs are typically small (~18-22 nucleotides) whereas mRNAs can be anywhere from a couple hundred basepairs to a few kilobases, making the barrier to clinically relevant cell delivery greater. This size disparity impacts the size and polydispersity index of an LNP. Studies have optimized the lipid:RNA ratio differently for siRNA and mRNA-based payloads. Due to their size, more siRNAs can typically be loaded into a LNP, with ideal lipid:RNA ratios being between 5-10, whereas ideal lipid:RNA ratios for mRNA are between 7.5-10. This relationship as well as interactions between other LNP characteristics have been investigated, focusing on the use of variants of cKK-E12, 7C1, and a new class of lipids, termed conformationally

constrained lipids, to alter the tropism of LNPs *in vivo*[9, 77]. Additionally, a diverse set of cholesterols[72, 78], both synthetic and natural, as well as phospholipids, have been identified as important components of potent LNPs. Despite challenges, it comes as no surprise that novel delivery vehicles for splenic ECs[9], splenic T-Cells[77], bone marrow ECs[79], splenic immune cells[80], and liver ECs[72] have been identified. For example, biodegradable variants of the potent ionizable lipid cKK-E12 have led to potent splenic immune cell delivery[80]. Similarly, formulating cKK-E12 with the charged phospholipid DOTAP instead of DOPE has led to potent lung delivery with minimal liver off-target delivery[81]. The impact of phospholipid charge on LNP tropism has been further evaluated[82]. Researchers developed a method known as Selective Organ Targeting (SORT) of nanoparticles in order to shift delivery of canonically liver-specific LNPs to the spleen and lung. Specifically, they showed that cationic phospholipids (e.g. DOTAP, DDAB, EPC) shift delivery to the lung while anionic phospholipids (e.g. 18PA, 18BMP) shift delivery to the spleen. Phospholipid charge may impact the overall LNP surface charge which has been shown to vary the types and amount of serum proteins that adhere to the LNP and drive tropism to different tissues[83, 84]. Often, efforts are made to reduce adsorption of serum proteins in the hopes of increasing circulation time and decreasing rapid clearance[83, 85]. For example, by varying the type and amount of PEG-lipid added to an LNP, researchers have shown that they can modulate LNP circulation time and tissue localization[86, 87]. The addition of PEG-lipids creates an aqueous shell around the LNP, reducing interaction with serum proteins and subsequent clearance. It is evident that the interplay between LNP components, the mole ratio between components, and the payload being delivered all play a role in the tropism and potency of clinically relevant LNPs.

While lipid nanoparticles have begun to make their way into the clinic – the first LNP-based siRNA delivery vehicle was approved in 2018 – viral based delivery strategies have been around for more than two decades. One caveat to LNP delivery is the current inability to access the nucleus of a cell directly. There is evidence to suggest that LNPs are endocytosed through multiple endocytosis mechanisms, including clathrin, caveolin, and micropinocytosis[88], and a small portion of the nucleic acid payload contained within the LNP ends up making it to the cytoplasm of a cell. Attempts at getting these payloads into the nucleus include adding nuclear localization sequences or hijacking endogenous cell machinery to import the nucleic acid. Similar to LNPs, viruses bind to specific receptors on the outside of a cell[89, 90] and get endocytosed via myriad endocytosis pathways[91, 92]. However, unlike LNPs, viruses are also efficient at getting their payload into the nucleus of a cell. Different DNA viruses have evolved NPC-dependent and NPC-independent methods of getting into the nucleus. For instance, recombinant AAVs are believed to achieve nuclear import through an NPC-dependent mechanism[92]. However, the rate of endosomal escape and nuclear import can vary between AAV serotypes. When long-term nuclear expression is the requirement – as is often the case with treatments that require delivery of a functional version of a gene that will be present for the duration of a patients' life – viruses are exceptionally good at delivering. AAVs in particular have gained steadfast appreciation for a number of reasons. These include their packaging capabilities (~4.7-5 kb), long-term gene expression without DNA integration in nondividing cells, and their relatively low immunogenicity compared to other viral vectors[93]. In tissues that are predominantly post-mitotic such as the liver, retina, and skeletal muscle, these

characteristics make AAVs a favourable long-term gene therapy vector. In fact, there are already clinically approved AAVs for spinal muscle atrophy[94] and retinal dystrophy[95].

However, some current viral-based strategies are troubled by the same issues as nanoparticles, namely delivery to on-target tissues outside of the liver and spleen without direct vector administration. Systemic delivery of AAV often leads to potent liver delivery – to obtain vector delivery outside of the liver, a high dose of the drug is required. This is true even for recombinant AAV serotypes that have tropism for non-hepatic tissues, such as AAV1, 6, and 9, which deliver to cardiomyocytes in the heart[92]. As another example, AAVs have been pursued as a viable option for delivery of functional dystrophin DNA to patients with muscular dystrophy – but, this approach requires delivery of the payload to muscle. While there are direct-administration-based approaches (e.g. intramuscular injection) where this technique would work well, potent systemic administration of an AAV intended to deliver to muscle with minimal off-target liver delivery is challenging[41, 96]. To this end, researchers are attempting to find novel AAV capsids that have increased tropism and selectivity for skeletal muscle[97]. While it is clear that the evolution of nucleic acid delivery vehicles has yielded potent clinical candidates, it has also uncovered the complexity associated with the success of these delivery vehicles. Discussed below are a few of the known challenges that nanomedicines encounter in an *in vivo* environment, with specific focus on LNPs, the topic of this thesis.

1.2.2 Complexity of lipid nanoparticle delivery

Nanomedicine delivery is a complex process regulated by the body[98]. Successful delivery requires a nanomedicine to protect the drug from degradation, avoid the systemic

immune system, avoid clearance organs[99], enter the desired tissue, select the right cell type within a complex tissue microenvironment and - if the drug requires cytoplasmic delivery - gain access to the cytoplasm without degrading in an organelle (**Fig. 1.1**). Given these barriers, it quickly becomes clear that a major challenge with nucleic acid therapeutics is their delivery to the correct cell or tissue. For example, nanoparticles carrying small molecules have been safely administered to patients[100], and siRNA delivered to hepatocytes by GalNAc conjugates[101] or lipid nanoparticles[1] have treated genetic disease. At the same time, the need for systemically administered nanomedicines that target non-hepatocytes is significant, since most systemically-administered drug delivery systems are still sequestered in the liver.

At each step, the nanomedicine must overcome defenses that have evolved to sequester and degrade foreign materials, making drug delivery inefficient. For example, a LNP that delivers siRNA to hepatocytes in mice, NHPs, and humans[1] was used to ask an important question: if a LNP carrying siRNA reaches the endosome of a target cell *in vivo,* what percentage of the siRNA accesses the cytoplasm? This LNP only released 2% of its siRNA into the cytoplasm[88]. Recognizing these inefficiencies, clinical advances in nanotechnology research[1, 100-103] are impressive. In order for nanoparticles to undergo a step function-like change in efficacy, we believe it is critical to answer four fundamental questions. First, how does nanoscale chemical structure, and in particular cholesterol structure, influence drug delivery *in vivo*? Second, which biological pathways govern nanoparticle delivery *in vivo*? Third, is screening directly *in vivo* the best approach? And fourth, can we quicken the pace at which clinically relevant lipid nanoparticles are

identified? These questions have been reviewed extensively elsewhere[6, 104, 105] and over three decades of work has resulted in a substantial body of knowledge.

It is well known that biological molecules and physical barriers govern on-target delivery, off-target delivery, toxicity, and the endosomal escape of LNPs. For example, labs have made important contributions to our understanding of the protein corona. We now know that the surface area:volume ratio of nanoparticles as well as hydrophilic/hydrophobic interactions make it thermodynamically likely[5] that diverse proteins will bind to nanoparticles after they are administered[106]. It is also known that the composition of this protein corona changes with time[5, 107] and the local microenvironment[4, 108, 109]. Finally, there is consensus that the protein corona can dictate interactions between a nanoparticle and the immune system[110], or a nanoparticle and a target cell[84, 111-113]. To further this knowledge, *in vivo* studies of lipid nanoparticle dynamics and cellular interactions at a high-throughput scale are necessary.

As a second example, it is clear that physical barriers can influence nanoparticle delivery. When a LNP is administered intravenously, serum protein opsonization can lead to rapid clearance due to nonspecific uptake facilitated by the mononuclear phagocyte system[114, 115]. For example, in the liver, Kupffer cell sequestration is an important barrier to functional LNP delivery to hepatocytes and other hepatic cell types. Cationic nanoparticles can be disassembled by the renal anionic basement membrane[116], and nanoparticles can have a difficult time accessing brain parenchyma, due to the capillary tight junctions and glial cells that make up the blood brain barrier[117-121]. By contrast, nanoparticles can access tissues like the liver because of discontinuous basement membranes in the hepatic sinusoids[117-121]. Notably, detailed *in vivo* analysis and

mathematical modeling was used to demonstrate that slow blood flow through the hepatic sinusoids increased nanoparticle uptake, potentially by increasing the likelihood nanoparticles extravagated out of the blood[99]. This suggests that altered blood flows can influence nanoparticle interactions in other diseases known to have unusual flow, including atherosclerotic plaques[122].

Endocytic pathways have also been shown to affect nanoparticle delivery. Many studies have identified genes that impact nanoparticle uptake *in vitro*[88, 123-125], and a growing number of publications have performed similar studies *in vivo*. For example, when nanoparticle pharmacokinetics were measured in healthy mice and tumor-bearing mice, authors observed a difference in delivery to non-tumor organs[126]. Similarly, nanoparticle pharmacokinetics changed in ApoE$^{-/-}$ and LDLR$^{-/-}$ mice[127], complementing previous data that demonstrate cholesterol trafficking pathways can influence *in vivo* nanoparticle delivery[65]. Small molecules that modulate specific pathways have also been shown to influence nanoparticle delivery. A recent publication demonstrated that the anti-malarial drug Chloroquine reduced nanoparticle uptake by macrophages[128], and that rapamycin delivered in nanoparticles increased the tolerability of other biologic drugs[129]. Nanoparticle chemical composition[130], shape[131-133], and size[134] can affect particle endocytosis, whereas genes related to lysosomal storage[135], anti-viral immune response[136], and mRNA translation[125] all impact the fate of a LNP once it is inside the cell. These broad examples represent a fraction of the work that has been performed at the interface of nanotechnology and biology over the previous decades[137] and highlight the complexity of nanoparticle delivery.

Despite these advances, questions regarding the interactions between nanotechnology and biology remain. For example, how do the structures of clinically relevant nanoparticles influence delivery? Are there universal rules that relate nanoparticle chemical traits to potency? How should nanoparticles be designed so that they reach an on-target tissue and cell type and do not get cleared by clearance organs? The following chapters describe the use of DNA barcoding as a high-throughput tool to study LNP delivery[7] and identify non-hepatocyte targeting LNPs (**Fig. 1.2**). Most current methods of LNP identification rely on *in vitro* testing to identify LNPs that work potently *in vivo*. This approach does not model any of the barriers to delivery previously described, making it a difficult approach to rely on when searching for LNPs whose functional actions translate across the species line. Throughout this work, high-throughput LNP barcoding is used to bypass *in vitro* testing and assess LNP functional delivery and biodistribution within an animal model as well as to identify novel, clinically relevant, LNPs with new tropisms.

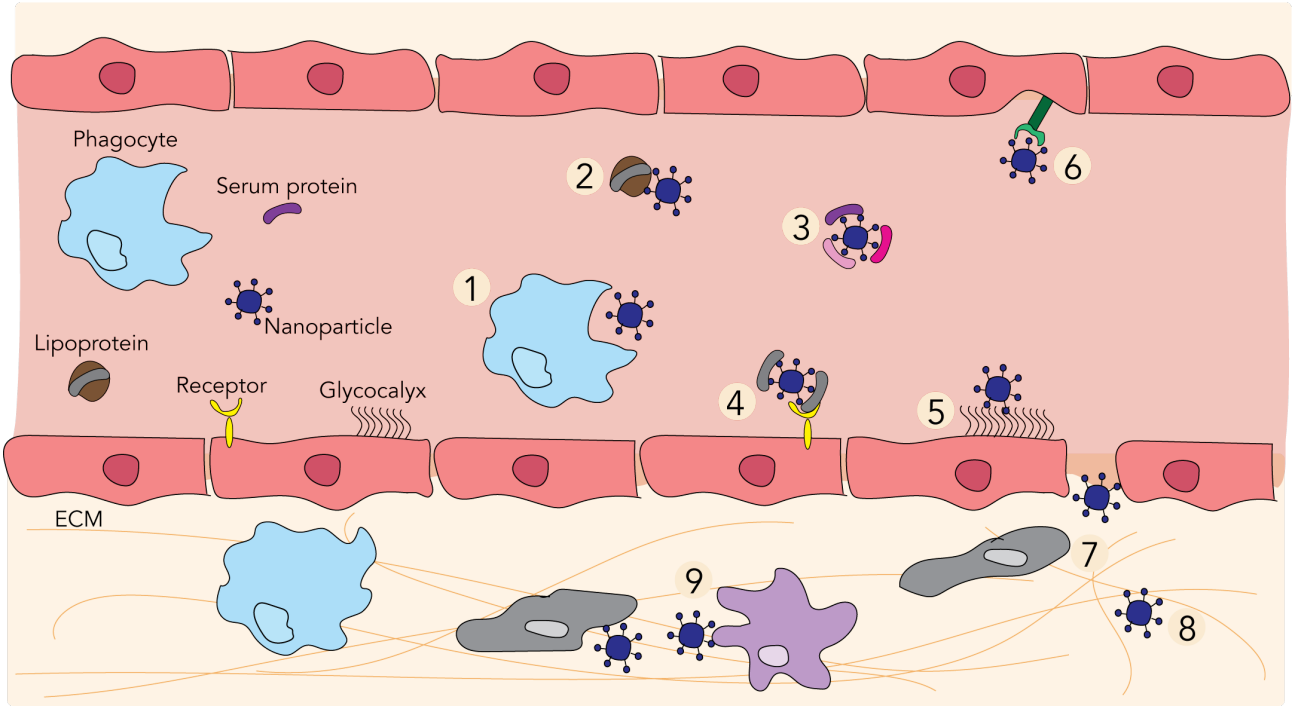


Figure 1.1 LNP delivery can be viewed as a complex phenotype affected by many cells and biomolecules. (A) Nanoparticles are (1) cleared by circulating immune cells and tissue resident immune cells. Due to their high surface area: volume ratio, nanoparticles interface with (2) lipoproteins and (3) other biomolecules that make up the protein corona. The corona, in turn, can (4) alter how nanoparticles bind target cells. Interestingly, depending on its composition, the nanoparticle corona can promote or inhibit cell targeting. While reaching target cells, nanoparticles also interact with (5) a dense 'forest' of cell surface glycoproteins and glycolipids, collectively termed the glycocalyx. Alternatively, nanoparticles may interact (6) directly with cell surface receptors. Nanoparticles can also exit the bloodstream; this process is affected by (7) the permeability of vascular endothelial cells. Within the target tissue, nanoparticles interact with (8) proteoglycans in the extracellular matrix (ECM), or (9) cells within the tissue itself. (B) DNA- and RNA-driven gene expression dictates nanoparticle behavior by controlling the synthesis and processing of proteins, sugars, and lipids. As a result, high throughput quantification of the five biomolecules could improve our understanding of biological pathways that affect nanoparticle delivery. Two methods are typically used: next generation sequencing, which quantifies DNA and RNA, and mass spectroscopy, which quantifies lipids, carbohydrates, and proteins. The scale at which DNA and RNA can be analyzed is currently greater than the scale at which lipids, carbohydrates, and proteins can be analyzed.

s

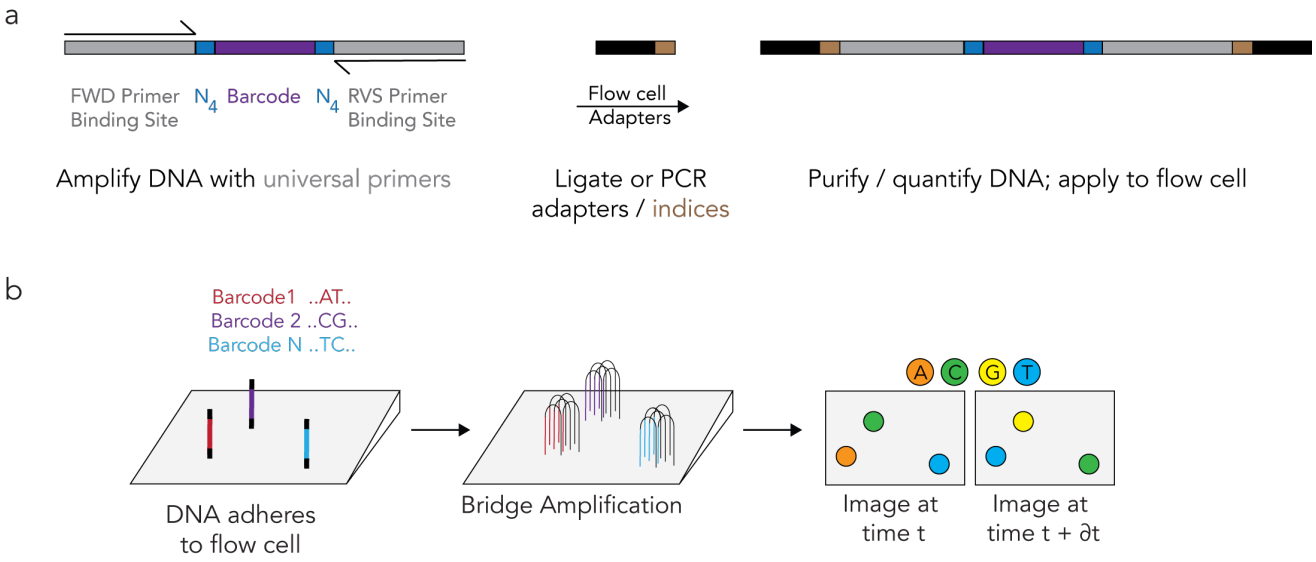


Figure 1.2 Next generation sequencing for high-throughput DNA barcoding. (a) DNA barcodes are most effective when amplified with universal primers. The barcode sequence is often a smaller number of nucleotides located in a larger DNA sequence. After amplification, sample indices and sequencing adapters are added using PCR or ligation. The sequencing adapters (b) bind the DNA to the sequencing flow cell. In solid state bridge amplification, each individual DNA sequence is amplified on the flow cell to create a cluster. The sequence of each cluster is read using fluorescent nucleotides. In this way, flow cells can generate hundreds of millions of data reads in a single experiment.

CHAPTER 2. A DIRECT COMPARISON OF IN VITRO AND IN VIVO LIPID NANOPARTICLE DELIVERY

The work presented here is an excerpt from Paunovska* K, Sago* CD, Monaco CM, Hudson WH, Gamboa Castro M, Rudoltz TG, Kalathoor S, Vanover DA, Santangelo PJ, Ahmed R, Bryksin AV, Dahlman JE (2018). "A Direct Comparison of *in Vitro* and *in Vivo* Nucleic Acid Delivery Mediated by Hundreds of Nanoparticles Reveals a Weak Correlation." *Nano Letters.*

2.1 Introduction

The transport of foreign nucleic acids is carefully regulated, making systemic drug delivery inefficient[98, 114, 138]. To deliver nucleic acids, thousands of chemically distinct lipid nanoparticles have been designed[64, 68, 71, 76, 139, 140]. LNP chemical diversity is imparted 2 ways. First, thousands of distinct biomaterials have been synthesized[64, 68, 71, 76, 139, 140]. Second, each biomaterial can be formulated into hundreds of LNPs by adding PEG, cholesterol, DOPE, and other constituents at different mole ratios. LNPs are screened *in vitro* before a small number of finalists is tested *in vivo*[64, 68, 71, 76, 139, 140]; in a typical example, we measured how well 2,000 LNPs delivered siRNA to HeLa cells in static cell culture before analyzing 5 LNPs *in vivo*[76]. Similar studies have been performed with LNPs composed of small amine-, sugar-, or peptide-based materials[64, 68, 71, 139, 140].

To successfully deliver nucleic acids after systemic administration, nanoparticles must overcome complex obstacles that are difficult to model *in vitro*. Nanoparticles must protect the DNA or RNA from degradation, avoid clearance, target the desired cell within a complex microenvironment, and gain access to the cytoplasm, without stimulating a systemic immune response. A significant fraction of the drug can be lost at each point, and as a result, every step is important to model. Physical variables also influence delivery. Cationic nanoparticles can be disassembled by the renal anionic basement membrane[116], and nanoparticles can be blocked from accessing brain parenchyma, due to the capillary

tight junctions and glial cells that make up the blood brain barrier[117-121]. By contrast, nanoparticles can access tissues like the liver via porous ECs and discontinuous basement membranes in the hepatic sinusoids[117-121]. In addition to these barriers, *in vivo* analysis and mathematical modeling have demonstrated that blood flow rates affect nanoparticle targeting by affecting the likelihood a particle extravasates[99].

Despite these differences, nanoparticles are often screened for delivery *in vitro.* Large scale *in vitro* nanoparticle screens typically utilize cells that are easy to expand (e.g. HeLa)[64, 68, 71, 76, 139, 140]. These cells have genotypes and phenotypes that differ from cells *in vivo.* Cells can also undergo significant changes in gene expression when cultured[141]. Many of these changes may be driven by exposure to a combination of foreign serum (e.g., FBS) and static fluid flow, which most cell types are not exposed to *in vivo*. Given that *in vitro* and *in vivo* delivery require the nanoparticle to overcome different physiological obstacles, and that endocytosis is likely to be affected by gene expression changes that occur when cells are removed from their natural microenvironment, we hypothesized that *in vivo* LNP delivery would not be predicted *in vitro* using common cell culture conditions.

The field can currently synthesize nanomaterials at a rate several orders of magnitude higher than the rate at which we can test nanomaterials for drug delivery *in vivo.* Recently, we reported a nanoparticle DNA barcoding system[7] to increase the number of LNPs we could study at once *in vivo*. We used a microfluidic device to barcode LNPs (**Fig. 2.1A**)[142]; each LNP was formulated to carry a unique DNA barcode. We pooled stable LNPs, administered them to animals, and deep sequenced the barcodes to quantify the delivery of up to 30 LNPs simultaneously[7]. This original paper focused exclusively on control experiments designed to characterize the system. Specifically, we demonstrated that barcoded LNPs can be made so they do not mix in solution, that DNA sequencing readouts were linear with respect to the amount of injected DNA, that DNA barcode delivery recapitulates the behavior of previously characterized LNPs, that delivery does

not change with DNA sequence, and that delivery of DNA barcodes to hepatocytes *in vivo* modeled siRNA delivery to hepatocytes *in vivo*[7].

We now report that the same LNP barcoding system, herein named JORDAN, can elucidate fundamental questions about nanoparticle delivery. We quantified how well 281 LNPs delivered DNA barcodes to endothelial cells and macrophages, both *in vitro* and *in vivo*. We focused on endothelial cells and macrophages for three reasons. First, both cell types are implicated in many diseases[143, 144]. Second, since they are ubiquitous, we could measure delivery to the same cell type in multiple tissues. Third, endothelial cells are more 'accessible' upon intravenous injection than tissue macrophages. We reasoned, incorrectly, that delivery to more accessible cells would be more predictable *in vitro*. Our data strongly suggest that *in vitro* LNP delivery to endothelial cells and macrophages using static cell culture does not predict *in vivo* LNP delivery to the same cell types.

We then used the JORDAN system to investigate how different LNPs distribute within a clearance organ (the spleen). By measuring how 85 LNPs delivered barcodes to 8 different splenic cell types, we found that cells derived from myeloid progenitors tended to be targeted to by similar LNPs; cells derived from lymphoid progenitors tended to be targeted by different LNPs. We then identified LNP1, which delivered barcodes to all 8 cell types we studied in the spleen. We confirmed the splenic targeting of LNP1 using fluorescently labeled DNA. The approach we have described can be extended to study (i) how well any *in vitro* system (e.g., tissue-on-a-chip) predicts delivery *in vivo,* and (ii) how different cells are targeted within a tissue.

2.2 Results

2.2.1 Rationally designed DNA barcodes for in vivo LNP delivery

We rationally designed DNA barcodes in order to study the delivery of many LNPs at once (**Fig. 2.1A-C**). Each DNA barcode contained phosphorothioate linkages in order to reduce exonuclease activity, and universal primer sites for unbiased PCR amplification (**Fig. 2.1C**)[7]. The 8 nucleotide 'barcode' region was located in the middle of the 56 nucleotide DNA sequence. Of the 4^8 possible DNA barcode combinations, we designed 240 to work with Illumina sequencing machines (**Fig. 2.2A**). We amplified barcodes using universal primers and labeled individual samples with Illumina dual-indexed adapters that enabled sample multiplexing (**Fig. 2.2B**). For each experiment, we calculated the 'normalized delivery', using the administered LNP solution as a DNA input (**Fig. 2.1D, Fig. 2.2C**). We also added new LNP quality controls to reduce the likelihood LNPs mixed together. Specifically, we analyzed the size of each individual LNP using DLS. Based on our experience studying LNPs[7, 31, 76, 145-147], we only pooled stable LNPs with good autocorrelation curves and diameters between 20 and 300 nm (**Fig. 2.2D**). We then tested whether barcoded LNPs entered cells. We formulated Alexa647-tagged DNA barcodes in a previously characterized[7, 31, 76, 145-149] LNP named 7C1. Barcodes entered iMAECs within 15 minutes and were observed inside the cell at 1.5 and 72 hours after administration (**Fig. 2.1E, Fig. 2.2E, F**).

2.2.2 LNPs can be screened using high-throughput barcoding

We then formulated 144 LNPs, systematically varying PEG structure. We synthesized 2 biomaterials called 'lipomers', which are lipid-amines conjugates created by reacting epoxide, acrylate, or methacrylate-terminated lipids with oligoamines[7, 64, 76, 150, 151]. Both lipomers were formulated into 72 LNPs, using 9 different PEGs, and 8 different PEG mole percentages, for a total of 144 LNPs (**Fig. 2.3A, Fig. 2.4A**). 112 out of 144

formulations formed stable LNPs and were pooled (**Fig. 2.4B**). We administered the 112 stable LNPs, as well as a naked DNA barcode - which served as a negative control - to cells at a total DNA dose of 4, 20, and 100 ng/well, in a 24 well plate. Concurrently, we administered the LNPs to mice via a tail vein injection at a total dose of 0.5 mg/kg DNA. We isolated DNA from cells or tissues 72 hours later, a time point we chose to minimize the influence of dynamic endocytic processes[76, 88, 124, 152, 153]. The 4 ng total DNA dose equaled an average DNA dose of 0.035 ng/well/LNP, demonstrating the sensitivity of the DNA barcoding system. We administered the LNPs to iMAECs, and mouse macrophages (RAWs). We chose iMAECs since they are isolated directly from the murine heart, and have been shown to recapitulate endothelial cell signaling and function[154]. We chose RAWs since they are a commonly used cell line.

2.2.3 In vitro LNP delivery does not recapitulate in vivo delivery

We examined positive and negative controls to evaluate whether this dataset was robust. The naked barcode (negative control) performed poorly compared to LNP-delivered DNA in all 18 samples (**Fig. 2.3C, D**). LNP delivery in iMAECs and RAWs treated with 20 ng total DNA predicted LNP delivery in iMAECs and RAWs treated with 4 or 100 ng DNA with high precision ($R^2 > 0.9$) (**Fig. 2.3E, Fig. 2.4B-D**). Put another way, in this positive control experiment, delivery to iMAECs *in vitro* at 1 dose predicted delivery to iMAECs at 2 other doses.

We then investigated whether *in vitro* LNP delivery predicted *in vivo* LNP delivery (**Fig. 2.5A**). We compared normalized delivery in iMAECs and RAWs to endothelial cells and macrophages isolated from mice injected with LNPs. We isolated endothelial cells and

macrophages from mice using a FACS protocol we previously established[7, 76, 145, 148] (**Fig. 2.6A-C**). *In vitro* iMAEC delivery did not predict *in vivo* delivery to heart, lung, or bone marrow endothelial cells (**Fig. 2.6J-L**). Similarly, delivery to RAW cells *in vitro* did not predict delivery to heart, lung, or bone marrow macrophages (**Fig. 2.6S-U***)*. To validate these results, we synthesized two additional LNP libraries (**Fig. 2.6D-G**). Library two consisted of 120 LNPs, of which 105 were found to be stable by DLS and included (**Fig. 2.6E**). In this library, we systematically varied the lipid tail and amine reacted to make the lipomer component of each LNP. Library three consisted of 156 LNPs, of which 64 were found to be stable by DLS and included (**Fig. 2.6G**). In this library, we systematically varied the PEG tail length and MW, using three different tail lengths (C14, C16, C18) and 2 different MWs (350, 2000 Da). Results from libraries two and three recapitulated results from library one; *in vitro* delivery to endothelial cells and macrophages did not predict *in vivo* delivery to the same cell types. In total, we performed three experiments, formulating 420 LNPs, of which 281 were stable and included **(Fig. 2.6H-I)**. Results from each individual experiment are plotted in Fig. 2.6J-AA. Combined results from all 3 experiments are plotted in **Fig. 2.5B-G**.

We considered the possibility that our results were due to a poor choice of cell line or time-point. To exclude this possibility, we investigated to what extent the (i) cell line and (ii) experimental time point altered the predictivity of *in vitro* delivery. We performed these experiments using library two. We administered library two to iMAECs, RAWs, and mice, and measured delivery 4, 48, and 72 hours after LNP administration (**Fig. 2.6M-O, V-X, BB-LL)**. At all three time points, we also administered LNPs to three primary human

endothelial cell lines: HAECs, HAVECs, and HUVECs. We observed no strong relationship between *in vitro* and *in vivo* delivery (**Fig. 2.6MM-UU**).

We then analyzed our data set, with the goal of quantifying the 'efficiency' of traditional *in vitro* screening. Put another way, if a LNP library is screened *in vitro* and a small number of LNPs is selected for *in vivo* analysis, how likely is it that the best *in vivo* candidate is selected? We first calculated the percentage of an *in vitro* library required to select the top 5, 10, 15, or 20% of *in vivo* LNPs. To ensure the top 5% of *in vivo* LNPs were selected, more than 50% of the *in vitro* library would need to be selected (**Fig. 2.7A, Fig. 2.8A**). We then analyzed how well the best *in vivo* LNPs performed *in vitro*. We ranked LNPs based on their *in vitro* performance. We then colored the LNPs that performed in the top 10% *in vivo* (**Fig. 2.8D**). Some LNPs that performed well *in vivo* also performed well *in vitro*. However, in many cases, LNPs that performed well *in vivo* did not rank highly *in vitro*; these LNPs would likely be discarded after an *in vitro* screen. Third, we evaluated how the top *in vitro* LNPs performed *in vivo*. Top ranked *in vitro* LNPs did not consistently perform well *in vivo* (**Fig. 2.7B, Fig. 2.8B**). Based on this result, we asked a 4th question: if we selected the top 3, 5, or 20 *in vitro* LNPs, how likely were we to pick the 1st, 1st and 2nd, or 1st, 2nd, and 3rd ranked *in vivo* LNPs? The odds of finding the top LNP *in vivo* were 11%, 22%, and 44% using the top 3, 5, and 20 *in vitro* LNPs, respectively; the odds of finding the top 2, or top 3 *in vivo* LNPs were lower (**Fig. 2.7C, Fig. 2.8C**). Taken together, these data strongly suggest that *in vitro* delivery may not predict systemic *in vivo* delivery. While they do not directly implicate all *in vitro* systems or all cell types, they do strongly suggest each *in vitro* system should be validated using many nanoparticles before being used as the basis for nanoparticle discovery.

2.2.4 High-throughput screening enables us to analyze LNP-property relationships

The JORDAN system generates large nanoparticle datasets; the size of these datasets enabled us to analyze the relationship between LNP properties and *in vivo* delivery statistically. We plotted DNA barcode delivery as a function of each material property. In total, we analyzed 309 relationships between LNP structure and *in vivo* delivery. We found that the lipomer alkyl tail length, lipomer amine structure, and PEG MW were most likely to influence LNP delivery (**Fig. 2.8E**). These results suggest that the structure of the amine-lipid compound, as well as the degree of LNP PEG both strongly influence LNP targeting. These results provide an important insight into LNP library design and substantiate previously reported nanoparticle research[69, 86, 155]. One important limitation is that we were not able to identify a mathematical framework with assumptions that allowed us to analyze how multiple LNP chemical variables interacted with one another. This future work is important, given that changing 1 LNP parameter often impacts another (e.g., adding more PEG to the LNP concomitantly reduces cholesterol).

2.2.5 LNP delivery to immune cells clusters by progenitor type

Nanoparticle biodistribution is quantified using *ex vivo* tissue fluorescence; however, it is still unclear how different cell types within a tissue microenvironment are targeted[76, 99]. More specifically, it is unclear which cell types tend to be targeted by similar LNPs. To address this question, we focused on the spleen; LNPs can deliver nucleic acids to[156], and be cleared by[99], the spleen. We formulated a 4th LNP library. Library 4 consisted of 144 LNPs, of which 85 were found to be stable by DLS, and included (**Fig. 2.10A, B**). We administered library 4 to RAWs and mice, and isolated 8 different cell types using

FACS (**Fig. 2.9A, Fig. 2.10C**). We performed unbiased Euclidean clustering, which is used to compare how many experimental groups relate to one another. Recapitulating our results from libraries 1-3, Euclidean clustering separated clearly separated *in vivo* delivery to all cell types (including macrophages) from *in vitro* delivery to RAWs (**Fig. 2.9B, Fig. 2.9C, Fig. 2.9G**).

More interestingly, the 7 immune cell sub-types clustered into cells derived from (i) myeloid progenitors and (ii) lymphoid progenitors, respectively. Plasmacytoid and conventional DCs were clustered most closely with one another, and also clustered closely with macrophages and neutrophils (**Fig. 2.9B, Fig. 2.9C**). All 4 cell types derive from a common progenitor. T cells, B cells, and natural killer cells, which derive from a different progenitor, clustered together (**Fig. 2.9B, Fig. 2.9C**). To quantify this clustering, we measured the correlation between all 8 *in vivo* cell types and RAWs (**Fig. 2.9C**). Conventional DCs and plasmacytoid DCs, which clustered together, were highly correlated (R^2 value = 0.90) (**Fig. 2.9C, 2.9D**), as were B cells and T cells (R^2 value = 0.88) (**Fig. 2.9C, 2.9E**). Cells derived from myeloid progenitors (e.g., conventional DCs) and lymphoid progenitors (e.g., T cells) had a much weaker correlation (R^2 value = 0.2) (**Fig. 2.9F**). These data demonstrate a unique capability for the JORDAN system; directly comparing how dozens of LNPs deliver nucleic to 7 different cell types would be very challenging using traditional one-by-one methods (e.g., fluorescence).

Using delivery data generated from this screen, we identified two LNPs for additional characterization (**Fig. 2.11A**). Barcodes delivered by LNP1 were enriched in all 8 splenic cell types, relative to barcodes delivered by LNP2 (**Fig. 2.11B, C**). We formulated LNP1 and LNP2 separately, using a Cy5.5-tagged DNA barcode, and injected mice

intravenously with 0.75 mg/kg DNA barcode. LNP1-treated mice had 12.1x more splenetic Cy5.5 *ex vivo* fluorescence than LNP2-treated mice, recapitulating the barcode readouts (**Fig. 2.11D**).

2.3 Discussion

We found that *in vivo* delivery to macrophages and endothelial cells is not predicted *in vitro* using common cell culture conditions. Modeling all the factors (e.g., blood flow, vascular heterogeneity, systemic and local immune cells, unwanted delivery to clearance organs) that influence nanoparticles *in vivo* is challenging. These results have important implications for nanoparticle design, given that nanoparticles are typically selected *in vitro*. We compared delivery to 5 different cell lines, and 6 different *in vivo* cell types; it will be important to determine whether these results extend to other cell types and cell culture conditions. For example, it was previously shown that delivery in primary hepatocytes was more predictive of *in vivo* delivery than a hepatoma cell line[157]. At first glance, our results may seem to contradict this work. We believe they do not; both studies underscore the importance of characterizing how well a given *in vitro* system predicts a desired *in vivo* outcome. To this end, we believe JORDAN is well positioned to optimize organ-on-a-chip[158] and organoids[159] designed to predict *in vivo* behavior by acting as a positive control.

JORDAN is agnostic to cell type and animal model. This allowed us to easily study drug delivery to 8 different splenic cell types in a single experiment. Testing how many LNPs target several cell types may lead to interesting discoveries. For example, our results suggest cells derived from myeloid progenitors tend to be targeted by the similar LNPs. This provides preliminary evidence that gene expression patterns that promote LNP

delivery to phagocytic cells may be identified. However, these results need to be validated using other nanoparticles.

We noticed that there were practical advantages to using the JORDAN system. Testing many nanoparticles at once reduces experimental variation that occurs when experiments performed over months are compared to one another. We previously screened LNPs one by one[76]; it was difficult to ensure kits, reagents, and cell passage number were perfectly consistent. By testing many LNPs on the same day, it is easier to reduce unintentional experimental bias. At the same time, the JORDAN system has limitations. JORDAN is unlikely to work with unstable or toxic LNPs; it is critical to characterize particles before pooling them. JORDAN measures biodistribution, which is required, but not sufficient, for intracellular delivery. It will also be important to prevent PCR contamination. Finally, like all high throughput screening systems, lead candidates need to be independently verified. For example, we identified LNP1, which performed well in our barcode screen, and LNP2, which performed poorly in our barcode screen, before confirming their activity one by one (**Figs. 2.9, 2.11**). As part of our original barcoding study[7], we performed a similar confirmation experiment using Factor 7 siRNA.

Despite these caveats, we believe our data demonstrate that JORDAN is a powerful new tool to help researchers understand *in vivo* drug delivery. We also believe the differences between *in vitro* and *in vivo* delivery suggest that *in vivo* screening may accelerate the rate at which clinically relevant LNPs can be discovered. To help other labs use the JORDAN system, we have published an open source, 'living' document, which details reagents, protocols and our LNP bioinformatics pipeline on dahlmanlab.org.

2.4 Materials and Methods

Nanoparticle Formulation. Nanoparticles were formulated in a microfluidic device by mixing DNA with lipomer, PEG, cholesterol, and a helper lipid, as previously described[7, 31, 76, 142, 145-149, 160]. Nanoparticles were made with variable molar ratios of these constituents. The genetic drug (in this case, DNA barcode) was diluted in 10 mM citrate buffer (Teknova), and loaded into a syringe (Hamilton Company). The materials making up the nanoparticle (lipomer, cholesterol, PEG, and helper lipid) were diluted in 100% ethanol, and loaded into a second syringe. The citrate phase and ethanol phase were mixed together in a microfluidic device, at rates of 600 uL/min and 200 uL/min, respectively, to form LNPs. We used the following helper lipids: DOPE (Avanti Lipids, 850725), and DOPC (Avanti Lipids, 850375).

DNA barcoding. Each chemically distinct LNP was formulated to carry its own unique DNA barcode (**Fig. 2.1A, B**). For example, LNP1 carried DNA barcode 1, while the chemically distinct LNP2 carried DNA barcode 2. The DNA barcodes were designed rationally with several characteristics, as we previously described[7]. We purchased 56 nucleotide single stranded DNA sequences from IDT (**Fig. 2.1C, Fig. 2.2A**). We included 2 universal 21 and 20 nucleotide primer regions in addition to a random 7 nucleotide ('7N') region that is unique to each piece of DNA (**Fig. 2.2B**). Barcodes were distinguished using an 8 basepair (bp) sequence in the middle of the barcode. An 8 bp sequence can generate over 1,000,000 (4^8) unique barcodes; we selected 240 barcodes to prevent sequence bleaching on the Illumina MiniSeq™ machine. The 2 nucleotides on the 5' and 3' ends of the 56-nucleotide ssDNA sequence were modified with phosphorothioate linkages to reduce exonuclease degradation and improve DNA barcode stability.

Nanoparticle Characterization. LNP hydrodynamic diameter was measured using high throughput DLS (DynaPro Plate Reader II, Wyatt). LNPs were diluted in sterile 1X PBS to a concentration of ~0.0005 mg/mL and analyzed. LNPs were included if they met 3 criteria: diameter >20 nm, diameter <300 nm, and autocorrelation function with only 1 inflection point. Over the course of our experiments, ~65% of the LNPs we formulated met all 3 criteria. Particles that met these criteria were pooled and dialyzed with 1X phosphate buffered saline (PBS, Invitrogen), and were sterile filtered with a 0.22 μm filter.

Animal Experiments. All animal experiments were performed in accordance with the Georgia Institute of Technology's IACUC. Female C57BL/6J (#000664) mice were purchased from the Jackson Laboratory. All mouse weights before and after injection, are shown in (**Fig. 2.12A, B**). In all experiments, mice were aged 4-12 weeks, female, and N = 3 – 5 mice per group were injected intravenously via the lateral tail vein with the same pooled LNPs.

Nanoparticle dosing. Mice were injected with a total DNA dose of 0.5 mg/kg. As an example, if an experiment measured 100 nanoparticles, then on average, each nanoparticle was administered at a dose of 0.005 mg/kg. The nanoparticle dose was determined using NanoDrop (Thermo Scientific).

Cell Culture. *In vitro* experiments were performed using mouse macrophage cells (RAW 264.7, ATCC), mouse aortic endothelial cells (provided by Hanjoong Jo at Emory)[154], HAEC (Lonza), HUVEC (Lonza), and HAVEC (Lonza). In all cases, cells were maintained and cultured using previously established conditions. In all cases, cell media was supplemented by penicillin-streptomycin (500 U/mL penicillin G, 0.5 mg/mL

streptomycin) (PenStrep, VWR*)* and 10% (v/v) FBS (VWR). RAW cells were passaged with DMEM F-12 50/50 (Corning). iMAEC cells were passaged using DMEM with 1 g/L glucose, L-glutamine, and sodium pyruvate (Corning), supplemented by 1% (v/v) MEMNEAA (Sigma Aldrich), and 25 μg/mL ECGS (Emd Millipore). HAEC and HAVEC cells were passaged with MCDB 131 media without L-glutamine (VWR Scientific), supplemented by 1% (v/v) L-glutamate, 25 μg/mL ECGS, 0.1% (v/v) ascorbic acid, hydrocortisol, and the following growth factors: endothelial growth factor (EGF), vascular-endothelial growth factor (VEGF), fibroblast growth factor (FGF), and insulin-like growth factor (IGF). HUVEC cells were passaged with M199 media with Earle's salts and L-glutamine (Corning), supplemented by 1% (v/v) ECGS, L-glutamine, and 0.2% (v/v) heparin. In all cases, cells were seeded in a 24-well plate at a density of 40,000 cells/well. 24 hours later, LNPs were added with a total DNA dose of 4, 20, or 100 ng (**Fig. 2.3E, Fig. 2.4B-D**). Based on these results, cells were treated with 20 ng total DNA in all other experiments. Six hours after transfection, media was removed, and fresh media was added. 72 hours after transfection, media was removed and DNA was isolated using QuickExtract (EpiCentre).

Fixed-cell staining. Cells were plated onto 35 mm glass-bottom dishes (In Vitro Scientific) one day prior to particle delivery. Cells were fixed at the indicated time points with 4% paraformaldehyde (Electron Microscopy Sciences) for 10 min at room temperature before permeabilization with 0.2% Triton X-100 (Sigma-Aldrich) for 5 min at room temperature. To stain actin, cells were then incubated with Phalloidin-488 (Thermo Scientific) for 30 minutes at 37°C. Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (Life

Technologies), and coverslips were placed over the cells in the dish and mounted with Prolong Gold (Life Technologies).

Microscopy. Images were acquired with a Hamamatsu Flash 4.0 v2 sCMOS camera on a PerkinElmer UltraView spinning disk confocal microscope mounted to a Zeiss Axiovert 200M body with a 63x NA 1.4 plan-apochromat objective. Images were acquired with Volocity (PerkinElmer) with Z-stacks taken in 0.2 μm increments. For live-cell images, cells and dishes were kept at 37 °C during imaging by using a Chamlide TC-L live-cell stage-top environment with objective heater (Live Cell Instrument). All images were linearly contrast enhanced. Live-cell images were smoothed with a fine rolling ball filter in Volocity.

Cell Isolation. One time-course experiment was performed; delivery was analyzed 4, 48, and 72 hours after LNPs were administered. In all other cases, tissues and cells were isolated 72 hours after injection with LNPs. In all experiments, mice were perfused with 20 mL of 1X PBS through the right atrium. The heart, lungs, spleen, and bone marrow were isolated immediately following perfusion. Tissues were finely cut, and then placed in a digestive enzyme solution with Collagenase Type I (Sigma Aldrich), Collagenase XI (Sigma Aldrich), and Hyaluronidase (Sigma Aldrich). The digestive enzyme for heart included Collagenase IV (Sigma Aldrich)[145]. Tissues were digested for 45 minutes at 37ºC and 550 rpm. Digested tissues were passed through a 70 μm filter. Red blood cells were lysed using (RBC) lysis buffer. Cells were resuspended in FACS buffer (2% FBS in 1X PBS).

Heart, Lung, Bone Marrow Cell Staining. Cells were stained to identify specific cell populations and sorted using the BD FACSFusion and BD FACS Aria IIIu cell sorters in the Georgia Institute of Technology cellular analysis core. Antibodies used for staining were CD31 (Clone 390, BioLegend), CD45.2 (Clone 104, BioLegend), and CD11b (Clone M1/70, BioLegend). We defined cell populations in the following manner: macrophages ($CD31^{-}CD45^{+}CD11b^{+}$), heart and lung endothelial cells ($CD31^{+}CD45^{-}$), bone marrow endothelial cells ($CD31^{+}$), immune cells excluding macrophages ($CD31^{-}CD45^{+}CD11b^{-}$), and other cells ($CD31^{-}CD45^{-}$) **(Fig. 2.6A-C).**

Splenetic Cell Staining and Isolation. Splenocytes were isolated by digesting sliced spleens in 0.1 U/ml collagenase (Sigma Aldrich) in Hank's Balanced Salt Solution (Corning) for 30 minutes at 37°C. Digestion was stopped by addition of EDTA to 5 mM, and the resulting mixture was passed through a cell strainer. Red blood cells were removed with ACK lysing buffer (Lonza), and cells were washed in FACS buffer (PBS supplemented with 2% FBS and 2 mM EDTA; Corning and HyClone) and re-strained. Cells were stained by conventional methods in FACS buffer as previously described[161]. Antibodies used for staining were CD19 (clone 1D3, eBioscience), CD3 (clone 17A2, eBioscience), CD31 (clone 390, BioLegend), CD45 (clone 104, BioLegend), NK1.1 (clone PK136, eBioscience), CD11b (clone M1/70, BioLegend), CD11c (clone N418, BioLegend), Siglec H (clone 551, BioLegend), F4/80 (clone BM8, BioLegend), and Ly-6G (clone 1A8, BioLegend). Cells were also stained with LIVE/DEAD viability dyes (Thermo Fisher) to exclude dead cells. Described splenocyte cell types (**Fig. 2.9A, Fig. 2.10D**) from four mice were isolated by FACS on two FACSAria II cell sorters (BD Biosciences) at the Emory University School of Medicine Flow Cytometry Core.

PCR Amplification. All samples were amplified and prepared for sequencing. More specifically, 1μL of primers (5uM for Final Reverse/Forward, 0.5uM for Base Forward) were added to 5 μL of Kapa HiFi 2X master mix, 3 μL sterile H_2O, and 1 μL DNA template. The reaction was run for 30 cycles. When the PCR reaction did not produce clear bands, the primer concentrations, DNA template input, PCR temperature, and number of cycles were optimized for individual samples. The PCR amplicon was isolated with gel extraction.

Deep Sequencing. Illumina deep sequencing was conducted in Georgia Tech's Molecular Evolution core. Runs were performed on an Illumina Miniseq^{TM}. Primers were designed based on Nextera XT adapter sequences.

Data Normalization. Counts for each particle, per tissue, were normalized to unity (**Fig. 2.1D**) The DNA counts in each tissue were then normalized to 100%. For example, if a sample (e.g., heart1) had 500,000 total barcode reads, and 50,000 of them came from particle X, while 4,000 came from particle Y, then the normalized delivery for particle X and Y would be 10% and 0.8%, respectively. The barcoded LNP mixture we injected into the mouse was also sequenced. This 'input' DNA was used to normalize DNA counts from the cells and tissues (**Fig. 2.2C**).

Data Analysis. Sequencing results were processed using a custom python-based tool to extract raw barcode counts for each tissue. These raw counts were then normalized with an R script prior to further analysis. Statistical analysis was done using GraphPad Prism 7. Correlation analyses were run assuming a Gaussian distribution in order to obtain Pearson correlation coefficients. R^2 values (0 – 1) were computed by squaring Pearson correlation coefficients.

Data Access. The data, analyses, and scripts used to generate all figures in the paper are available upon request to J.E.D. or dahlmanlab.org.

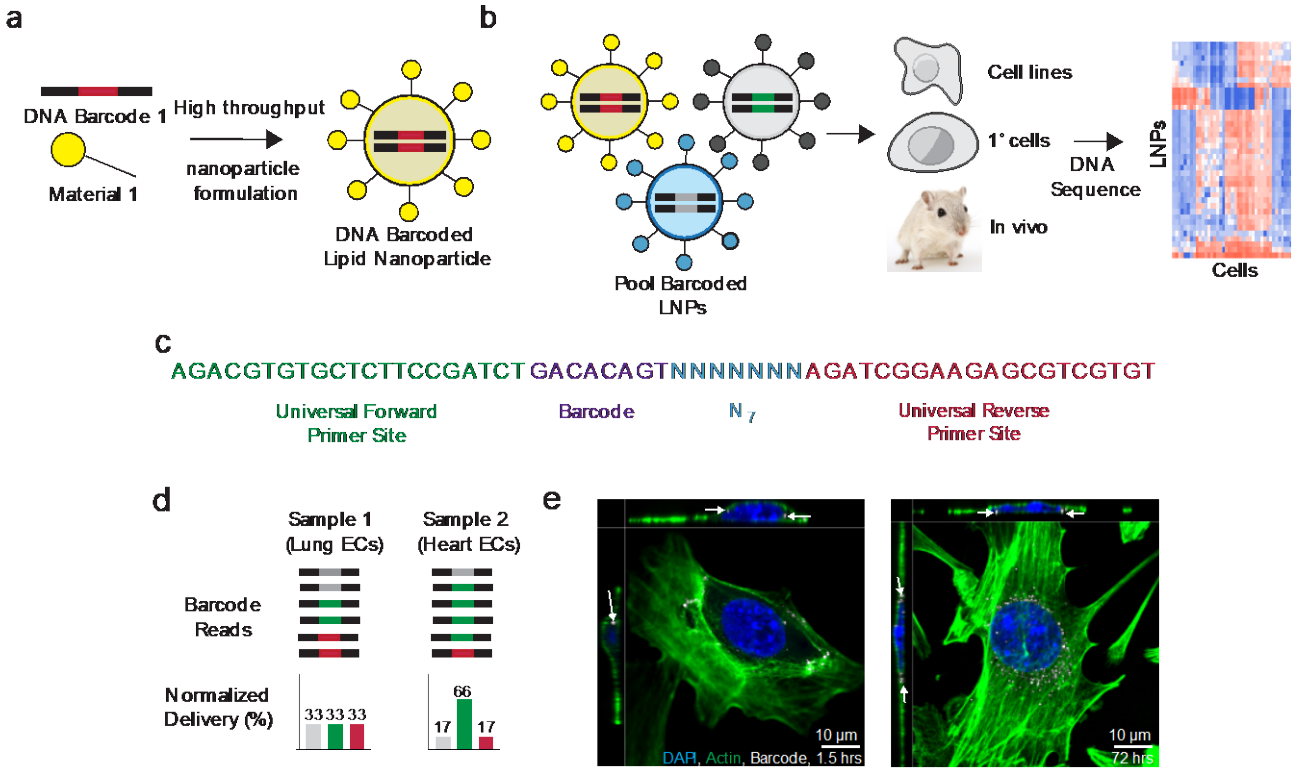


Figure 2.1 JORDAN, a system for high throughput *in vivo* nanoparticle analysis. (a) LNPs were formulated to carry DNA barcodes, before (b) stable LNPs were pooled together and administered to cells or mice. Cells were deep sequenced to quantify the relative delivery of all the LNPs simultaneously. (c) The DNA barcode was rationally designed with universal primer sites and a randomized 7 nucleotide region to minimize PCR bias. (d) Normalized delivery for every barcoded LNP was calculated. In this example schematic, all 3 barcodes were equally represented in Sample 1, while in Sample 2, the green barcode was overrepresented. We would hypothesize that the gray LNP delivered DNA more efficiently to Sample 2 than the yellow or blue LNP. The full data analysis to calculate normalized delivery is described in Fig. 2.2C. (e) Alexa-647 fluorescence 1.5 and 72 hours after cells were transfected with 20 ng of Alexa Fluor 647 tagged DNA barcode formulated into the LNP 7C1.

a

Forward Universal Site Barcode 7N Region Reverse Universal Site

A*G*A CGT GTG CTC TTC CGA TCT GAC ACA GTN NNN NNN AGA TCG GAA GAG CGT CGT* G*T

1	GAC ACA GT	43	GCT ACA AC	85	AGT ATG CC	127	CAT CTG CT	169	ACT GCT TG	204	CTT AGG AC
2	GCA TAA CG	44	CCG ATG TA	86	TAC TGC TC	128	CCA ACG AA	170	AGA AGC CT	205	CTT GCT AG
3	ACA GAG GT	45	TAG GAG CT	87	TGC TTG CT	129	CCA GTT GA	171	AGA TCG TC	206	GAA CGG TT
4	CCA CTA AG	46	AAC AAG GC	88	TCC ACG TT	130	CCG GAA TA	172	AGC GTG TA	207	GAA GTG CT
5	TGT TCC GT	47	CTC GGT AA	89	AAC CAG AG	131	CCT ACC TA	173	AGG CAA TG	208	GAC GTC AT
6	GAT ACC TG	48	AGC TTC AG	90	ACG AAC GA	132	CCT ATT GG	174	AGG TTC CT	209	GAG ACC AA
7	AGC CGT AA	49	TCA CCT AG	91	ATA GTC GG	133	CCT TGG AA	175	AGT CGA AG	210	GAG TAG AG
8	CTC CTG AA	50	CAA GTC GT	92	CCA TGA AC	134	CGA ATT GC	176	AGT GGC AA	211	GAT CCA CT
9	ACG AAT CC	51	CTG TAT GC	93	GAG CAA TC	135	CGA GTT AG	177	ATA CTG GC	212	GAT GGA GT
10	AAT GGT CG	52	AGT TCG CA	94	CAA CTT GG	136	CGC TGA TA	178	ATC CTT CC	213	GCA CAC AA
11	CGC TAC AT	53	ATC GGA GA	95	CCA CAA CA	137	CGG CAT TA	179	ATC TCC TG	214	GCA TTG GT
12	CCT AAG TC	54	AAG TCC TC	96	TGG TGA AG	138	CGT CAA GA	180	ATG CGC TT	215	GCC TTC TT
13	TTG CTT GG	55	TGG ATG GT	97	AAC ACG CT	139	CTA AGA CC	181	ATT CCG CT	216	GCT AAG GA
14	CCT GTC AA	56	AGG TGT TG	98	AAC AGG TG	140	CTA GGT TG	182	CAA GCC AA	217	GCT GAA TC
15	AGC CTA TC	57	GAC GAA CT	99	AAC CTA CG	141	CTC ACC AA	183	CAA TGC GA	218	GGA AGA GA
16	TGA TCA CG	58	GTT CTT CG	100	AAG ACA CC	142	CTC GAC TT	184	CAT CAA CC	219	GGA CTA CT
17	CCA CAT TG	59	TTC GCC AT	101	AAG CGA CT	143	CTC TCA GA	178	ATC CTT CC	220	GGA TGT AG
18	TCG AGA GT	60	CAA CTC CA	102	AAT CGC TG	144	CTG ATG AG	179	ATC TCC TG	221	GGA TTC AC
19	GGT CGT AT	61	AAC CGT GT	103	AAT TCC GG	145	TAC CTG CA	180	ATG CGC TT	222	GGT ATA GG
20	ACA GGC AT	62	CGG TTG TT	104	ACA CCT CA	146	CTT ACA GC	181	ATT CCG CT	223	GGT TAG CT
21	GTG ATC CA	63	CTA GCA GT	105	ACA GTT CG	147	CTT CGG TT	182	CAA GCC AA	224	GTA AGC AC
22	TTC GTA CG	64	ACC TCT TC	106	ACC ATG TC	148	GAA CGA AG	183	CAA TGC GA	225	GTA CGA TC
23	ATG ACA GG	65	TAC TAG CG	107	ACC GGT TA	149	GAA GAT CC	184	CAT CAA CC	226	GTC AAC AG
24	CGA CCT AA	66	ACA ACA GC	108	ACG AGA AC	150	GAC CGA TA	185	CAT GAG CA	227	GTC CTT GA
25	TAT GGC AC	67	CGC AAT GT	109	ACG CTT CT	151	GAG AAG GT	186	CCA ACT TC	228	GTC TGA GT
26	ATA ACG CC	68	CAG TGC TT	110	ACT CCT AC	152	GAG GCA TT	187	CCG AAG AT	229	GTG TGT TC
27	GTA GTA CC	69	TCT AGG AG	111	ACT GCG AA	153	GAT CAG AC	188	CCG TAA CT	230	GTT CCA TG
28	CGC GTA TT	70	GAT TGT CC	112	AGA ACC AG	154	GAT GCT AC	189	CCT AGA GA	231	TAA GCG CA
29	ATC CAC GA	71	GGT ACG AA	113	AGA TAC GG	155	AAC AAC CG	190	CCT CAT CT	232	TAC AGA GC
30	TAA CGT CG	72	CTT CAC TG	114	AGC GAG AT	156	AAC AGT CC	191	CGA ACA AC	233	TAG CAG GA
31	CCT TCC AT	73	ATA GGT CC	115	AGG AGG TT	157	AAC GCA CA	192	CGA CAC TT	234	TAG CTT CC
32	GAT CAA GG	74	ACC GAC AA	116	AGG TAG GA	158	AAG ACC GT	193	CGA TCG AT	235	TAG TGC CA
33	AAG CAT CG	75	AAC ACT GG	117	AGT CAG GT	159	AAG CGT TC	194	CGG AGT AT	236	TAT GAC CG
34	AGG ATA GC	76	ACC ATA GG	118	AGT GCA TC	160	AAT GAC GC	195	CGG TAA TC	237	TCA CTC GA
35	GGC TCA AT	77	TCG ATG AC	119	ATA CGC AG	161	ACA AGA CG	196	CGT CCA TT	238	TCA GTA GG
36	TTC ACG GA	78	GAC TTG TG	120	ATC CGT TG	162	ACA CGA GA	197	CTA CAA GG	239	TCC GAT CA
37	GGC GAA TA	79	CCG TTA TG	121	ATC GTG GT	163	ACA GCA AG	198	CTA TCC AC	240	TCG AAC CT
38	AAG TGC AG	80	CAA CGA GT	122	ATG CCT AG	164	ACC GAA TG	199	CTC AGA AG		
39	GCA ATT CC	81	TTA CCG AC	123	ATT AGC CG	165	ACC TAG AC	200	CTC GTT CT		
40	CTT CGC AA	82	GAG AGT AC	124	CAA GAA GC	166	ACG ATC AG	201	CTC TGG AT		
41	CAT TGA CG	83	CTG TAC CA	125	CAA TCA GG	167	ACG GAC TT	202	CTG CCA TA		
42	TCT GGA CA	84	TGA GCT GT	126	CAT ACT CG	168	ACT CGA TC	203	CTG AAC GT		

b

Index Name	Index Sequence	Index Name	Index Sequence
N701	TAAGGCCA	S502	ATAGAGAG
N702	CGTACTAG	S503	AGAGGATA
N703	AGGCAGAA	S505	CTCCTTAC
N704	TCCTGAGC	S506	TATGCAGT
N705	GGACTCCT	S507	TACTCCTT
N706	TAGGCATG	S508	AGGCTTAG
N707	CTCTCTAC	S510	ATTAGACG
N710	CGAGGCTG	S511	CGGAGAGA
N711	AAGAGGCA	S513	CTAGTCGA
N712	GTAGAGGA	S515	AGCTAGAA
N714	GCTCATGA	S516	ACTCTAGG
N715	ATCTCAGG	S517	TCTTACGC
N716	ACTCGCTA	S518	CTTAATAG
N718	GGAGCTAC	S520	ATAGCCTT
N719	GCGTAGTA	S521	TAAGGCTC
N720	CGGAGCCT	S522	TCGCATAA
N721	TACGCTGC		
N722	ATGCGCAG		
N723	TAGCGCTC		
N724	ACTGAGCG		
N726	CCTAAGAC		
N727	CGATCAGT		
N728	TGCAGCTA		
N729	TCGACGTC		

Binds to Forward Universal Site

Binds to Reverse Universal Site

Illumina Index

Universal Reverse Primer (N7XX)	AATGATACGGCGACCACCGAGATCTACACTAAGGCCAACACTCTTTCCCTACACGACGCTCTTCCGATCT
Base Forward Primer	TGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Universal Forward Primer (S5XX)	CAAGCAGAAGACGGCATACGAGATATAGAGAGGTGACTGGAGTTCAGACGTGTG

c

Step 1) Raw Counts

		Sample Name								
LNP #	*Barcode*	*Heart 1*	*Heart 2*	*Heart 3*	*Lung 1*	*Lung 2*	*Lung 3*	*INPUT 1*	*INPUT 2*	*INPUT 3*
1	GACACAGT	100	250	200	200	150	100	100	110	150
2	GCATAACG	120	300	150	75	100	120	110	100	190
3	ACAGAGGT	150	260	75	125	90	150	130	150	130
4	CCACTAAG	160	350	100	150	150	175	140	180	150
5	TGTTCCGT	200	100	75	90	200	100	150	190	140
6	GATACCTG	130	150	140	50	55	60	180	200	170
7	AGCCGTAA	300	350	400	100	125	130	120	135	150
8	CTCCTGAA	200	300	250	50	60	55	120	130	135
9	ACGAATCC	100	50	100	300	350	250	115	110	125
10	AATGGTCG	300	350	300	50	60	50	135	130	150

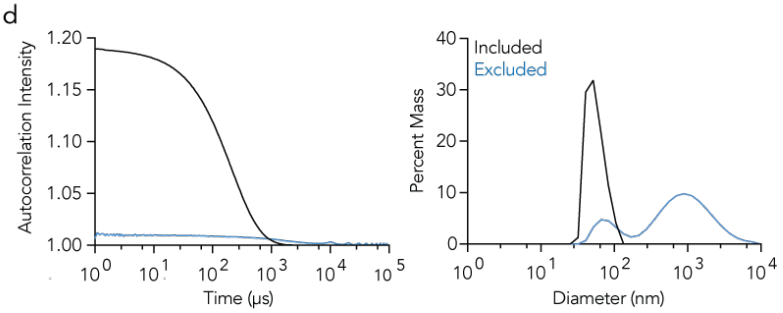
Step 2) Counts Relative to Input Row Mean

		Sample Name						
LNP #	*Barcode*	*Heart 1*	*Heart 2*	*Heart 3*	*Lung 1*	*Lung 2*	*Lung 3*	*INPUT MEAN*
1	GACACAGT	0.83	2.08	1.67	1.67	1.25	0.83	120.00
2	GCATAACG	0.90	2.25	1.13	0.56	0.75	0.90	133.33
3	ACAGAGGT	1.10	1.90	0.55	0.91	0.66	1.10	136.67
4	CCACTAAG	1.02	2.23	0.64	0.96	0.96	1.12	156.67
5	TGTTCCGT	1.25	0.63	0.47	0.56	1.25	0.63	160.00
6	GATACCTG	0.71	0.82	0.76	0.27	0.30	0.33	183.33
7	AGCCGTAA	2.22	2.59	2.96	0.74	0.93	0.96	135.00
8	CTCCTGAA	1.56	2.34	1.95	0.39	0.47	0.43	128.33
9	ACGAATCC	0.86	0.43	0.86	2.57	3.00	2.14	116.67
10	AATGGTCG	2.17	2.53	2.17	0.36	0.43	0.36	138.33

Step 3) Counts Normalized to Column Sum + Bioreplicate Mean

		Sample Name							
LNP #	*Barcode*	*Heart Mean*	*Heart 1*	*Heart 2*	*Heart 3*	*Lung Mean*	*Lung 1*	*Lung 2*	*Lung 3*
1	GACACAGT	10.33	6.60	11.70	12.68	13.50	18.52	12.51	9.47
2	GCATAACG	9.44	7.13	12.64	8.56	8.00	6.25	7.51	10.23
3	ACAGAGGT	7.85	8.70	10.69	4.17	9.74	10.16	6.59	12.48
4	CCACTAAG	8.50	8.09	12.55	4.85	10.97	10.64	9.58	12.70
5	TGTTCCGT	5.66	9.91	3.51	3.57	8.62	6.25	12.51	7.11
6	GATACCTG	5.34	5.62	4.60	5.81	3.25	3.03	3.00	3.72
7	AGCCGTAA	18.24	17.61	14.56	22.54	9.48	8.23	9.27	10.95
8	CTCCTGAA	13.43	12.35	13.13	14.82	4.63	4.33	4.68	4.87
9	ACGAATCC	5.24	6.79	2.41	6.52	27.65	28.57	30.02	24.36
10	AATGGTCG	15.96	17.19	14.21	16.49	4.16	4.02	4.34	4.11

d



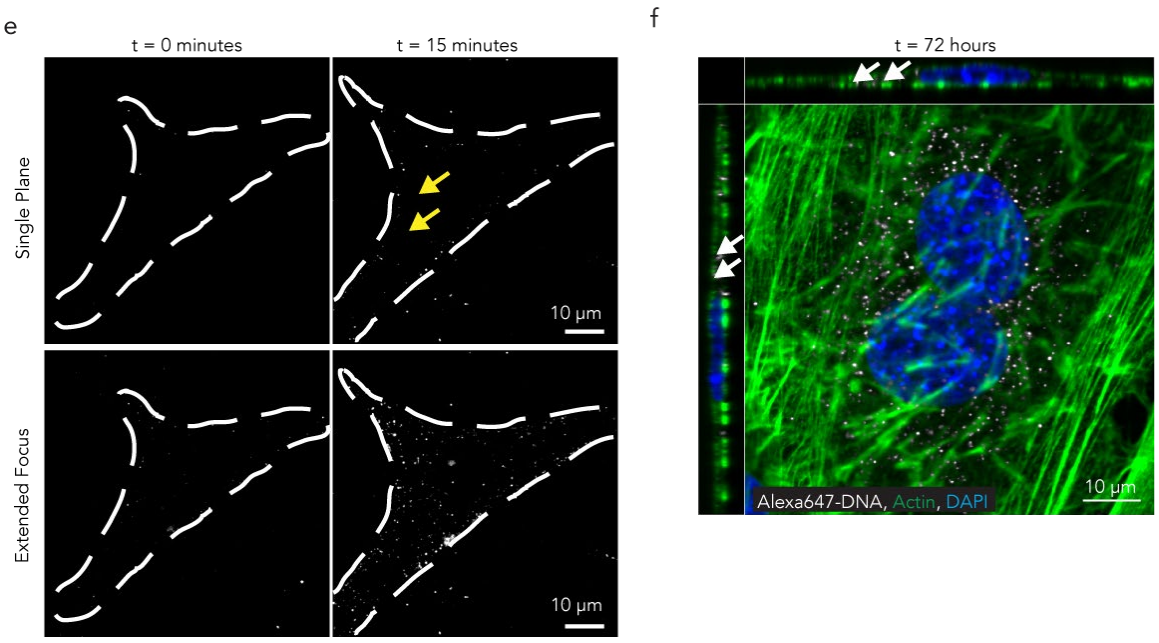


Figure 2.2 JORDAN is a rationally designed system for high throughput in vivo LNP analysis. (a) 240 barcode sequences were chosen to comply with the Illumina MiniSeq machine. (b) Primers used to amplify barcodes from in vitro and in vivo samples. (c) Data from all experiments was normalized using the procedure shown. (d) Pooled LNPs were stable, had good autocorrelation curves, and single-peak diameter distributions between 20 and 300 nm. All other LNPs were discarded. (e) Alexa Fluor 647 tagged DNA barcode localizes to cell surface after 15 minutes and (f) can be seen in the cell cytoplasm after 72 hours.

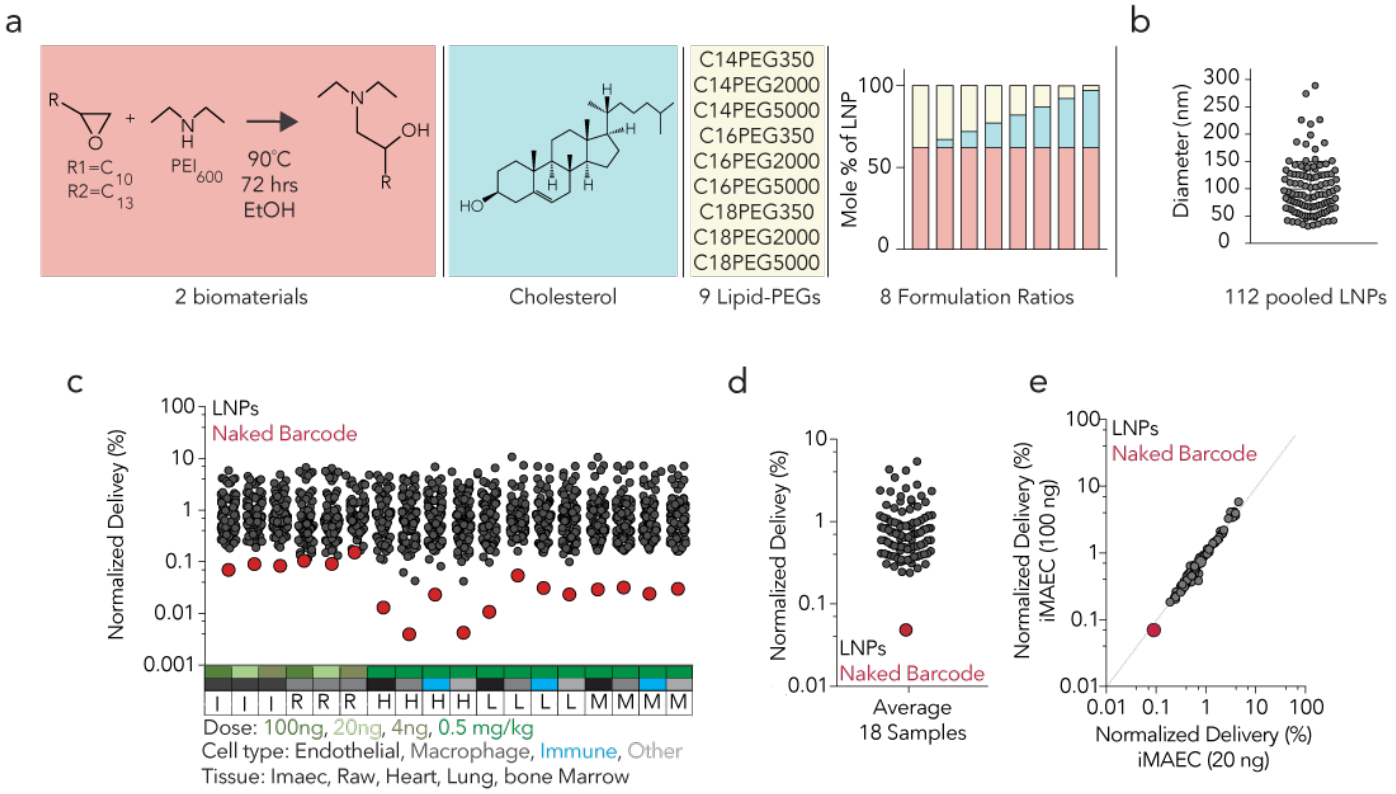
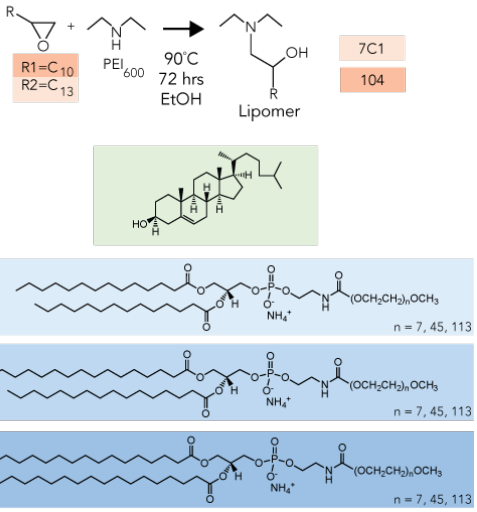


Figure 2.3. To analyze the robustness of our system (a) We formulated 144 chemically distinct LNPs, (b) pooled stable LNPs and administered them to two cell lines (RAWs, iMAECs) and mice. (c) Normalized delivery for all LNPs and naked barcode in 18 cell and tissue types. (d) Average normalized delivery for all 18 samples. The naked barcode delivered DNA less efficiently than all LNPs. (e) Normalized delivery in iMAECs 72 hours after 20 ng or 100 ng total DNA was administered. *In vitro* delivery to iMAECs at 20 ng/well predicted *in vitro* delivery to iMAECs at 100 ng/well.

a

LNP #	Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %
1	7C1	PEI 600	C15 Epoxy	10	C14	350	62	0	38
2	7C1	PEI 600	C15 Epoxy	10	C14	350	62	5	33
3	7C1	PEI 600	C15 Epoxy	10	C14	350	62	10	28
4	7C1	PEI 600	C15 Epoxy	10	C14	350	62	15	23
5	7C1	PEI 600	C15 Epoxy	10	C14	350	62	20	18
6	7C1	PEI 600	C15 Epoxy	10	C14	350	62	25	13
7	7C1	PEI 600	C15 Epoxy	10	C14	350	62	30	8
8	7C1	PEI 600	C15 Epoxy	10	C14	350	62	35	3
9	104	PEI 600	C12 Epoxy	10	C14	350	62	0	38
10	104	PEI 600	C12 Epoxy	10	C14	350	62	5	33
11	104	PEI 600	C12 Epoxy	10	C14	350	62	10	28
12	104	PEI 600	C12 Epoxy	10	C14	350	62	15	23
13	104	PEI 600	C12 Epoxy	10	C14	350	62	20	18
14	104	PEI 600	C12 Epoxy	10	C14	350	62	25	13
15	104	PEI 600	C12 Epoxy	10	C14	350	62	30	8
16	104	PEI 600	C12 Epoxy	10	C14	350	62	35	3
17	7C1	PEI 600	C15 Epoxy	10	C16	350	62	0	38
18	7C1	PEI 600	C15 Epoxy	10	C16	350	62	5	33
19	7C1	PEI 600	C15 Epoxy	10	C16	350	62	10	28
20	7C1	PEI 600	C15 Epoxy	10	C16	350	62	15	23
21	7C1	PEI 600	C15 Epoxy	10	C16	350	62	20	18
22	7C1	PEI 600	C15 Epoxy	10	C16	350	62	25	13
23	7C1	PEI 600	C15 Epoxy	10	C16	350	62	30	8
24	7C1	PEI 600	C15 Epoxy	10	C16	350	62	35	3
25	104	PEI 600	C12 Epoxy	10	C16	350	62	0	38
26	104	PEI 600	C12 Epoxy	10	C16	350	62	5	33
27	104	PEI 600	C12 Epoxy	10	C16	350	62	10	28
28	104	PEI 600	C12 Epoxy	10	C16	350	62	15	23
29	104	PEI 600	C12 Epoxy	10	C16	350	62	20	18
30	104	PEI 600	C12 Epoxy	10	C16	350	62	25	13
31	104	PEI 600	C12 Epoxy	10	C16	350	62	30	8
32	104	PEI 600	C12 Epoxy	10	C16	350	62	35	3
33	7C1	PEI 600	C15 Epoxy	10	C18	350	62	0	38
34	7C1	PEI 600	C15 Epoxy	10	C18	350	62	5	33
35	7C1	PEI 600	C15 Epoxy	10	C18	350	62	10	28
36	7C1	PEI 600	C15 Epoxy	10	C18	350	62	15	23
37	7C1	PEI 600	C15 Epoxy	10	C18	350	62	20	18
38	7C1	PEI 600	C15 Epoxy	10	C18	350	62	25	13
39	7C1	PEI 600	C15 Epoxy	10	C18	350	62	30	8
40	7C1	PEI 600	C15 Epoxy	10	C18	350	62	35	3
41	104	PEI 600	C12 Epoxy	10	C18	350	62	0	38
42	104	PEI 600	C12 Epoxy	10	C18	350	62	5	33
43	104	PEI 600	C12 Epoxy	10	C18	350	62	10	28
44	104	PEI 600	C12 Epoxy	10	C18	350	62	15	23
45	104	PEI 600	C12 Epoxy	10	C18	350	62	20	18
46	104	PEI 600	C12 Epoxy	10	C18	350	62	25	13
47	104	PEI 600	C12 Epoxy	10	C18	350	62	30	8
48	104	PEI 600	C12 Epoxy	10	C18	350	62	35	3
49	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	0	38
50	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	5	33
51	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	10	28
52	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	15	23
53	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	20	18
54	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	25	13
55	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	30	8
56	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	35	3
57	104	PEI 600	C12 Epoxy	10	C14	2000	62	0	38
58	104	PEI 600	C12 Epoxy	10	C14	2000	62	5	33
59	104	PEI 600	C12 Epoxy	10	C14	2000	62	10	28
60	104	PEI 600	C12 Epoxy	10	C14	2000	62	15	23
61	104	PEI 600	C12 Epoxy	10	C14	2000	62	20	18
62	104	PEI 600	C12 Epoxy	10	C14	2000	62	25	13
63	104	PEI 600	C12 Epoxy	10	C14	2000	62	30	8
64	104	PEI 600	C12 Epoxy	10	C14	2000	62	35	3
65	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	0	38
66	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	5	33
67	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	10	28
68	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	15	23
69	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	20	18
70	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	25	13
71	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	30	8
72	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	35	3
73	104	PEI 600	C12 Epoxy	10	C16	2000	62	0	38
74	104	PEI 600	C12 Epoxy	10	C16	2000	62	5	33
75	104	PEI 600	C12 Epoxy	10	C16	2000	62	10	28
76	104	PEI 600	C12 Epoxy	10	C16	2000	62	15	23
77	104	PEI 600	C12 Epoxy	10	C16	2000	62	20	18
78	104	PEI 600	C12 Epoxy	10	C16	2000	62	25	13
79	104	PEI 600	C12 Epoxy	10	C16	2000	62	30	8
80	104	PEI 600	C12 Epoxy	10	C16	2000	62	35	3
81	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	0	38
82	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	5	33
83	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	10	28
84	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	15	23
85	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	20	18
86	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	25	13
87	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	30	8
88	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	35	3
89	104	PEI 600	C12 Epoxy	10	C18	2000	62	0	38
90	104	PEI 600	C12 Epoxy	10	C18	2000	62	5	33
91	104	PEI 600	C12 Epoxy	10	C18	2000	62	10	28
92	104	PEI 600	C12 Epoxy	10	C18	2000	62	15	23
93	104	PEI 600	C12 Epoxy	10	C18	2000	62	20	18
94	104	PEI 600	C12 Epoxy	10	C18	2000	62	25	13
95	104	PEI 600	C12 Epoxy	10	C18	2000	62	30	8
96	104	PEI 600	C12 Epoxy	10	C18	2000	62	35	3
97	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	0	38
98	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	5	33
99	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	10	28
100	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	15	23
101	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	20	18
102	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	25	13
103	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	30	8
104	7C1	PEI 600	C15 Epoxy	10	C14	5000	62	35	3
105	104	PEI 600	C12 Epoxy	10	C14	5000	62	0	38
106	104	PEI 600	C12 Epoxy	10	C14	5000	62	5	33
107	104	PEI 600	C12 Epoxy	10	C14	5000	62	10	28
108	104	PEI 600	C12 Epoxy	10	C14	5000	62	15	23
109	104	PEI 600	C12 Epoxy	10	C14	5000	62	20	18
110	104	PEI 600	C12 Epoxy	10	C14	5000	62	25	13
111	104	PEI 600	C12 Epoxy	10	C14	5000	62	30	8
112	104	PEI 600	C12 Epoxy	10	C14	5000	62	35	3
113	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	0	38
114	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	5	33
115	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	10	28
116	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	15	23
117	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	20	18
118	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	25	13
119	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	30	8
120	7C1	PEI 600	C15 Epoxy	10	C16	5000	62	35	3
121	104	PEI 600	C12 Epoxy	10	C16	5000	62	0	38
122	104	PEI 600	C12 Epoxy	10	C16	5000	62	5	33
123	104	PEI 600	C12 Epoxy	10	C16	5000	62	10	28
124	104	PEI 600	C12 Epoxy	10	C16	5000	62	15	23
125	104	PEI 600	C12 Epoxy	10	C16	5000	62	20	18
126	104	PEI 600	C12 Epoxy	10	C16	5000	62	25	13
127	104	PEI 600	C12 Epoxy	10	C16	5000	62	30	8
128	104	PEI 600	C12 Epoxy	10	C16	5000	62	35	3
129	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	0	38
130	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	5	33
131	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	10	28
132	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	15	23
133	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	20	18
134	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	25	13
135	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	30	8
136	7C1	PEI 600	C15 Epoxy	10	C18	5000	62	35	3
137	104	PEI 600	C12 Epoxy	10	C18	5000	62	0	38
138	104	PEI 600	C12 Epoxy	10	C18	5000	62	5	33
139	104	PEI 600	C12 Epoxy	10	C18	5000	62	10	28
140	104	PEI 600	C12 Epoxy	10	C18	5000	62	15	23
141	104	PEI 600	C12 Epoxy	10	C18	5000	62	20	18
142	104	PEI 600	C12 Epoxy	10	C18	5000	62	25	13
143	104	PEI 600	C12 Epoxy	10	C18	5000	62	30	8
144	104	PEI 600	C12 Epoxy	10	C18	5000	62	35	3



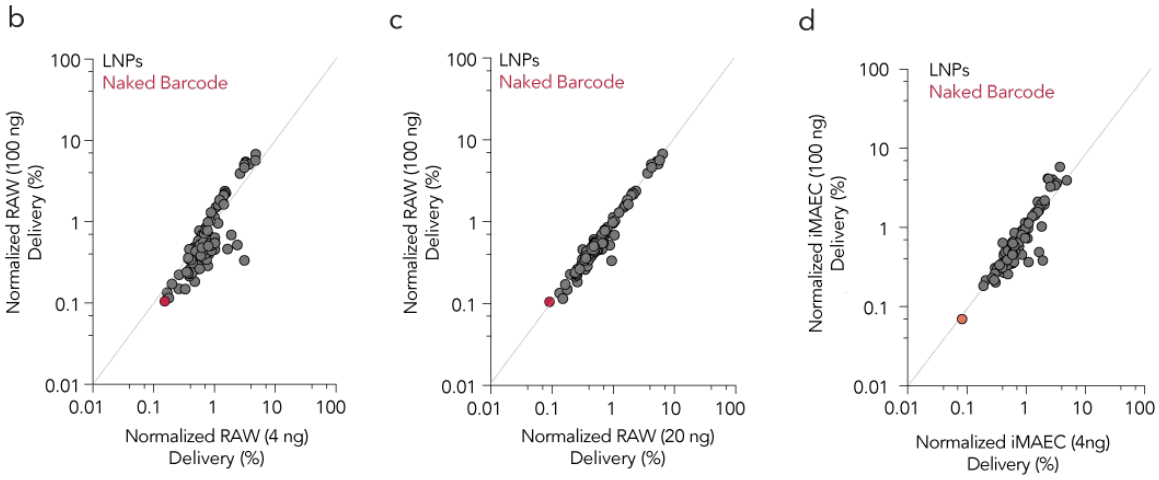


Figure 2.4 The data generated by JORDAN are robust. (a) Library 1 was designed to vary PEG chain length (C14, C16, C18) and PEG MW (350, 2000, 5000), using two different lipomers. Chemical formulas for lipomers as well as all other LNP components are shown. (b,c) Normalized delivery in RAWs 72 hours after 4 ng, 20 ng, or 100 ng total DNA was administered. (d) Normalized delivery in IMAECs 72 hours after 4 ng or 100 ng total DNA was administered.

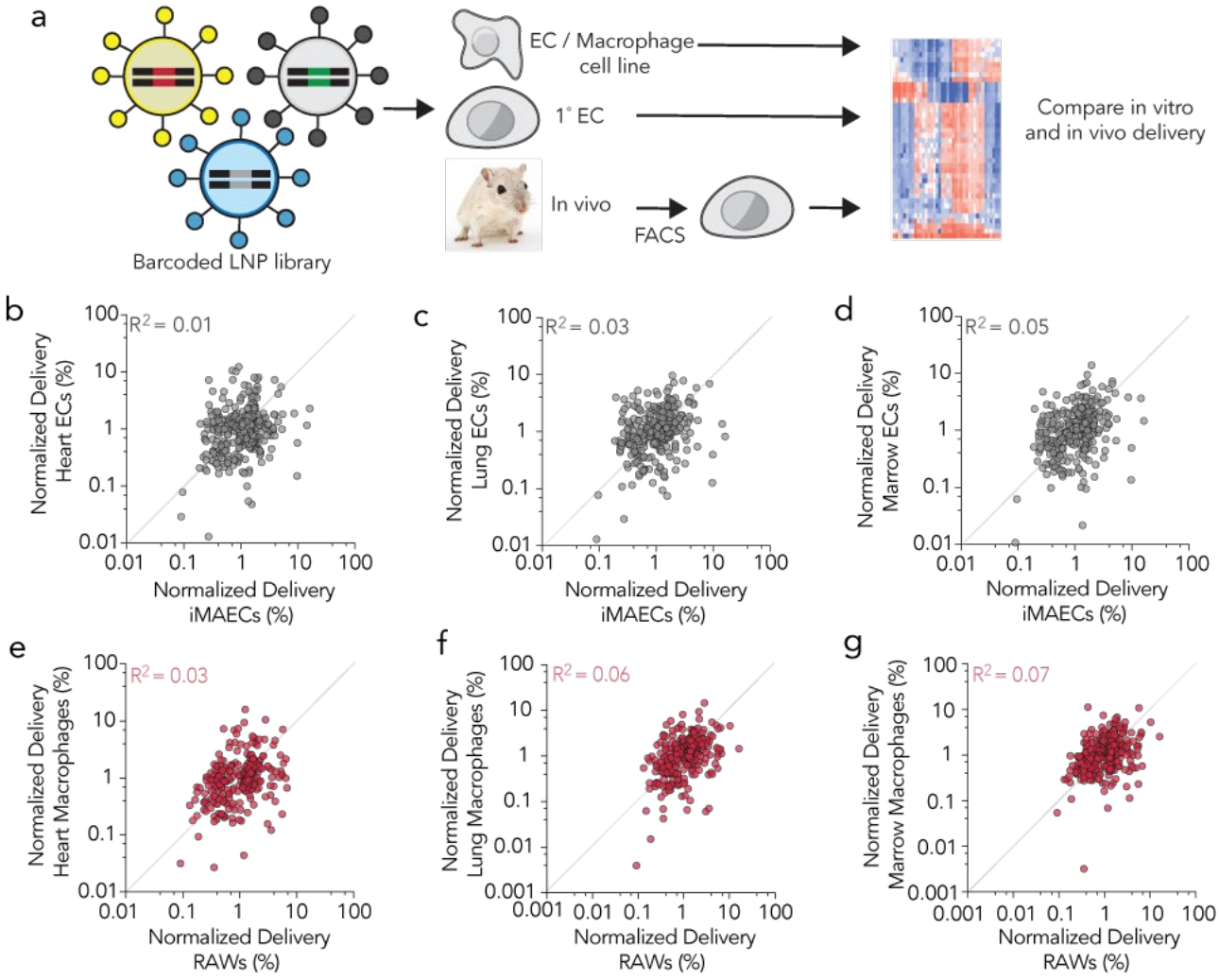
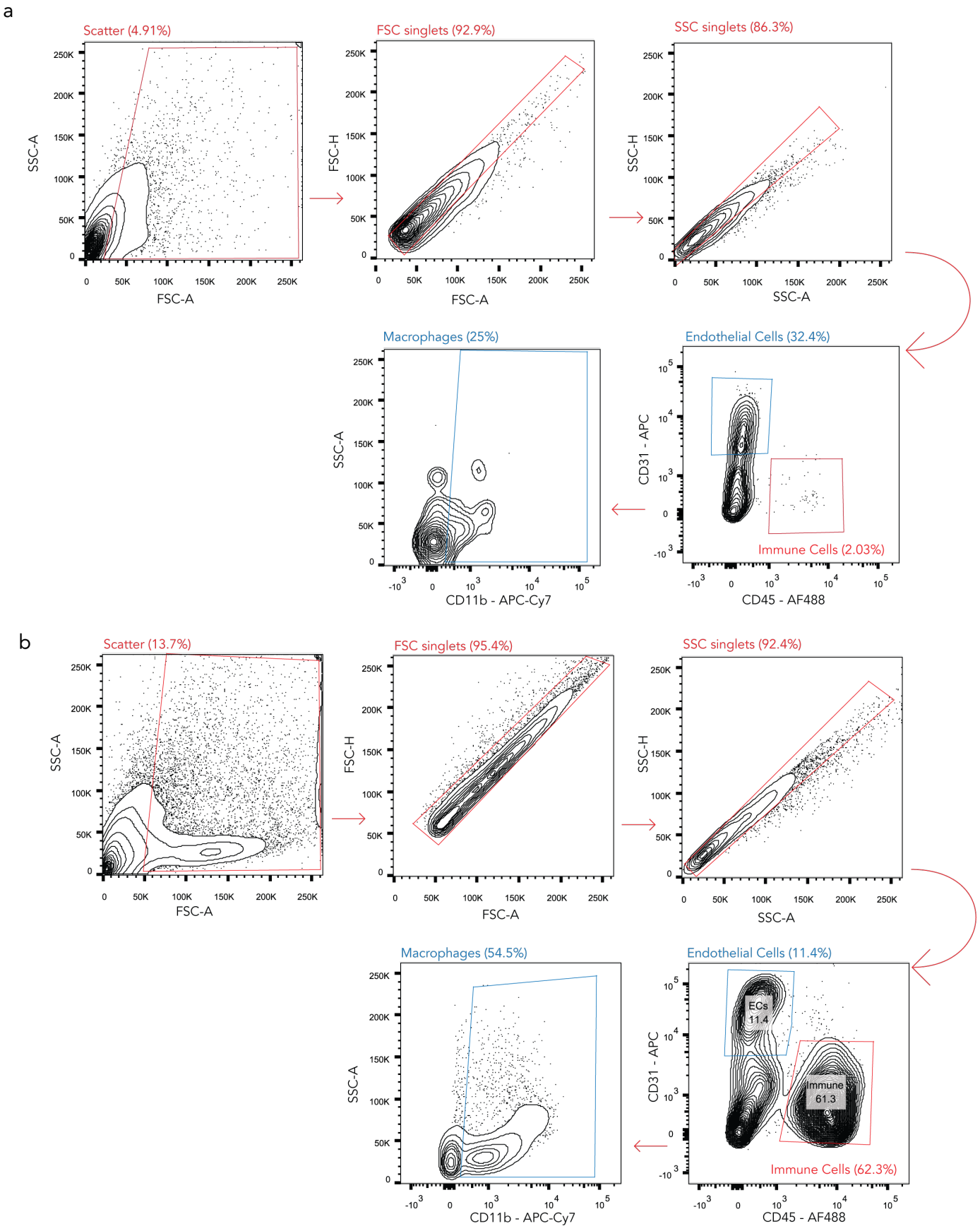
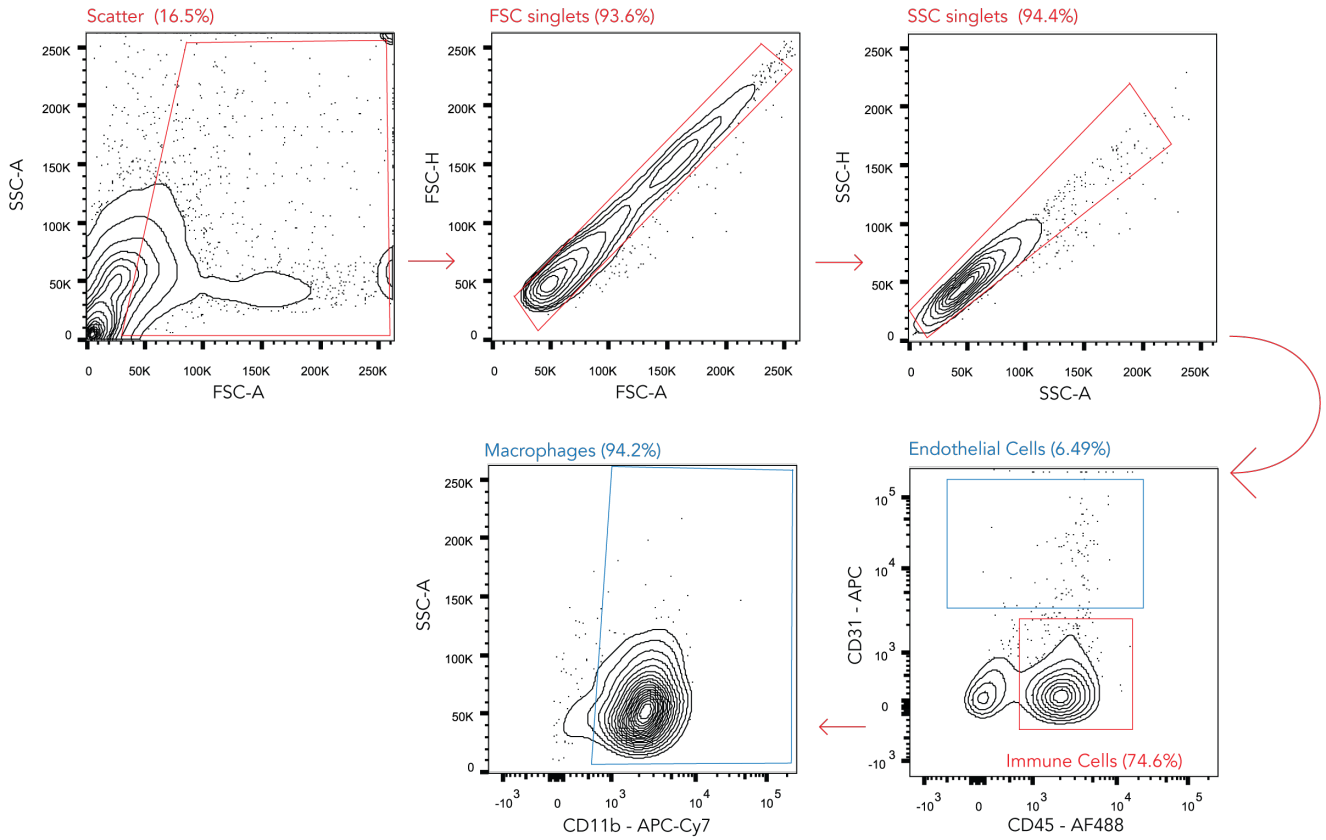


Figure 2.5 A direct comparison between in vitro and in vivo nanoparticle delivery. (a) 420 LNPs were formulated and delivery was compared between in vivo FACS sorted cells, primary cells, and cell lines. (b-d) Normalized delivery of LNPs in iMAECs and heart, lung, and bone marrow endothelial cells. (e-g) Normalized delivery of LNPs in RAWs and heart, lung, and bone marrow macrophages. In both cases, in vitro LNP delivery does not predict in vivo delivery.

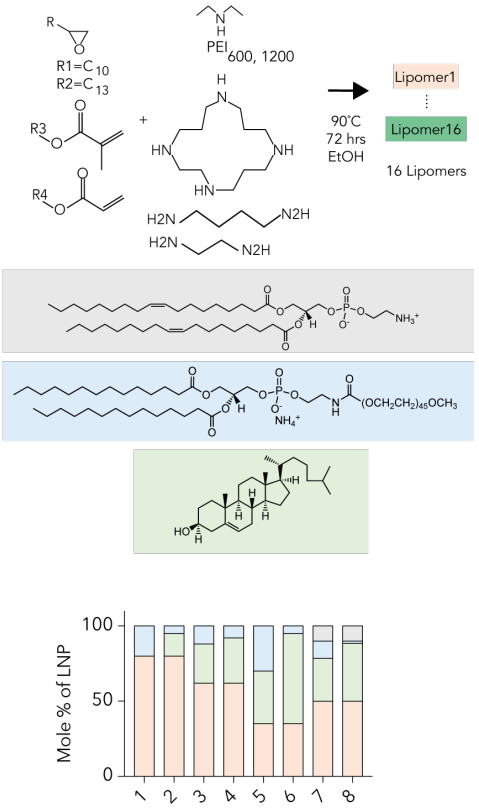

a
Scatter (4.91%)
SSC-A
FSC-A
FSC singlets (92.9%)
FSC-H
FSC-A
SSC singlets (86.3%)
SSC-H
SSC-A
Macrophages (25%)
SSC-A
CD11b - APC-Cy7
Endothelial Cells (32.4%)
CD31 - APC
Immune Cells (2.03%)
CD45 - AF488
b
Scatter (13.7%)
SSC-A
FSC-A
FSC singlets (95.4%)
FSC-H
FSC-A
SSC singlets (92.4%)
SSC-H
SSC-A
Macrophages (54.5%)
SSC-A
CD11b - APC-Cy7
Endothelial Cells (11.4%)
ECs
11.4
Immune
61.3
CD31 - APC
Immune Cells (62.3%)
CD45 - AF488

c

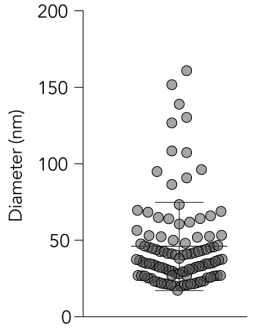

Scatter (16.5%)
SSC-A
FSC-A
FSC singlets (93.6%)
FSC-H
FSC-A
SSC singlets (94.4%)
SSC-H
SSC-A
Endothelial Cells (6.49%)
CD31 - APC
Immune Cells (74.6%)
CD45 - AF488
Macrophages (94.2%)
SSC-A
CD11b - APC-Cy7

d

LNP #	Lipomer	Amine	Lipid	Helper Lipid	Mass	PEG	PEG MW	Lipomer Mole	Cholesterol Mole	PEG Mole %	Helper Lipid Mole %
1	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	80	0	20	0
2	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	80	15	5	0
3	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	62	26	12	0
4	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	62	30	8	0
5	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	35	35	30	0
6	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	35	60	5	0
7	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
8	7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
9	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
10	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
11	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
12	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
13	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
14	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
15	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
16	8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
17	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	80	0	20	0
18	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	80	15	5	0
19	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	62	26	12	0
20	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	62	30	8	0
21	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	35	35	30	0
22	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	35	60	5	0
23	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	50	28.5	11.5	10
24	69	PEI 600	Stearyl Methacrylate	DOPE	10	C14	2000	50	38.5	1.5	10
25	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
26	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
27	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
28	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
29	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
30	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
31	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
32	104	PEI 600	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
33	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
34	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
35	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
36	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
37	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
38	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
39	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
40	48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
41	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
42	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
43	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
44	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
45	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
46	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
47	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
48	92	1,4,8,12-Tetraazacyclopentadecane	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
49	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
50	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
51	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
52	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
53	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
54	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
55	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
56	58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
57	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
58	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
59	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
60	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
61	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
62	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
63	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
64	100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
65	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
66	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
67	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
68	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
69	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
70	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
71	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
72	78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
73	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	80	0	20	0
74	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	80	15	5	0
75	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	62	26	12	0
76	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	62	30	8	0
77	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	35	35	30	0
78	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	35	60	5	0
79	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	50	28.5	11.5	10
80	79	PEI 1200	Stearyl Methacrylate	DOPE	10	C14	2000	50	38.5	1.5	10
81	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
82	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
83	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
84	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
85	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
86	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
87	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
88	102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
89	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
90	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
91	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
92	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
93	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
94	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
95	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
96	68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
97	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
98	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
99	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
100	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
101	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
102	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
103	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
104	106	PEI 1200	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10
105	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0
106	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0
107	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	62	26	12	0
108	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0
109	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0
110	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0
111	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10
112	88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10
113	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	80	0	20	0
114	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0
115	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	62	26	12	0
116	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0
117	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0
118	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	35	60	5	0
119	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	50	28.5	11.5	10
120	108	Triethylenetetramine	C12 Epoxy	DOPE	10	C14	2000	50	38.5	1.5	10

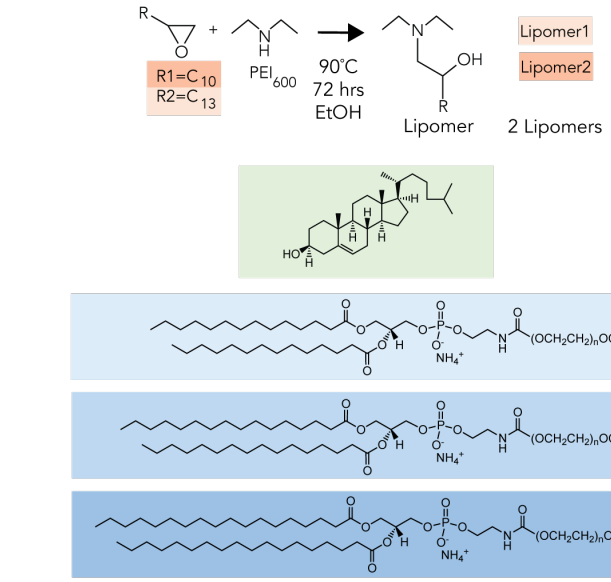


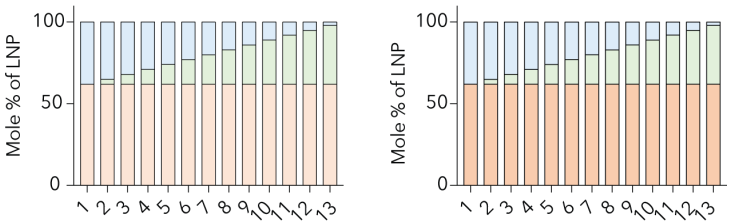
e



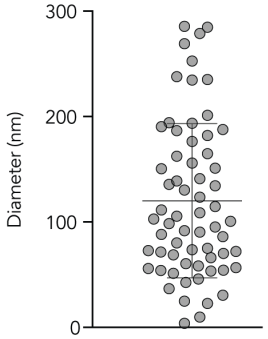
f

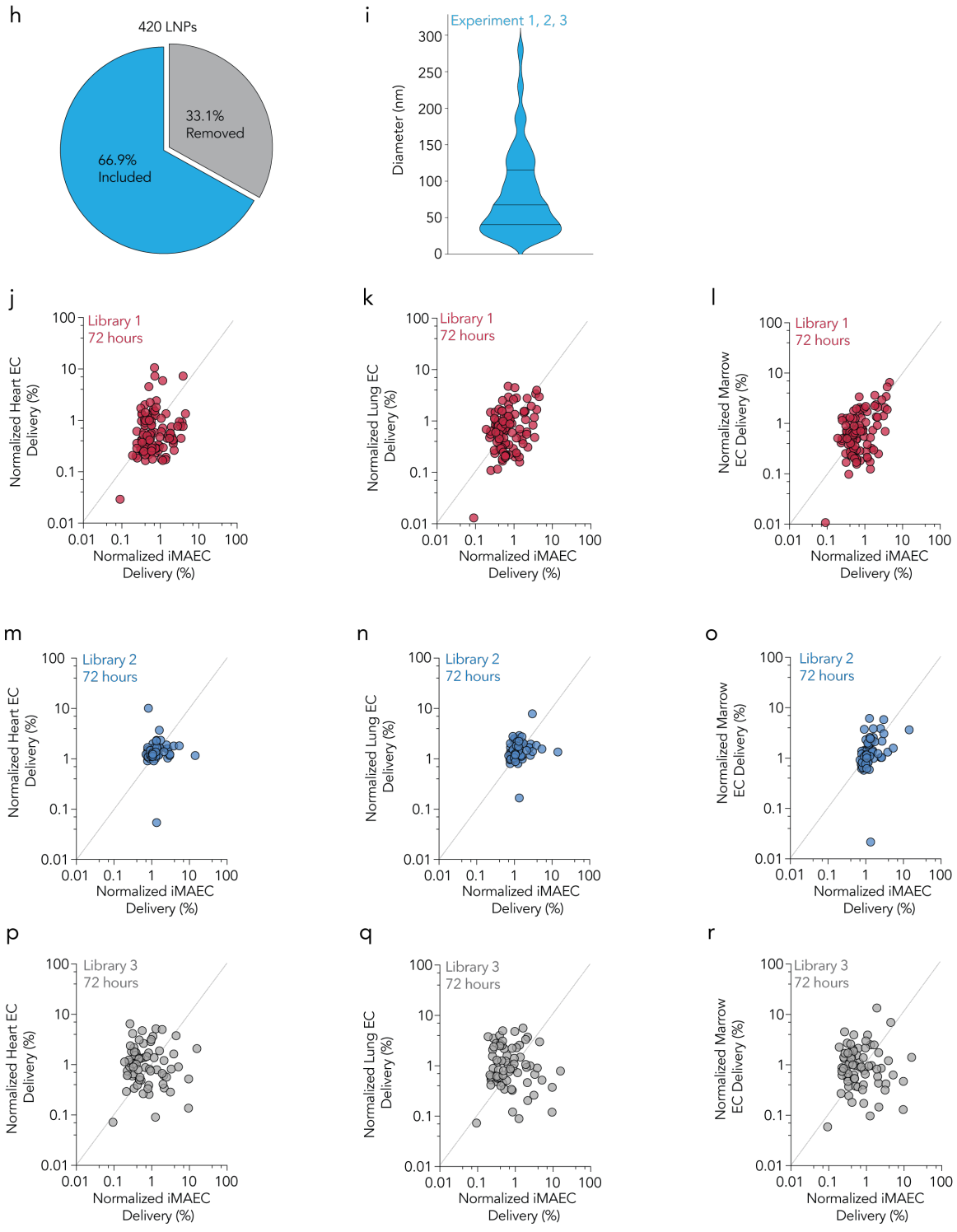
LNP #	Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %
1	104	PEI 600	C12 Epoxy	10	C14	2000	62	0	38
2	104	PEI 600	C12 Epoxy	10	C14	2000	62	3	35
3	104	PEI 600	C12 Epoxy	10	C14	2000	62	6	32
4	104	PEI 600	C12 Epoxy	10	C14	2000	62	9	29
5	104	PEI 600	C12 Epoxy	10	C14	2000	62	12	26
6	104	PEI 600	C12 Epoxy	10	C14	2000	62	15	23
7	104	PEI 600	C12 Epoxy	10	C14	2000	62	18	20
8	104	PEI 600	C12 Epoxy	10	C14	2000	62	21	17
9	104	PEI 600	C12 Epoxy	10	C14	2000	62	24	14
10	104	PEI 600	C12 Epoxy	10	C14	2000	62	27	11
11	104	PEI 600	C12 Epoxy	10	C14	2000	62	30	8
12	104	PEI 600	C12 Epoxy	10	C14	2000	62	33	5
13	104	PEI 600	C12 Epoxy	10	C14	2000	62	36	2
14	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	0	38
15	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	3	35
16	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	6	32
17	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	9	29
18	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	12	26
19	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	15	23
20	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	18	20
21	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	21	17
22	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	24	14
23	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	27	11
24	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	30	8
25	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	33	5
26	7C1	PEI 600	C15 Epoxy	10	C14	2000	62	36	2
27	104	PEI 600	C12 Epoxy	10	C16	2000	62	0	38
28	104	PEI 600	C12 Epoxy	10	C16	2000	62	3	35
29	104	PEI 600	C12 Epoxy	10	C16	2000	62	6	32
30	104	PEI 600	C12 Epoxy	10	C16	2000	62	9	29
31	104	PEI 600	C12 Epoxy	10	C16	2000	62	12	26
32	104	PEI 600	C12 Epoxy	10	C16	2000	62	15	23
33	104	PEI 600	C12 Epoxy	10	C16	2000	62	18	20
34	104	PEI 600	C12 Epoxy	10	C16	2000	62	21	17
35	104	PEI 600	C12 Epoxy	10	C16	2000	62	24	14
36	104	PEI 600	C12 Epoxy	10	C16	2000	62	27	11
37	104	PEI 600	C12 Epoxy	10	C16	2000	62	30	8
38	104	PEI 600	C12 Epoxy	10	C16	2000	62	33	5
39	104	PEI 600	C12 Epoxy	10	C16	2000	62	36	2
40	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	0	38
41	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	3	35
42	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	6	32
43	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	9	29
44	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	12	26
45	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	15	23
46	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	18	20
47	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	21	17
48	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	24	14
49	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	27	11
50	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	30	8
51	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	33	5
52	7C1	PEI 600	C15 Epoxy	10	C16	2000	62	36	2
53	104	PEI 600	C12 Epoxy	10	C18	2000	62	0	38
54	104	PEI 600	C12 Epoxy	10	C18	2000	62	3	35
55	104	PEI 600	C12 Epoxy	10	C18	2000	62	6	32
56	104	PEI 600	C12 Epoxy	10	C18	2000	62	9	29
57	104	PEI 600	C12 Epoxy	10	C18	2000	62	12	26
58	104	PEI 600	C12 Epoxy	10	C18	2000	62	15	23
59	104	PEI 600	C12 Epoxy	10	C18	2000	62	18	20
60	104	PEI 600	C12 Epoxy	10	C18	2000	62	21	17
61	104	PEI 600	C12 Epoxy	10	C18	2000	62	24	14
62	104	PEI 600	C12 Epoxy	10	C18	2000	62	27	11
63	104	PEI 600	C12 Epoxy	10	C18	2000	62	30	8
64	104	PEI 600	C12 Epoxy	10	C18	2000	62	33	5
65	104	PEI 600	C12 Epoxy	10	C18	2000	62	36	2
66	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	0	38
67	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	3	35
68	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	6	32
69	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	9	29
70	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	12	26
71	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	15	23
72	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	18	20
73	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	21	17
74	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	24	14
75	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	27	11
76	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	30	8
77	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	33	5
78	7C1	PEI 600	C15 Epoxy	10	C18	2000	62	36	2
79	104	PEI 600	C12 Epoxy	10	C14	350	62	0	38
80	104	PEI 600	C12 Epoxy	10	C14	350	62	3	35
81	104	PEI 600	C12 Epoxy	10	C14	350	62	6	32
82	104	PEI 600	C12 Epoxy	10	C14	350	62	9	29
83	104	PEI 600	C12 Epoxy	10	C14	350	62	12	26
84	104	PEI 600	C12 Epoxy	10	C14	350	62	15	23
85	104	PEI 600	C12 Epoxy	10	C14	350	62	18	20
86	104	PEI 600	C12 Epoxy	10	C14	350	62	21	17
87	104	PEI 600	C12 Epoxy	10	C14	350	62	24	14
88	104	PEI 600	C12 Epoxy	10	C14	350	62	27	11
89	104	PEI 600	C12 Epoxy	10	C14	350	62	30	8
90	104	PEI 600	C12 Epoxy	10	C14	350	62	33	5
91	104	PEI 600	C12 Epoxy	10	C14	350	62	36	2
92	7C1	PEI 600	C15 Epoxy	10	C14	350	62	0	38
93	7C1	PEI 600	C15 Epoxy	10	C14	350	62	3	35
94	7C1	PEI 600	C15 Epoxy	10	C14	350	62	6	32
95	7C1	PEI 600	C15 Epoxy	10	C14	350	62	9	29
96	7C1	PEI 600	C15 Epoxy	10	C14	350	62	12	26
97	7C1	PEI 600	C15 Epoxy	10	C14	350	62	15	23
98	7C1	PEI 600	C15 Epoxy	10	C14	350	62	18	20
99	7C1	PEI 600	C15 Epoxy	10	C14	350	62	21	17
100	7C1	PEI 600	C15 Epoxy	10	C14	350	62	24	14
101	7C1	PEI 600	C15 Epoxy	10	C14	350	62	27	11
102	7C1	PEI 600	C15 Epoxy	10	C14	350	62	30	8
103	7C1	PEI 600	C15 Epoxy	10	C14	350	62	33	5
104	7C1	PEI 600	C15 Epoxy	10	C14	350	62	36	2
105	104	PEI 600	C12 Epoxy	10	C16	350	62	0	38
106	104	PEI 600	C12 Epoxy	10	C16	350	62	3	35
107	104	PEI 600	C12 Epoxy	10	C16	350	62	6	32
108	104	PEI 600	C12 Epoxy	10	C16	350	62	9	29
109	104	PEI 600	C12 Epoxy	10	C16	350	62	12	26
110	104	PEI 600	C12 Epoxy	10	C16	350	62	15	23
111	104	PEI 600	C12 Epoxy	10	C16	350	62	18	20
112	104	PEI 600	C12 Epoxy	10	C16	350	62	21	17
113	104	PEI 600	C12 Epoxy	10	C16	350	62	24	14
114	104	PEI 600	C12 Epoxy	10	C16	350	62	27	11
115	104	PEI 600	C12 Epoxy	10	C16	350	62	30	8
116	104	PEI 600	C12 Epoxy	10	C16	350	62	33	5
117	104	PEI 600	C12 Epoxy	10	C16	350	62	36	2
118	7C1	PEI 600	C15 Epoxy	10	C16	350	62	0	38
119	7C1	PEI 600	C15 Epoxy	10	C16	350	62	3	35
120	7C1	PEI 600	C15 Epoxy	10	C16	350	62	6	32
121	7C1	PEI 600	C15 Epoxy	10	C16	350	62	9	29
122	7C1	PEI 600	C15 Epoxy	10	C16	350	62	12	26
123	7C1	PEI 600	C15 Epoxy	10	C16	350	62	15	23
124	7C1	PEI 600	C15 Epoxy	10	C16	350	62	18	20
125	7C1	PEI 600	C15 Epoxy	10	C16	350	62	21	17
126	7C1	PEI 600	C15 Epoxy	10	C16	350	62	24	14
127	7C1	PEI 600	C15 Epoxy	10	C16	350	62	27	11
128	7C1	PEI 600	C15 Epoxy	10	C16	350	62	30	8
129	7C1	PEI 600	C15 Epoxy	10	C16	350	62	33	5
130	7C1	PEI 600	C15 Epoxy	10	C16	350	62	36	2
131	104	PEI 600	C12 Epoxy	10	C18	350	62	0	38
132	104	PEI 600	C12 Epoxy	10	C18	350	62	3	35
133	104	PEI 600	C12 Epoxy	10	C18	350	62	6	32
134	104	PEI 600	C12 Epoxy	10	C18	350	62	9	29
135	104	PEI 600	C12 Epoxy	10	C18	350	62	12	26
136	104	PEI 600	C12 Epoxy	10	C18	350	62	15	23
137	104	PEI 600	C12 Epoxy	10	C18	350	62	18	20
138	104	PEI 600	C12 Epoxy	10	C18	350	62	21	17
139	104	PEI 600	C12 Epoxy	10	C18	350	62	24	14
140	104	PEI 600	C12 Epoxy	10	C18	350	62	27	11
141	104	PEI 600	C12 Epoxy	10	C18	350	62	30	8
142	104	PEI 600	C12 Epoxy	10	C18	350	62	33	5
143	104	PEI 600	C12 Epoxy	10	C18	350	62	36	2
144	7C1	PEI 600	C15 Epoxy	10	C18	350	62	0	38
145	7C1	PEI 600	C15 Epoxy	10	C18	350	62	3	35
146	7C1	PEI 600	C15 Epoxy	10	C18	350	62	6	32
147	7C1	PEI 600	C15 Epoxy	10	C18	350	62	9	29
148	7C1	PEI 600	C15 Epoxy	10	C18	350	62	12	26
149	7C1	PEI 600	C15 Epoxy	10	C18	350	62	15	23
150	7C1	PEI 600	C15 Epoxy	10	C18	350	62	18	20
151	7C1	PEI 600	C15 Epoxy	10	C18	350	62	21	17
152	7C1	PEI 600	C15 Epoxy	10	C18	350	62	24	14
153	7C1	PEI 600	C15 Epoxy	10	C18	350	62	27	11
154	7C1	PEI 600	C15 Epoxy	10	C18	350	62	30	8
155	7C1	PEI 600	C15 Epoxy	10	C18	350	62	33	5
156	7C1	PEI 600	C15 Epoxy	10	C18	350	62	36	2



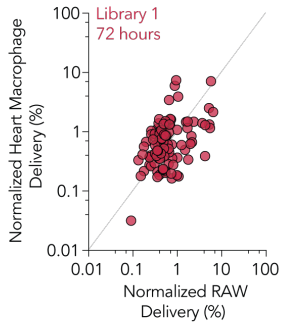


g

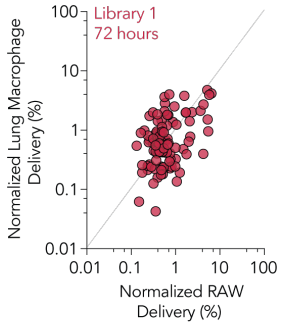



h
420 LNPs
33.1%
Removed
66.9%
Included
i
Experiment 1, 2, 3
Diameter (nm)
300
250
200
150
100
50
0
j
Library 1
72 hours
Normalized Heart EC
Delivery (%)
k
Normalized Lung EC
Delivery (%)
l
Normalized Marrow
EC Delivery (%)
Normalized iMAEC
Delivery (%)
m
Library 2
72 hours
n
o
p
Library 3
72 hours
q
r
100
10
1
0.1
0.01
0.01
0.1
1
10
100

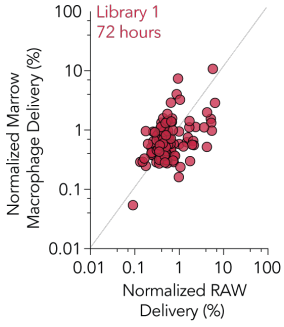
s


Library 1
72 hours
Normalized Heart Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

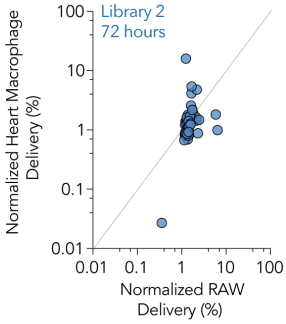
t


Library 1
72 hours
Normalized Lung Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

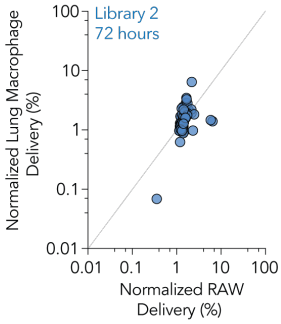
u


Library 1
72 hours
Normalized Marrow Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

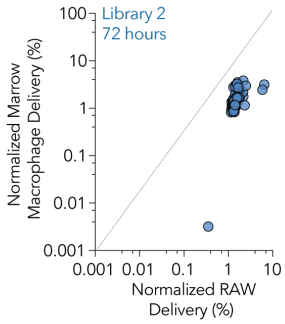
v


Library 2
72 hours
Normalized Heart Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

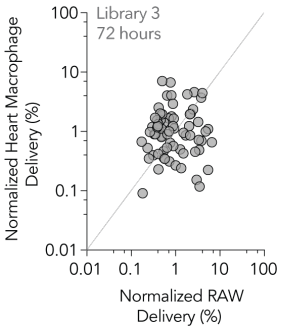
w


Library 2
72 hours
Normalized Lung Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

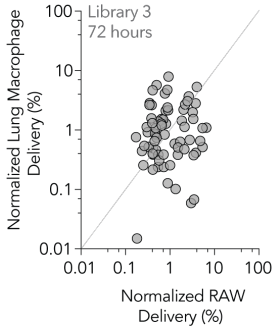
x


Library 2
72 hours
Normalized Marrow Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.001
0.001
0.01
0.1
1
10

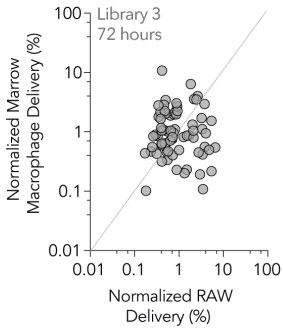
y

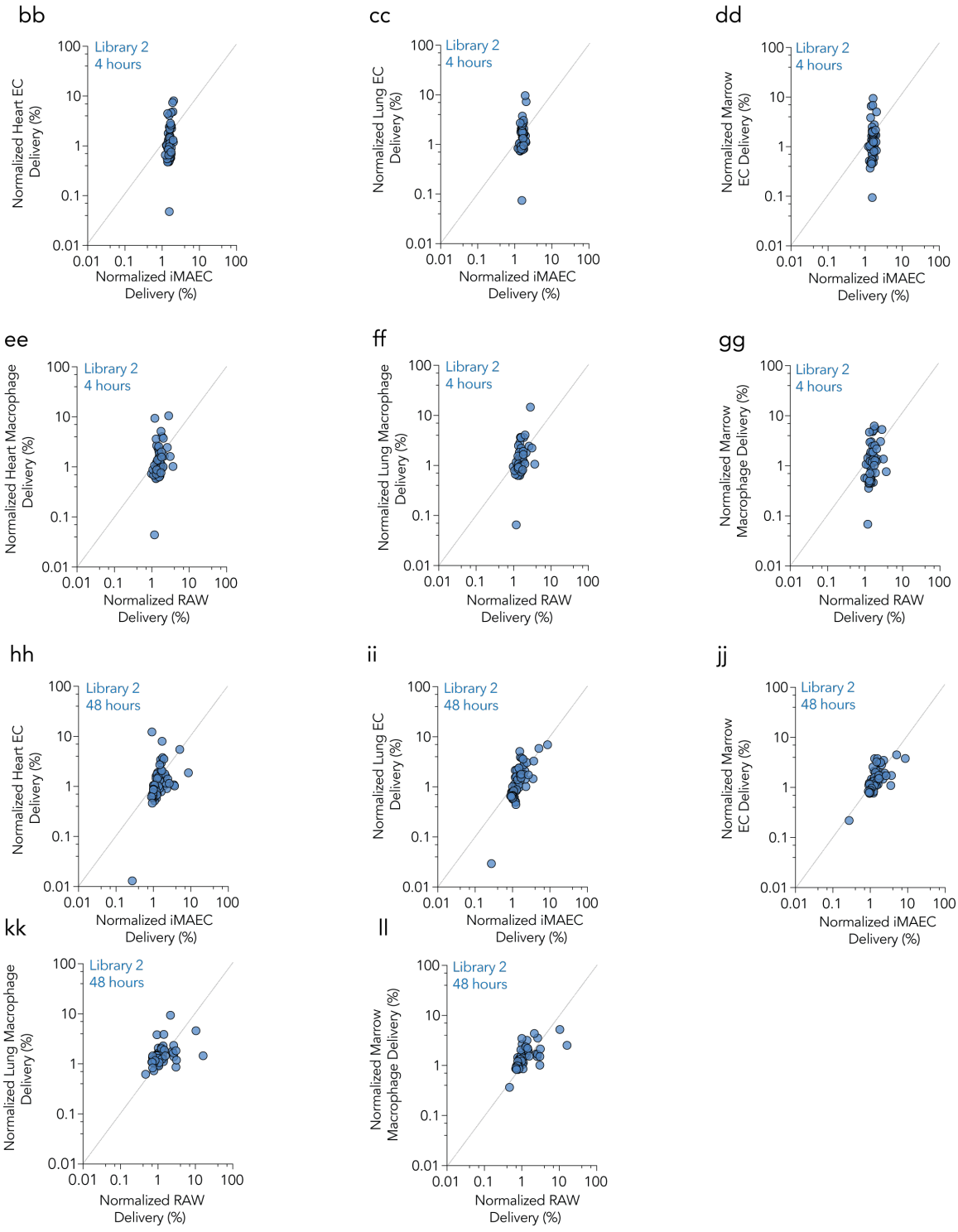

Library 3
72 hours
Normalized Heart Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

z


Library 3
72 hours
Normalized Lung Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100

aa


Library 3
72 hours
Normalized Marrow Macrophage Delivery (%)
Normalized RAW Delivery (%)
100
10
1
0.1
0.01
0.01
0.1
1
10
100


bb
Library 2
4 hours
Normalized Heart EC Delivery (%)
Normalized iMAEC Delivery (%)
cc
Library 2
4 hours
Normalized Lung EC Delivery (%)
Normalized iMAEC Delivery (%)
dd
Library 2
4 hours
Normalized Marrow EC Delivery (%)
Normalized iMAEC Delivery (%)
ee
Library 2
4 hours
Normalized Heart Macrophage Delivery (%)
Normalized RAW Delivery (%)
ff
Library 2
4 hours
Normalized Lung Macrophage Delivery (%)
Normalized RAW Delivery (%)
gg
Library 2
4 hours
Normalized Marrow Macrophage Delivery (%)
Normalized RAW Delivery (%)
hh
Library 2
48 hours
Normalized Heart EC Delivery (%)
Normalized iMAEC Delivery (%)
ii
Library 2
48 hours
Normalized Lung EC Delivery (%)
Normalized iMAEC Delivery (%)
jj
Library 2
48 hours
Normalized Marrow EC Delivery (%)
Normalized iMAEC Delivery (%)
kk
Library 2
48 hours
Normalized Lung Macrophage Delivery (%)
Normalized RAW Delivery (%)
ll
Library 2
48 hours
Normalized Marrow Macrophage Delivery (%)
Normalized RAW Delivery (%)
0.01
0.1
1
10
100

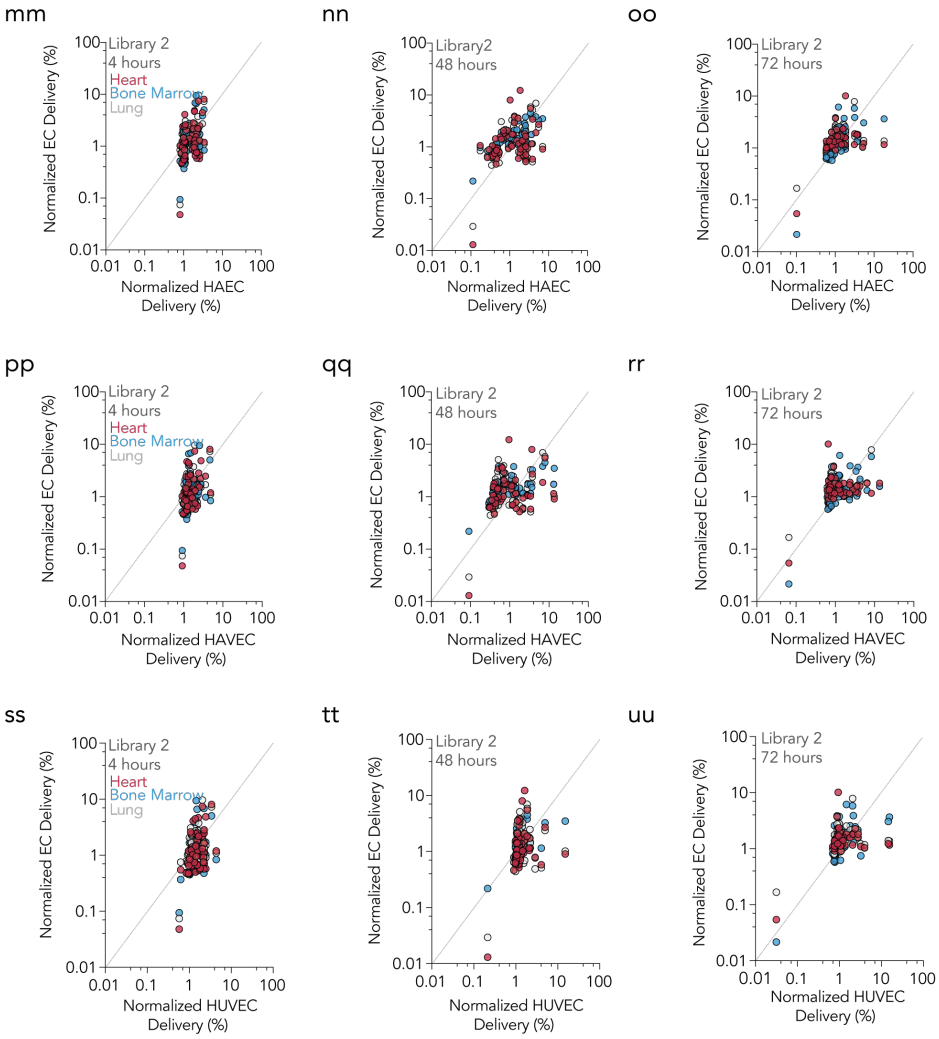


Figure 2.6 Gating strategy for isolation of endothelial cells and macrophages from (a) mouse heart, (b) mouse lung, and (c) mouse bone marrow. (d) Library 2 was designed to test the efficiency of different lipomers. 15 different lipomers were tested at varying lipomer mole percentages. Chemical formulas for lipomers are shown. (e) LNP diameters (nm) for stable pooled LNPs from library 2. (f) Library 3 was designed to further investigate the effect of PEG lipid length (C14, C16, C18) and MW (350, 2000). Chemical formulas for lipomers are shown. (g) LNP diameters (nm) for stable pooled LNPs from library 3. (h) For experiments in (a), (b), and (c), 420 LNPs were formulated, of which 281 formed stable LNPs and were pooled. (i) Diameter distribution for 281 stable LNPs that were pooled. Lines on violin plot represent 75th quartile, mean, and 25th quartile, respectively. Normalized heart, lung, and bone marrow endothelial cell delivery versus normalized iMAEC delivery for (j-l) library 1, (m-o) library 2, (p-r) library 3. Normalized heart, lung, or bone marrow macrophage delivery versus normalized RAW delivery for (s-u) library 1, (v-x) library 2, (yaa) library 3. Normalized heart, lung, and bone marrow endothelial cell delivery versus normalized (bb-dd, hh-jj) iMAEC and (ee – gg, kk-ll) RAW delivery 4 hours and 48 hours after LNP administration, respectively. (mm-oo) Normalized HAEC, (pp-rr) HAVEC, and (ss-uu) HUVEC delivery 4 hrs, 48 hrs, and 72 hrs after LNP administration versus heart, lung, and bone marrow macrophages at the same timepoints.

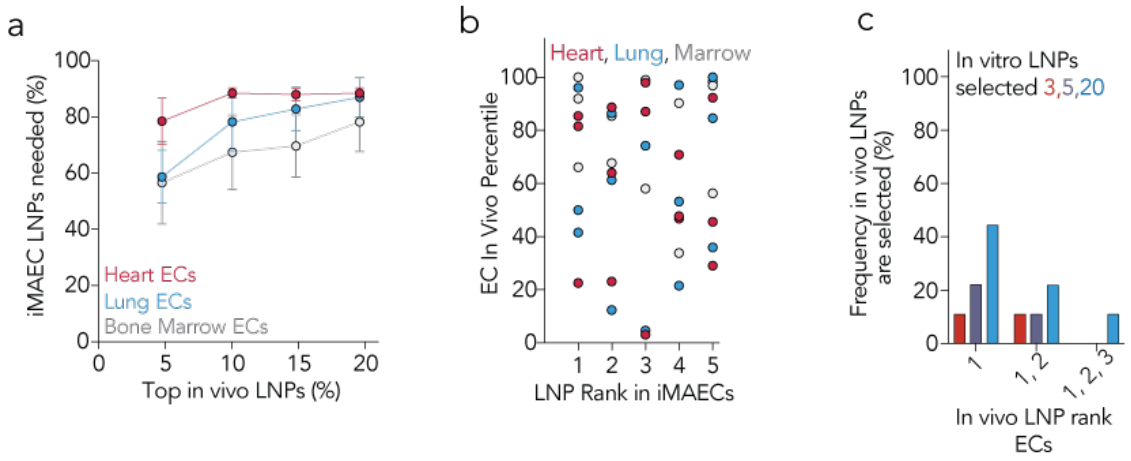
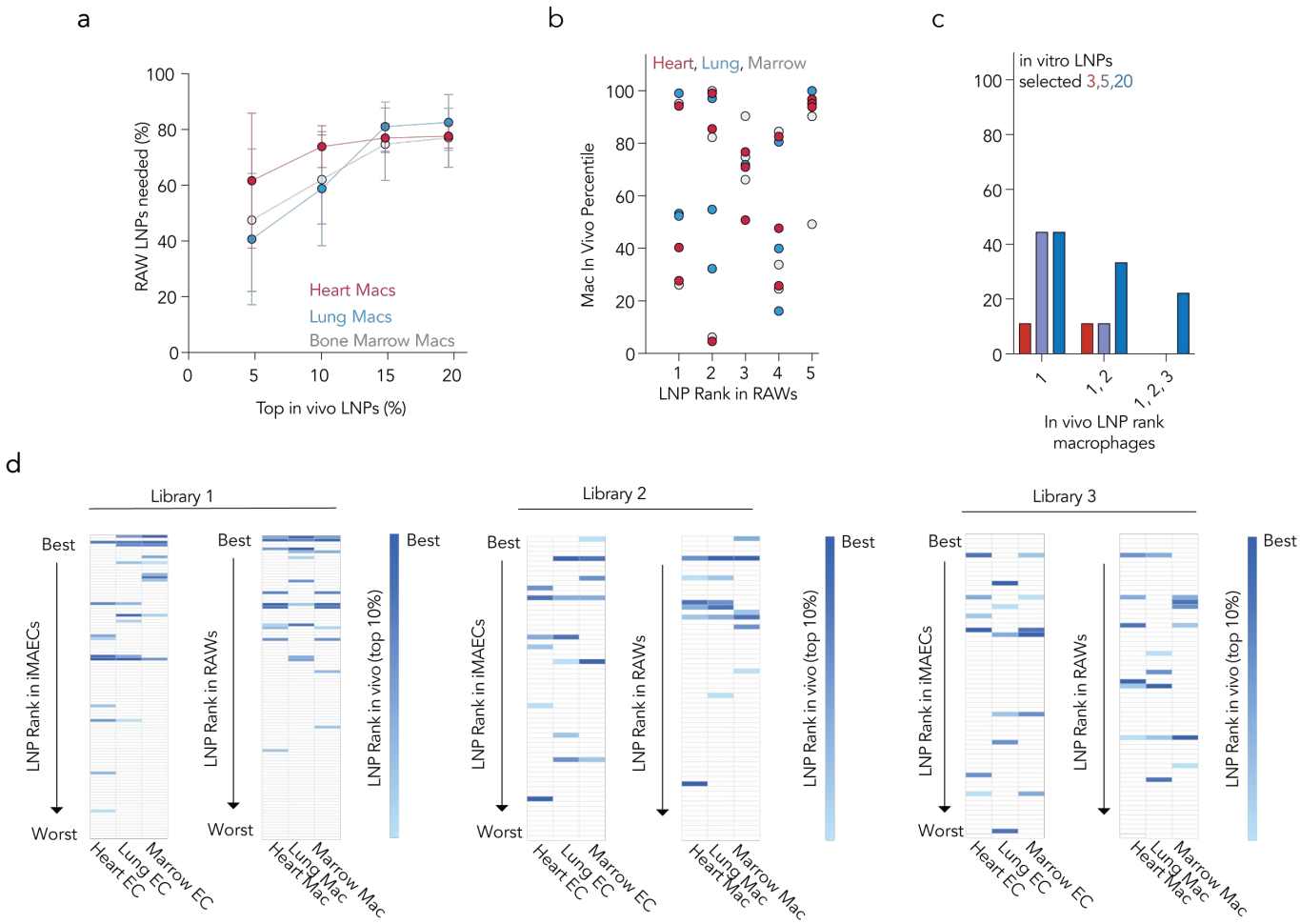


Figure 2.7 Quantifying the efficiency with which *in vitro* screens predict *in vivo* delivery. (a) Percentage of *in vitro* LNPs required to encompass the top 5, 10, 15, and 20% of the LNPs in heart, lung, and bone marrow endothelial cells *in vivo*. For example, over 50% of the *in vitro* library would be required to ensure the top 5% of the *in vivo* LNPs were selected. (b) LNP rank *in vivo* in heart, lung, and bone marrow endothelial cells, for the top 5 *in vitro* ranked LNPs. (c) Frequency with which the 1st, 1st and 2nd, or 1st, 2nd, and 3rd *in vivo* LNPs would be chosen by selecting the top 3, 5, and 20 LNPs *in vitro*.


a
RAW LNPs needed (%)
Top in vivo LNPs (%)
Heart Macs
Lung Macs
Bone Marrow Macs
b
Heart, Lung, Marrow
Mac In Vivo Percentile
LNP Rank in RAWs
c
in vitro LNPs
selected 3,5,20
In vivo LNP rank
macrophages
d
Library 1
Library 2
Library 3
Best
Worst
LNP Rank in iMAECs
LNP Rank in RAWs
LNP Rank in vivo (top 10%)
Heart EC
Lung EC
Marrow EC
Heart Mac
Lung Mac
Marrow Mac

e

Library 1													
Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Diameter (nm)	iMAEC	Heart Ecs	Lung Ecs	Marrow Ecs
7C1	PEI 600	C15 Epoxy	10	C14	350	62	20	18	143.4	1		5	1
7C1	PEI 600	C15 Epoxy	10	C16	350	62	15	23	117.6	3	3	3	2
7C1	PEI 600	C15 Epoxy	10	C14	350	62	5	33	218.8	4		4	4
7C1	PEI 600	C15 Epoxy	10	C16	350	62	20	18	198.7	8			6
7C1	PEI 600	C15 Epoxy	10	C18	5000	62	35	3	83.9	10		8	10
7C1	PEI 600	C15 Epoxy	10	C14	350	62	0	38	120.2	14			7
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	5	33	33.4	15			3
7C1	PEI 600	C15 Epoxy	10	C18	350	62	20	18	136.2	16			8
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	30	8	154.3	24	4	7	
7C1	PEI 600	C15 Epoxy	10	C14	5000	62	20	18	73.8	28		2	9
7C1	PEI 600	C15 Epoxy	10	C14	5000	62	30	8	77.1	30		9	
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	10	28	50.3	35	6		
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	20	18	89.8	36	9		
104	PEI 600	C12 Epoxy	10	C14	2000	62	5	33	144.2	42	2	6	
104	PEI 600	C12 Epoxy	10	C18	350	62	5	33	56.3	43	1	1	5
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	25	13	111.6	59	7		
104	PEI 600	C12 Epoxy	10	C18	2000	62	10	28	47.8	64	5	10	
104	PEI 600	C12 Epoxy	10	C18	350	62	10	28	68.2	82	8		
104	PEI 600	C12 Epoxy	10	C16	350	62	30	8	78.2	95	10		

Library 1													
Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Diameter (nm)	RAWs	Heart Macs	Lung Macs	Marrow Macs
7C1	PEI 600	C15 Epoxy	10	C14	350	62	20	18	143.4	1	7	2	6
7C1	PEI 600	C15 Epoxy	10	C16	350	62	15	23	117.6	2	2	4	1
7C1	PEI 600	C15 Epoxy	10	C14	350	62	5	33	218.8	5	6	1	
7C1	PEI 600	C15 Epoxy	10	C16	350	62	20	18	198.7	6		8	7
7C1	PEI 600	C15 Epoxy	10	C16	350	62	0	38	134.0	8		10	
7C1	PEI 600	C15 Epoxy	10	C18	5000	62	35	3	83.9	16		5	
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	30	8	154.3	20	4		4
104	PEI 600	C12 Epoxy	10	C18	350	62	5	33	56.3	24	1	9	2
104	PEI 600	C12 Epoxy	10	C14	2000	62	5	33	144.2	25	3		3
7C1	PEI 600	C15 Epoxy	10	C14	5000	62	20	18	73.8	31	10	3	
7C1	PEI 600	C15 Epoxy	10	C18	2000	62	10	28	50.3	32	8		10
104	PEI 600	C12 Epoxy	10	C18	2000	62	10	28	47.8	36	5		5
7C1	PEI 600	C15 Epoxy	10	C14	5000	62	25	13	74.6	42		7	
7C1	PEI 600	C15 Epoxy	10	C14	5000	62	30	8	77.1	43		6	
104	PEI 600	C12 Epoxy	10	C14	350	62	30	8	88.5	47			8
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	5	33	67.6	66			9
104	PEI 600	C12 Epoxy	10	C18	350	62	30	8	172.8	74	9		

f

Library 2															
Lipomer	Amine	Lipid	Phospholipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Helper Lipid Mole %	Diameter (nm)	iMAEC	Heart Ecs	Lung Ecs	Marrow Ecs
68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0	21.5	1			6
7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	80	0	20	0	33.3	5		1	2
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10	69.7	9			3
100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	80	15	5	0	37.0	11	4		
78	PEI 1200	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0	35.0	13	2	4	4
100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0	30.3	20	3	2	
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10	34.3	22	5		
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	36	60	5	0	31.2	25		6	1
88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10	64.0	34	6		
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	62	30	8	0	66.1	40		5	
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0	53.2	45		3	5
102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0	26.7	53	1		

Library 2															
Lipomer	Amine	Lipid	Phospholipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Helper Lipid Mole %	Diameter (nm)	RAWs	Heart Macs	Lung Macs	Marrow Macs
68	PEI 600	Lauryl Acrylate	DOPE	10	C14	2000	35	35	30	0	21.5	1			4
8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0	21.7	17	5	4	2
8	1,4,8,12-Tetraazacyclopentadecane	Lauryl Acrylate	DOPE	10	C14	2000	50	38.5	1.5	10	53.0	19			3
48	Spermine Free Base	Lauryl Acrylate	DOPE	10	C14	2000	80	0	20	0	22.9	16			5
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10	34.3	14	2	3	
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	80	15	5	0	53.2	15	4	2	
58	Spermidine	Lauryl Acrylate	DOPE	10	C14	2000	35	60	5	0	31.2	28			6
88	Triethylenetetramine	Lauryl Acrylate	DOPE	10	C14	2000	50	28.5	11.5	10	64.0	33		6	
100	Spermine Free Base	C12 Epoxy	DOPE	10	C14	2000	62	30	8	0	30.3	9	6	5	
102	Spermidine	C12 Epoxy	DOPE	10	C14	2000	35	35	30	0	26.7	51	1		
7C1	PEI 600	C15 Epoxy	DOPE	10	C14	2000	80	0	20	0	33.3	5	3	1	1

g

Library 3													
Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Diameter (nm)	iMAEC	Heart Ecs	Lung Ecs	Marrow Ecs
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	0	38	72.0	5	5		
7C1	PEI 600	C15 Epoxy	10	C16	350	62	24	14	201.2	7			1
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	18	20	105.6	11		6	
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	9	29	71.6	14	3		2
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	6	32	58.3	16		1	6
104	PEI 600	C12 Epoxy	10	C16	350	62	27	11	108.7	18	2		
7C1	PEI 600	C15 Epoxy	10	C16	350	62	12	26	234.6	19			5
104	PEI 600	C12 Epoxy	10	C18	2000	62	30	8	56.8	21	6		
7C1	PEI 600	C15 Epoxy	10	C16	2000	62	30	8	53.4	22		3	
104	PEI 600	C12 Epoxy	10	C16	350	62	21	17	75.0	29			4
7C1	PEI 600	C15 Epoxy	10	C16	2000	62	15	23	42.6	33			3
104	PEI 600	C12 Epoxy	10	C18	2000	62	9	29	182.2	38		2	
7C1	PEI 600	C15 Epoxy	10	C16	2000	62	6	32	36.8	45		4	
104	PEI 600	C12 Epoxy	10	C14	2000	62	18	20	237.9	52	4		
104	PEI 600	C12 Epoxy	10	C18	2000	62	0	38	150.6	56	1		
104	PEI 600	C12 Epoxy	10	C18	2000	62	18	20	111.5	64		5	
Library 3													
Lipomer	Amine	Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Diameter (nm)	RAWs	Heart Macs	Lung Macs	Marrow Macs
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	21	17	135.8	5	4	3	
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	0	38	72.0	14	3		3
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	6	32	58.3	15			5
7C1	PEI 600	C15 Epoxy	10	C16	2000	62	30	8	53.4	16			4
7C1	PEI 600	C15 Epoxy	10	C14	2000	62	24	14	60.4	20	5		2
104	PEI 600	C12 Epoxy	10	C18	2000	62	21	17	186.6	26		1	
104	PEI 600	C12 Epoxy	10	C16	2000	62	12	26	176.3	30		4	
104	PEI 600	C12 Epoxy	10	C16	350	62	27	11	108.7	32	6		
104	PEI 600	C12 Epoxy	10	C18	2000	62	0	38	150.6	33	2	6	
104	PEI 600	C12 Epoxy	10	C14	2000	62	18	20	237.9	45	1	2	6
104	PEI 600	C12 Epoxy	10	C16	350	62	0	38	193.8	51			1
104	PEI 600	C12 Epoxy	10	C14	2000	62	12	26	51.4	53		5	

h

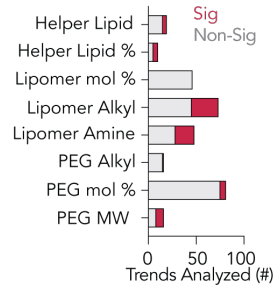
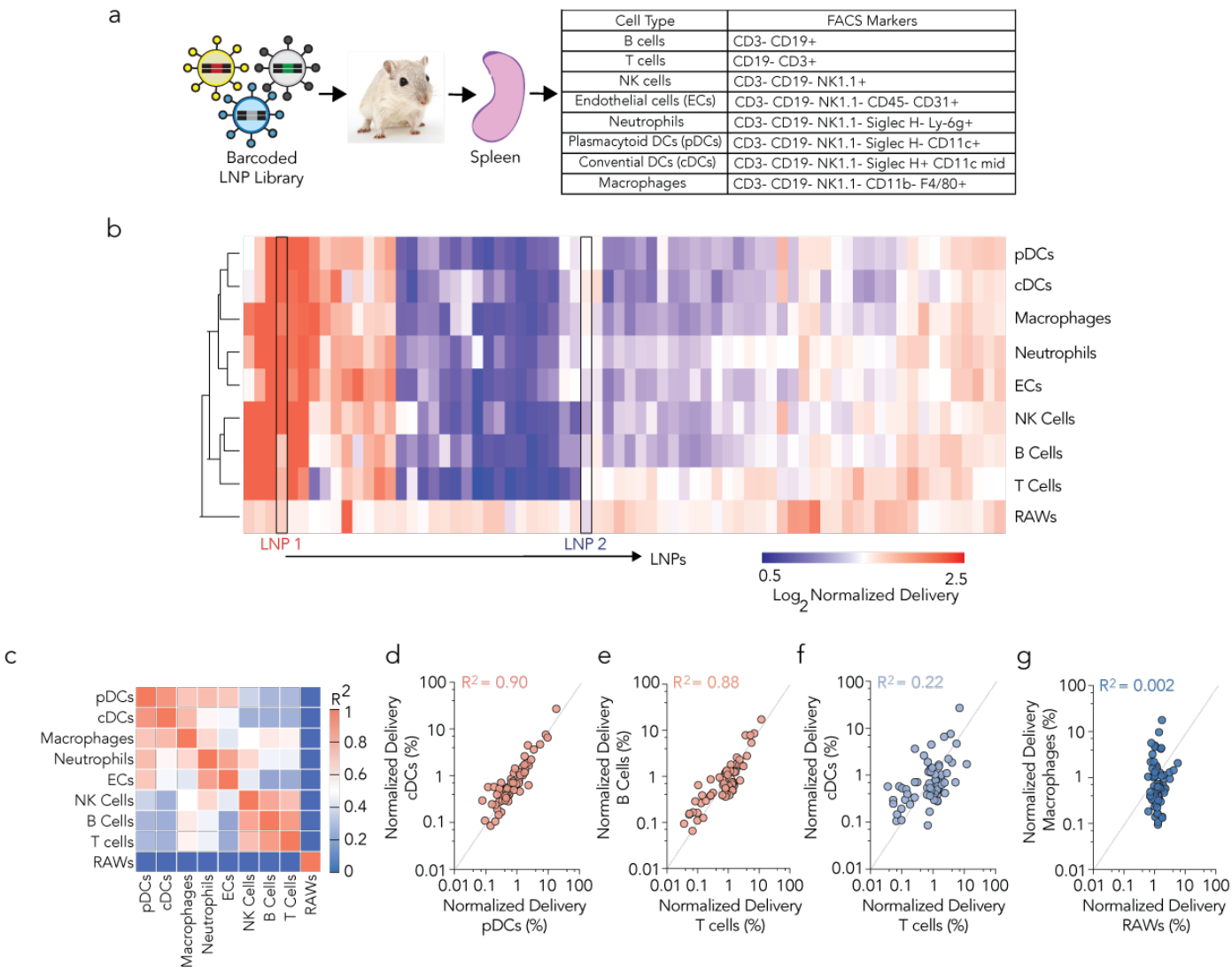


Figure 2.8 The amount (%) of total LNPs from an in vitro screen required to encompass the top 5, 10, 15, and 20% of LNPs in heart, lung, and bone marrow macrophages in vivo. For example, nearly 60% of the in vitro library would be required to ensure that the top 5% of in vivo LNPs were found. (b) LNP rank in vivo in heart, lung, and bone marrow macrophages, for the top 5 in vitro performing LNPs. (c) Frequency with which the 1, 1 and 2, or 1, 2, and 3 in vivo LNPs are chosen when the top 3, 5, and 20 LNPs in vitro. (d) LNP rank in heart, lung, and bone marrow macrophages, plotted by in vitro ranking. The columns are sorted from best (LNP 1) to worst (LNP 104) in vitro. Top performing in vivo LNPs are colored. (e-g) The compositions and diameters of the colored LNPs in Fig. S4D are listed for each library for heart, lung, and bone marrow ECs and macrophages. (h) The number of significant ($p<0.05$, ANOVA) and non-significant relationships between LNP structure and cell targeting. Lipomer alkyl tail, amine structure, and PEG MW were most likely to influence LNP delivery.

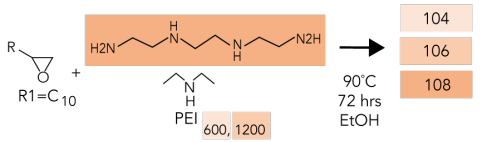


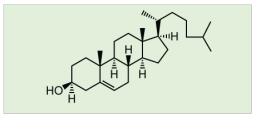
Cell Type	FACS Markers
B cells	CD3- CD19+
T cells	CD19- CD3+
NK cells	CD3- CD19- NK1.1+
Endothelial cells (ECs)	CD3- CD19- NK1.1- CD45- CD31+
Neutrophils	CD3- CD19- NK1.1- Siglec H- Ly-6g+
Plasmacytoid DCs (pDCs)	CD3- CD19- NK1.1- Siglec H- CD11c+
Convential DCs (cDCs)	CD3- CD19- NK1.1- Siglec H+ CD11c mid
Macrophages	CD3- CD19- NK1.1- CD11b- F4/80+

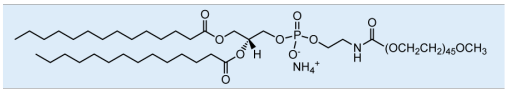
Figure 2.9 High throughput analysis of delivery to splenic microenvironment. (a) 144 LNPs were formulated; 85 stable LNPs were pooled and administered to WT mice. 72 hours later, cell types were isolated from the spleen using FACS. (b) Unbiased clustering of LNPs in each cell type, generated by a Euclidean distance algorithm. RAWs (macrophage *in vitro*) clustered separately from all 8 *in vivo* cell types, and both DC populations clustered together. (c) R^2 values for all 8 *in vivo* cell types as well as RAWs. Normalized delivery in (d) plasmacytoid and conventional DCs, (e) B cells and T cells, (f) conventional DCs and T cells, and (g) RAWs and splenic macrophages.

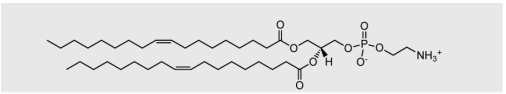
a

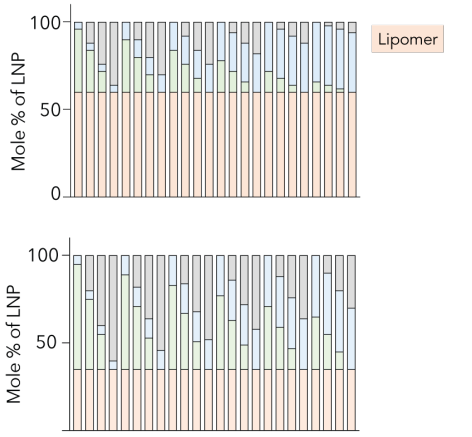
LNP #	Lipomer	Amine	Lipid	Helper Lipid	Mass Ratio	PEG Chain	PEG MW	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Helper Lipid Mole %
1	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	36	4	0
2	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	24	4	12
3	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	12	4	24
4	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	4	36
5	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	30	10	0
6	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	20	10	10
7	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	10	10	20
8	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	10	30
9	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	24	16	0
10	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	16	16	8
11	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	8	16	16
12	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	16	24
13	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	18	22	0
14	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	12	22	6
15	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	6	22	12
16	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	22	18
17	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	12	28	0
18	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	8	28	4
19	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	4	28	8
20	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	28	12
21	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	6	34	0
22	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	4	34	2
23	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	2	34	4
24	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	60	0	34	6
25	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	36	4	0
26	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	24	4	12
27	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	12	4	24
28	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	4	36
29	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	30	10	0
30	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	20	10	10
31	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	10	10	20
32	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	10	30
33	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	24	16	0
34	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	16	16	8
35	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	8	16	16
36	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	16	24
37	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	18	22	0
38	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	12	22	6
39	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	6	22	12
40	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	22	18
41	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	12	28	0
42	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	8	28	4
43	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	4	28	8
44	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	28	12
45	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	6	34	0
46	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	4	34	2
47	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	2	34	4
48	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	60	0	34	6
49	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	36	4	0
50	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	24	4	12
51	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	12	4	24
52	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	4	36
53	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	30	10	0
54	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	20	10	10
55	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	10	10	20
56	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	10	30
57	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	24	16	0
58	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	16	16	8
59	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	8	16	16
60	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	16	24
61	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	18	22	0
62	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	12	22	6
63	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	6	22	12
64	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	22	18
65	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	12	28	0
66	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	8	28	4
67	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	4	28	8
68	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	28	12
69	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	6	34	0
70	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	4	34	2
71	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	2	34	4
72	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	60	0	34	6
73	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	60	5	0
74	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	40	5	20
75	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	20	5	40
76	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	5	60
77	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	54	11	0
78	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	36	11	18
79	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	18	11	36
80	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	11	54
81	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	48	17	0
82	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	32	17	16
83	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	16	17	32
84	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	17	48
85	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	42	23	0
86	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	28	23	14
87	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	14	23	28
88	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	23	42
89	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	36	29	0
90	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	24	29	12
91	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	12	29	24
92	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	29	36
93	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	30	35	0
94	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	20	35	10
95	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	10	35	20
96	104	PEI 600	C12-Epoxy	DOPE	10	C14	2000	35	0	35	30
97	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	60	5	0
98	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	40	5	20
99	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	20	5	40
100	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	5	60
101	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	54	11	0
102	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	36	11	18
103	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	18	11	36
104	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	11	54
105	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	48	17	0
106	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	32	17	16
107	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	16	17	32
108	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	17	48
109	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	42	23	0
110	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	28	23	14
111	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	14	23	28
112	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	23	42
113	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	36	29	0
114	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	24	29	12
115	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	12	29	24
116	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	29	36
117	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	30	35	0
118	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	20	35	10
119	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	10	35	20
120	106	PEI 1200	C12-Epoxy	DOPE	10	C14	2000	35	0	35	30
121	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	60	5	0
122	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	40	5	20
123	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	20	5	40
124	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	5	60
125	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	54	11	0
126	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	36	11	18
127	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	18	11	36
128	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	11	54
129	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	48	17	0
130	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	32	17	16
131	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	16	17	32
132	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	17	48
133	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	42	23	0
134	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	28	23	14
135	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	14	23	28
136	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	23	42
137	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	36	29	0
138	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	24	29	12
139	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	12	29	24
140	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	29	36
141	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	30	35	0
142	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	20	35	10
143	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	10	35	20
144	108	Triethylenetetramine	C12-Epoxy	DOPE	10	C14	2000	35	0	35	30



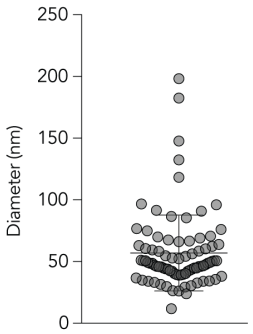








b



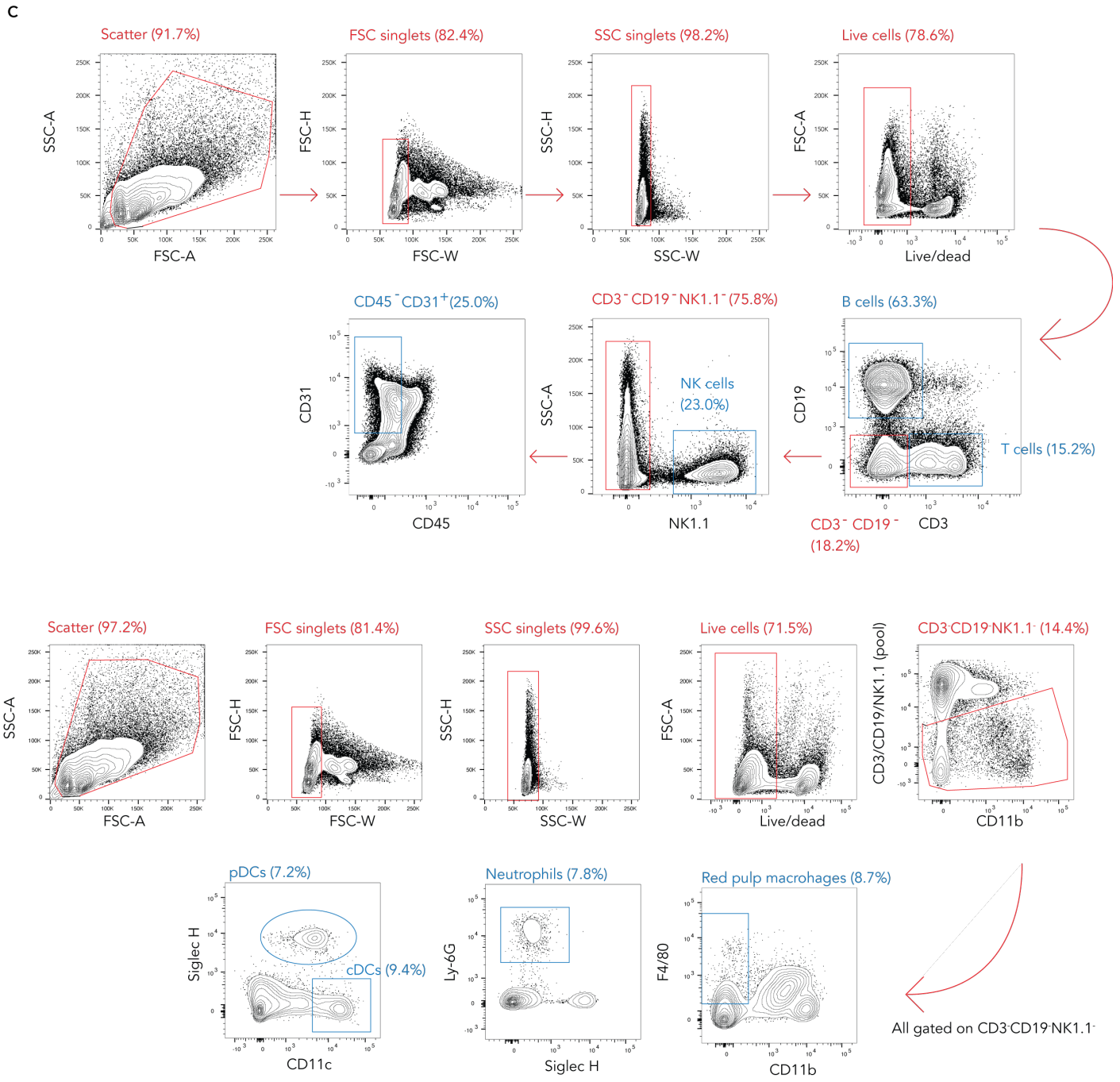


Figure 2.10 (a) LNPs used for spleen biodistribution are shown. Three different lipomers with varying amines. Two low MW Poly(ethlyeneimine) (PEI - MW600 and MW1200), and triethylenetetramine were chosen. Chemical formulas for lipomers as well as other LNP components are shown. (b) Stable LNPs were pooled. (c) Gating strategy for isolation of immune cell subsets from mouse spleen by FACS. Red gates are parents of subsequent populations, and blue gates represent sorted populations.

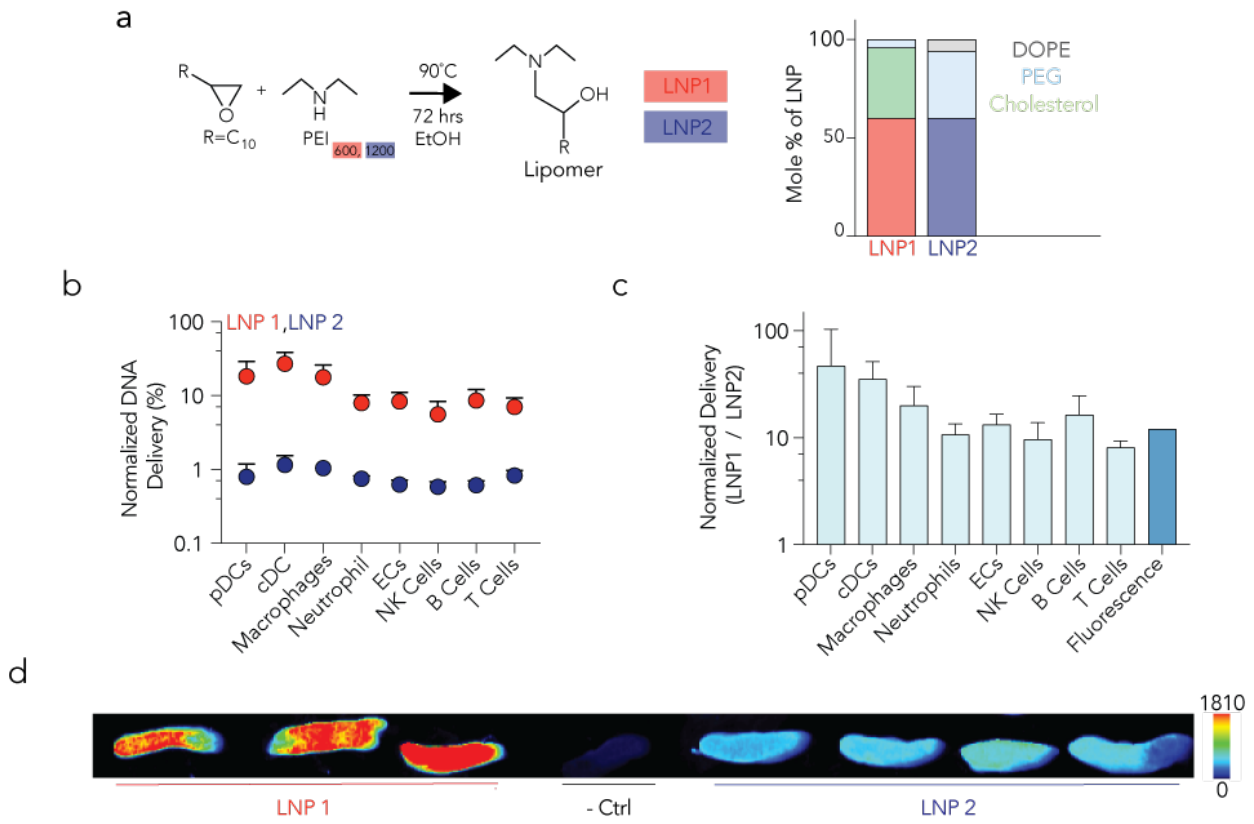


Figure 2.11 (a) Chemical compositions of LNP1 and LNP2. (b) Normalized delivery of 2 LNPs; 1 with high (LNP1) and 1 with low (LNP2) normalized delivery. (c) The ratio of normalized delivery (LNP1/LNP2) in each cell type using barcodes, as well as fluorescence of the spleen whole tissue. (d) Cy5.5 fluorescence in splenic whole tissue 3 hours after mice were injected with either LNP1 or LNP2. LNP1 fluorescence was higher, as predicted by the barcoding data.

a

Library 1			
Strain	Procedure	W (0 hrs) (g)	W (72 hrs) (g)
C57BL/6	IV-pooled NPs	19.2	20.4
C57BL/6	IV-pooled NPs	19.6	19.7
C57BL/6	IV-pooled NPs	17.3	18
C57BL/6	IV-pooled NPs	20.1	21

Library 2			
Strain	Procedure	W (0 hrs) (g)	W (4 hrs) (g)
C57BL/6	IV-pooled LNP Inject	21.4	22.2
C57BL/6	IV-pooled LNP Inject	21.3	21.1
C57BL/6	IV-pooled LNP Inject	23.9	23.3
C57BL/6	IV-pooled LNP Inject	19.1	18.9
Strain	Procedure	W (0 hrs) (g)	W (48 hrs) (g)
C57BL/6	IV-pooled LNP Inject	20.7	22.07
C57BL/6	IV-pooled LNP Inject	22.4	22.6
C57BL/6	IV-pooled LNP Inject	20.7	21.4
C57BL/6	IV-pooled LNP Inject	22.7	22.6
Strain	Procedure	W (0 hrs) (g)	W (72 hrs) (g)
C57BL/6	IV-pooled LNP Inject	20.3	21.7
C57BL/6	IV-pooled LNP Inject	21.7	21.8
C57BL/6	IV-pooled LNP Inject	21.2	21.7
C57BL/6	IV-pooled LNP Inject	21	21.5

Library 3			
Strain	Procedure	W (0 hrs) (g)	W (72 hrs) (g)
C57BL/6	IV-pooled LNP Inject	18.1	19.5
C57BL/6	IV-pooled LNP Inject	19	20
C57BL/6	IV-pooled LNP Inject	19.3	19.3
C57BL/6	IV-pooled LNP Inject	18.5	19.9

Splenic Library			
Strain	Procedure	W (0 hrs) (g)	W (72 hrs)
C57BL/6	IV-pooled LNP Inject	23.8	24.4
C57BL/6	IV-pooled LNP Inject	22.2	22.9
C57BL/6	IV-pooled LNP Inject	24.2	25
C57BL/6	IV-pooled LNP Inject	22.5	23.5

b

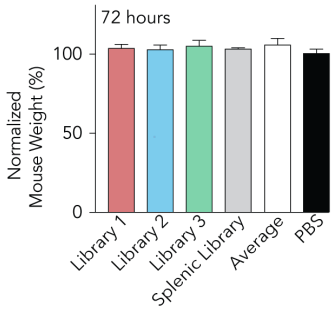


Figure 2.12 (a) Mouse weights before, W (0 hrs), and after injection, W (72 hrs), are shown for each experiment. (b) Normalized mouse weights for each experiment compared to PBS injected mice.

CHAPTER 3. ANALYZING IN VIVO DRUG DELIVERY DATA POINTS REVEALS CHOLESTEROL STRUCTURE IMPACTS NANOPARTICLE DELIVERY

The work presented here is an excerpt from Paunovska K, Carmen JG, Lokugamage MP, Sago CD, Sato M, Lando GN, Gamboa Castro M, Bryksin AV, Dahlman JE (2018). "Analyzing 2000 *in vivo* Drug Delivery Data Points Reveals Cholesterol Structure Impacts Nanoparticle Delivery." *ACS Nano.*

3.1 Introduction

In vivo drug delivery is a complex process that is difficult to predict.[98, 114] The relationship between *in vitro* and *in vivo* delivery can be non-existent,[8] demonstrating the utility of testing hundreds of nanoparticles *in vivo*.[8] Recently, DNA barcode-based technologies have enabled scientists to study many nanoparticles *in vivo* simultaneously.[7, 8, 162] Here we sought to improve nucleic acid delivery mediated by LNPs by systematically studying the relationship between LNP chemical structure and *in vivo* delivery.

We focused on cholesterol variants in the LNP for several reasons. Many labs have studied how the structure of hundreds of cationic or ionizable lipid-like biomaterial in LNPs affects delivery *in vitro*;[64, 65, 68, 71, 76, 163] *in vivo* structure function studies using more than a few LNPs have not been published. LNPs are created ('formulated') by mixing these lipid-like biomaterials with other constituents, most often PEG-lipids[69] and unmodified cholesterol. However, cholesterol is naturally oxidized or esterified *in vivo*. Oxidized cholesterol is typically found within ox-LDL. LDL oxidation is partially driven by diet, the presence of ROS, and other factors.[164] Esterification of cholesterol occurs at different sites (*e.g.* peripheral tissues, liver), enabling more compact storage and transportation of

cholesterol.[165] Esterification of cholesterol from peripheral tissues is mediated by lecithin-cholesterol acyltransferase on the surface of nascent HDL, and by acyl-COA-cholesterol acyltransferase intracellularly.[166, 167] These cholesterol 'variants' are also actively trafficked to cells including hepatocytes, endothelial cells, and macrophages.[165]

The amount of different cholesterol variants also changes with many common diseases (*e.g.,* high cholesterol, atherosclerosis, hyperlipidemia, diabetes),[168] suggesting that LNP trafficking may change with the disease state of the patient. For instance, ox-LDL is pro-atherosclerotic, pro-inflammatory, and contributes to the amount of ROS in the bloodstream.[169-172] Increases in ox-LDL are indicators of diseases such as atherosclerosis and heart disease.[173-176] Despite these facts, the relationship between cholesterol structure and *in vivo* LNP delivery remains unexplored.

We hypothesized that the structure of cholesterol included in LNPs affects targeting *in vivo.* This hypothesis has important implications. It suggests LNP targeting can be tuned using naturally- or synthetically-derived cholesterol variants; this is critical given the need for LNPs that deliver RNAs to cell types other than hepatocytes.[3] It also implies that LNPs may behave differently in patients with aberrant cholesterol levels. For instance, patients with aberrant metabolisms may have increased amounts of oxidative stress, leading to a higher presence of oxidized cholesterol, creating positive feedback.[177] In patients with dyslipidemia, this positive feedback loop can start as a change in lipoprotein core structure, particularly a decrease in cholesteryl esters and cholesterol and an increased chance of oxidation.[178] This is important given the growing clinical use of LNPs that deliver siRNAs[179] and the high percentage of patients that have aberrant cholesterol levels. We tested our hypothesis in wild type mice as well as LDLR and VLDLR knockout mice. Both

mouse models are regularly used to study cholesterol dysfunction.[180, 181] LDLR$^{-/-}$ and VLDLR$^{-/-}$ mice are typically given high fat diets to induce metabolic disease. Herein we examined nanoparticle delivery in strain- and age-matched WT controls for mice fed a normal diet, such that the only difference would be LDLR or VLDLR expression.

Other groups have studied the relationship between nanoparticles and gene expression[84, 127]. Our current work complements these studies but is distinct. These studies used a small number of nanoparticles to test the hypothesis that a specific gene influenced nanoparticle delivery. By contrast, we used >100 LNPs to test hypothesis that cholesterol structure affected LNP delivery. We formulated 141 LNPs with 6 cholesterol variants based on natural lipoproteins. We administered all the LNPs *in vivo* at once to WT, LDLR$^{-/-}$, or VLDLR$^{-/-}$ mice using high throughput LNP DNA barcoding,[7, 8] and isolated cells using FACS. Using a nanoparticle bioinformatics pipeline, we found that LNPs formulated with esterified cholesterol increased LNP distribution relative to LNPs with regular or oxidized cholesterol in WT mice when averaging nanoparticle distribution across all cell types analyzed. Based on the *in vivo* screen, we identified an LNP enriched in hepatic endothelial cells. As predicted by the *in vivo* nanoparticle barcoding screen, the LNP efficiently delivered therapeutic acids to hepatic endothelial cells, which have been refractory to systemic nanoparticle targeting.

3.2 Results

We used high throughput DNA barcoding[7, 8] to assess how cholesterol variants altered LNP biodistribution. LNPs can be made with a similar size and composition as LDL and VLDL (**Fig. 3.1A**), two lipoproteins which interact with the LDLR and VLDLR (**Fig.**

3.1B). We formulated 141 LNPs using esterified, oxidized, or unmodified cholesterol (**Fig. 3.1C, D, Fig. 3.2A**). All LNPs were formulated using the validated biomaterial 7C1.[76] Of the 141, 111 met our inclusion criteria: autocorrelation curves with 1 inflection point and hydrodynamic diameters between 20 nm and 200 nm, based on DLS (**Fig. 3.1E-F, Fig. 3.2B**). These 111 LNPs – along with a naked DNA barcode, which served as a negative control - were pooled together and intravenously administered to WT, $LDLR^{-/-}$, and $VLDLR^{-/-}$ mice at a total DNA dose of 0.5 mg/kg (0.0045 mg/kg/barcode). We sacrificed the mice 72 hours later and harvested

DNA from lung endothelial cells ($CD31^{+}CD45^{-}$), lung macrophages ($CD31^{-}CD45^{+}CD11b^{+}$), splenic endothelial cells, splenic macrophages, liver endothelial cells, and hepatocytes ($CD31^{-}CD45^{-}$) using FACS[8, 76, 145, 148, 149] (**Fig. 3.2C-E**). Seventy-two hours is sufficiently long for LNPs to be cleared from the bloodstream.[76] All 3 cell types play critical roles in cholesterol trafficking. Macrophages are known to uptake oxidized cholesterol *via* scavenger receptor-mediated endocytosis of ox-LDL, one of the initial steps in the formation of foam cells which are critical to the progression of atherosclerosis[182-184] Hepatocytes also play a critical role by synthesizing cholesterol in the liver and responding to internal increases or decreases in cholesterol by up or down-regulating production of LDLR.[185, 186] Finally, endothelial cells actively interact with serum lipoproteins to maintain cholesterol homeostasis.[187, 188] To assess how all LNPs delivered DNA at once, we amplified barcodes and deep sequenced them as we previously described[7, 8]. The readout for these DNA sequencing experiments is normalized delivery,[8] which is analogous to counts per million in RNA-seq experiments (**Fig. 3.2F**).

We first analyzed whether cholesterol structure affected LNP size. We measured the hydrodynamic diameter of all 111 LNPs individually. Oxidized, esterified, and unmodified cholesterol formulated LNPs that met our inclusion criteria 72-100%, 56-76%, and 80% of the time, respectively (**Fig. 3.2G**). Cholesteryl oleate had the lowest percent included LNPs and the tightest diameter distribution (22 – 115 nm) for cholesterol containing LNPs (**Fig. 3.2H**). The average LNP diameter did not change with cholesterol type (**Fig. 3.1H**). As an additional control for LNP size, we plotted normalized delivery against LNP diameter for all ~2,000 *in vivo* data points. The number of *in vivo* data points was calculated as shown in **Table 1**. As we reported previously,[7, 8] we found no relationship between LNP size and delivery (**Fig. 3.2I**). We then looked at whether this relationship improved if we split the LNPs by cholesterol variant and then plotted normalized delivery against LNP diameter (**Fig. 3.2J-P**). We did not observe an improvement in the R^2 value, suggesting that there was no trend between LNP size distribution and normalized delivery when breaking up the LNPs by cholesterol variant (**Fig. 3.2Q**). The diameter of the pooled LNPs was similar to the diameters of the individual LNPs (**Fig. 3.1F**). We also analyzed the delivery of naked barcode; as expected, this negative control was delivered much less efficiently than barcodes delivered by LNPs in all 18 samples (**Fig. 3.1G, H**).

3.2.1 Cholesterol variants impact LNP biodistribution

To assess how cholesterol structure affected delivery *in vivo* for all LNPs, we used unbiased Euclidian analysis to generate a nanoparticle targeting heatmap (**Fig. 3.3A**). Euclidean analysis is a common bioinformatics approach[189] that 'clusters' large data sets into experimental groups which behave similarly; it can be used to study barcoded LNPs.[8] The naked barcode – designated by an asterisk – was easily identified; it delivered barcodes

inefficiently in all samples. Euclidean analysis created 3 clusters; when compared to the 'center' cluster, the left- and right-most clusters had more purple, which designated higher normalized delivery (**Fig. 3.3A**). Based on this data visualization, we analyzed the cholesterol types in the left- and right-most (*i.e.* 'good') clusters, and the center (*i.e.* 'bad') cluster. LNPs formulated with esterified cholesterols were 1.4-fold enriched in the good clusters, relative to LNPs made with oxidized cholesterol. In other words, LNPs formulated with esterified cholesterols were 1.4-fold more likely to be in the left- or right-most clusters than oxidized cholesterols. LNPs made with oxidized cholesterols were enriched by 1.3-fold in the center cluster (**Fig. 3.3B**). Enrichment is described in **Fig. 3.4A**.

3.2.2 Esterified cholesterol outperforms oxidized cholesterol

Based on these analyses, we quantified normalized barcode delivery mediated by nanoparticles that contained esterified, unmodified, or oxidized cholesterols in all cell types in WT mice. Normalized barcode delivery mediated by LNPs made with esterified cholesterol was significantly higher than barcode delivery mediated by LNPs with regular cholesterol or oxidized cholesterol (**Fig. 3.3C, D, Table 2**). These analyses averaged delivery of each nanoparticle, including those that delivered barcodes inefficiently, across all cell types. However, many studies focus on top performing LNPs. We identified LNPs in the top 15% in each cell type and performed an enrichment analysis as described in **Fig. 3.4A**. LNP formulations that were enriched in each cell type in WT, VLDLR$^{-/-}$, and LDLR$^{-/-}$ mice are listed in **Fig. 3.4B-D**. We then analyzed whether particle size and biodistribution in top performing LNPs were correlated and found no significant relationship between the two (**Fig. 3.4E-K**). We performed this analysis for the whole animal (*i.e.*, all cell types, averaged); esterified cholesterol was consistently enriched in the top 15%. LNPs

formulated with esterified cholesterols were 2-fold more likely to be in the top 15% of LNPs than LNPs made with oxidized cholesterols in WT mice (**Fig. 3.3E, F**). Taken together, these data support the hypothesis that cholesterol structure affects LNP delivery *in vivo*.

3.2.3 LDLR and VLDLR affect LNP in vivo delivery globally

We performed the same analyses described above (**Fig. 3.3C-F**) for $LDLR^{-/-}$ and $VLDLR^{-/-}$ mice (**Fig. 3.4L-S**). Mimicking results in WT mice, oxidized cholesterols performed poorly relative to esterified and unmodified cholesterol in both knockout models when quantified using average normalized delivery and enrichment. Esterified cholesterol and unmodified cholesterol performed similarly when quantified using average normalized delivery. Enrichment in the top 15% varied; esterified cholesterol outperformed unmodified cholesterol in $LDLR^{-/-}$ mice, but not in $VLDLR^{-/-}$ mice. These results support previously published data demonstrating the cholesterol trafficking receptors can affect LNP delivery.[84]

Based on these initial analyses, we quantified the extent to which LNP delivery in $LDLR^{-/-}$ and $VLDLR^{-/-}$ mice differed from LNP delivery in WT mice. We plotted normalized delivery for all LNPs in all 6 cell types in WT, $LDLR^{-/-}$, and $VLDLR^{-/-}$ mice (**Fig. 3.5A**). If either gene affected delivery of the LNP library tested, then the R^2 value between the WT and knockout mice would decrease (**Fig. 3.5A**). The high throughput nature of barcoding enabled us to compare WT and knockout mice rigorously; each plot contains >650 *in vivo* data points (**Fig. 3.5B, C**). We found that both LDLR and VLDLR affected delivery; the R^2 values between WT and either $LDLR^{-/-}$ or $VLDLR^{-/-}$ mice was

0.37 and 0.50, respectively. We then evaluated whether there was a cell type-specific effect to these genes by analyzing the R^2 values between WT and $LDLR^{-/-}$ (**Fig. 3.6A-F**) or $VLDLR^{-/-}$ (**Fig. 3.6G-L**) mice for each of the 6 cell types individually. We did not observe clear patterns; the cell type specific effects of these genes on LNP delivery will need to be explored using different approaches in the future.

3.2.4 LNPs formulated with cholesteryl oleate deliver therapeutic RNA

DNA barcode readouts quantify nanoparticle biodistribution, which is required, but not sufficient, for functional RNA delivery into the cytoplasm. Cytoplasmic RNA delivery is necessary for successful RNA-interference as well as gene editing. RNA mediated *in vivo* genome editing is important for studying biological pathways and understanding the potential differential effects that genes have on different cell types. To analyze whether LNPs formulated with esterified cholesterols functionally delivered RNAs *in vivo,* we selected an LNP for further analysis. To directly compare esterified cholesterol and unmodified cholesterol – which is the current gold standard in the field - we chose an LNP molar ratio (**Fig. 3.7A**) that made up 3 of the top 5 LNPs in hepatic endothelial cells in our barcoding screen (**Fig. 3.7B**). Hepatic endothelial cells have – with few exceptions[151] – been difficult to target systemically, and as a result, have not been edited by Cas9 after systemic administration of sgRNAs. We formulated 2 LNPs with a 50: 29: 11: 10 molar ratio of 7C1, cholesterol, $C_{14}PEG_{2000}$, and 18:1 Lyso PC, respectively. LNP-oleate contained cholesteryl oleate, whereas LNP-unmod contained unmodified cholesterol. This molar ratio resulted in small, stable LNPs when formulated to carry siRNA and sgRNA (**Fig. 3.8A**). We considered formulating LNPs with cholesteryl stearate, however, LNPs formulated with stearate were stable less frequently (**Fig. 3.2H**).

We formulated LNP-oleate and LNP-unmod to carry a chemically modified[190] sgRNA targeting GFP (**Fig. 3.8B**) and injected these nanoparticles intravenously into mice that express SpCas9-P2A-GFP under a CAG promoter. Five days after a 1.0 mg/kg sgRNA injection, we isolated hepatic endothelial cells and hepatocytes using FACS and quantified indels using TIDE.[191] Delivery to hepatic endothelial cells was highly efficient, leading to 41% editing at the target GFP locus (**Fig. 3.7C**). LNP-unmod was efficient (31% indels), but less so than LNP-oleate. Oleate delivery was particularly specific; the indel ratio of hepatic endothelial cells: hepatocytes was 3 (**Fig. 3.7D, E**). By contrast, all previous systemically administered nanoparticle gene editing has occurred preferentially in hepatocytes.[39, 190, 192, 193] This is the first report of sgRNA-mediated *in vivo* editing in hepatic endothelial cells.

We assessed the activity of LNP-oleate and LNP-unmod using siRNA. siRNA-based therapeutics have successfully treated disease in hepatocytes; understanding how to target hepatic endothelial cells has the potential to lead to therapeutics that target endothelial cell driven disease. We intravenously injected WT mice with 1.5 mg/kg siRNA targeting the endothelial specific gene ICAM-2 (**Fig. 3.8C**). Both siICAM-2 and the control siRNA targeting Luciferase (siLuc) were chemically modified to reduce immune stimulation and promote on-target activity.[76, 145, 148] Three days after siICAM-2 treatment with LNP-oleate or LNP-unmod, ICAM-2 protein expression, measured by MFI, decreased by 74% and 75% respectively, in hepatic endothelial cells, compared to PBS- and siLuc-treated mice (**Fig. 3.7F**). Following treatment, mice injected with sgRNA or siRNA gained weight as quickly as PBS-treated mice (**Fig. 3.8D, E**).

3.3 Discussion

Despite being a universal problem that limits all genetic therapies,[15, 30, 194, 195] it is still difficult to predict which nanoparticles will deliver RNAs *in vivo*. Here we demonstrated that *in vivo* screening can be used to identify LNP traits that affect delivery. Our study was powered by strong statistical analyses; we compared nearly 2,000 *in vivo* drug delivery data points. These data support the hypothesis that modified cholesterols can affect nanoparticle targeting.

We identified an LNP formulation that efficiently targeted hepatic endothelial cells *in vivo*. The LNP preferentially delivered sgRNAs to hepatic endothelial cells 3X more efficiently than hepatocytes. This is uncommon; almost all reported LNPs preferentially target hepatocytes.[39, 64, 65, 68, 69, 71, 163, 190, 192, 193] Targeting hepatic endothelial cells is important given the active role they play in establishing the liver microenvironment and driving fibrosis, inflammation, primary tumor growth, and metastasis.[196] Although we do not know the mechanism for preferential targeting to hepatic endothelial cells over hepatocytes, literature suggests that LNPs interact with serum proteins, which may promote delivery to specific cell types. We anticipate future studies utilizing LNP-oleate to treat hepatic endothelial cell disease and study fundamental biological questions related to hepatic endothelial cell signaling. More generally, our data demonstrate that cholesterol can be viewed as another modular LNP component that can be rationally designed to improve *in vivo* delivery, demonstrating that DNA barcoding is a powerful tool that can identify material properties that influence nanoparticle delivery *in vivo*.

This study complements previous *in vitro* work relating siRNA delivery to the structure of the cationic or ionizable lipid-like compound.[64, 65, 68, 76, 163] This work also supports the idea that LNPs can be rationally designed with cholesterol structures that

closely mimic natural LDL, HDL, or VDLR to improve delivery,[197] or rationally designed to interact with natural cholesterol trafficking pathways. Given that cholesterol trafficking is perturbed in many diseases and as a side effect of commonly prescribed drugs,[168] this also suggests that the efficacy of a LNP may vary with the patient population. One important limitation to this work is that the mechanism by which delivery of LNPs with esterified cholesterol is improved remains unclear. We hypothesize that this effect is mediated by differential interactions with serum proteins and the protein corona.[198] Future studies detailing changes in target cell signaling and protein coronas will be required to confirm or disprove this proposed mechanism.

3.4 Materials and Methods

Nanoparticle Formulation. Nanoparticles were formulated in a microfluidic device by mixing DNA with 7C1, PEG, cholesterol, and a helper lipid, as previously described.[7, 31, 76, 142, 145-149, 160] Nanoparticles were made with variable mole ratios of these constituents. The nucleic acid (*e.g.* DNA barcode, siRNA, sgRNA) was diluted in 10 mM citrate buffer (Teknova) and loaded into a syringe (Hamilton Company). The materials making up the nanoparticle (7C1, cholesterol, PEG, and helper lipid) were diluted in 100% ethanol, and loaded into a second syringe. The citrate phase and ethanol phase were mixed together in a microfluidic device at 600 uL/min and 200 uL/min, respectively. Helper lipids were purchased from Avanti Polar Lipids.

DNA barcoding. Each chemically distinct LNP was formulated to carry its own distinct DNA barcode (**Fig. 3.1A, B**). For example, LNP1 carried DNA barcode 1, while the chemically distinct LNP2 carried DNA barcode 2. The DNA barcodes (IDT) were designed

rationally with universal primer sites and an 8 nucleotide barcode sequence, similar to what we previously described[7, 8]. Three nucleotides on the 5' and 3' ends were modified with phosphorothioates to reduce exonuclease degradation and improve DNA barcode stability. To ensure equal amplification of each sequence, we included universal forward and reverse primer regions on all barcodes. Each barcode was distinguished using a distinct 8nt sequence. An 8nt sequence can generate over 4^8 (65,536) distinct barcodes. We used 156 distinct 8nt sequences designed by to prevent sequence bleaching on the Illumina MiniSeq™ sequencing machine.

Nanoparticle Characterization. LNP hydrodynamic diameter was measured using DLS (DynaPro Plate Reader II, Wyatt). LNPs were diluted in sterile 1X PBS to a concentration of ~0.06 μg/mL, and analyzed. LNPs were included if they met 3 criteria: diameter >20 nm, diameter <200 nm, and autocorrelation function with only 1 inflection point. Particles that met these criteria were pooled and dialyzed in 1X phosphate buffered saline (PBS, Invitrogen), and sterile filtered with a 0.22 μm filter.

Animal Experiments. All animal experiments were performed in accordance with the Georgia Institute of Technology's IACUC. Female C57BL/6J (#000664), $LDLR^{-/-}$ (#002207), $VLDLR^{-/-}$ (#002529), and SpCas9 constitutive mice (#026179) were purchased from the Jackson Laboratory. All mice were purchased from the Jackson Laboratory. In all experiments, mice were aged 5-8 weeks, and N = 3 or 4 mice per group were injected intravenously *via* the lateral tail vein (**Table 3**). The nanoparticle concentration was determined using NanoDrop (Thermo Scientific).

Cell Isolation & Staining. Mice were perfused with 20 mL of 1X PBS through the right atrium. The lungs, spleen, and liver were isolated immediately following perfusion. Tissues were finely cut, and then placed in a digestive enzyme solution with Collagenase Type I (Sigma Aldrich), Collagenase XI (Sigma Aldrich), and Hyaluronidase (Sigma Aldrich) at 37ºC and 550 rpm for 45 minutes. The digestive enzyme for heart and spleen included Collagenase IV (Sigma Aldrich).[76, 145, 148] Digested tissues were passed through a 70 μm filter and red blood cells were lysed. Cells were stained to identify specific cell populations and sorted using the BD FACSFusion cell sorter in the Georgia Institute of Technology Cellular Analysis Core. Antibody clones used for staining were anti-CD31 (390, BioLegend), anti-CD45.2 (104, BioLegend), anti-CD11b (M1/70, BioLegend), and anti-CD102 (3C4, BioLegend).

Endothelial RNAi. C57BL/6J Mice were injected with 1.5 mg/kg siLuciferase or 1.5 mg/kg siICAM2 (AxoLabs). siRNAs were chemically modified at the 2' position to increase stability and specificity and negate immunostimulation. Both siGFP and siICAM2 sequences have been previously reported several times.[76, 145, 148] Seventy-two hours after injection, tissues were isolated and protein expression was quantified as MFI using flow cytometry. ICAM2 MFI in PBS-treated mice was normalized to 100 percent, and all treated groups were compared to this control group.

Endothelial Gene Editing. Constitutive SpCas9 mice were injected with LNP-unmod or LNP-oleate delivering e-sgGFP (AxoLabs) at a dose of 1.0mg/kg. 5 days after injection, tissues were isolated, and cell types were sorted using FACS. DNA was extracted using QuickExtract and sanger sequencing was conducted by Eton Biosciences. Indel formation was measured by TIDE (https://tide-calculator.nki.nl).

PCR Amplification. All samples were amplified and prepared for sequencing using a two-step, nested PCR protocol. More specifically, 1 μL of each primer (10 uM Reverse/Forward) were added to 5 μL of Kapa HiFi 2X master mix, 2 μL sterile H_2O, and 1 μL DNA template. This first PCR reaction was run for 30 cycles. The second PCR, to add Nextera XT chemistry, indices, and i5/i7 adapter regions was run for 5-10 cycles and used the product from 'PCR 1' as template If this initial PCR reaction did not produce clear bands, the primer concentrations, DNA template input, PCR temperature, and number of cycles were optimized for individual samples. The PCR amplicon was isolated using BluePippin (Sage Science).

Deep Sequencing. Illumina deep sequencing was conducted in Georgia Tech's Molecular Evolution core. Runs were performed on an Illumina Miniseq™. Primers were designed based on Nextera XT adapter sequences.

Data Normalization. Counts for each particle, per tissue, were normalized to the barcoded LNP mixture injected into mice, as previously described.[7] This 'input' DNA was used to normalize DNA counts from the cells and tissues.

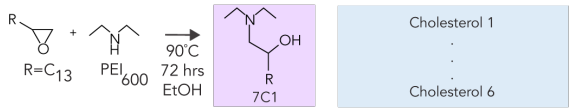
Data Analysis. Sequencing results were processed using a custom python-based tool to extract raw barcode counts for each tissue. These raw counts were then normalized with an R script prior to further analysis. Statistical analyses were done using GraphPad Prism 7. Correlation analyses were run assuming a Gaussian distribution in order to obtain Pearson correlation coefficients. R^2 values (0 – 1) were computed by squaring Pearson correlation coefficients. Data is plotted as mean ± standard error mean unless otherwise stated.

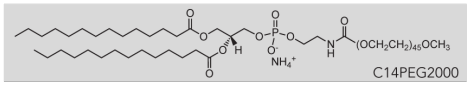
Data Access. The data, analyses, and scripts used to generate all figures in the paper are available upon request to J.E.D. or dahlmanlab.org.

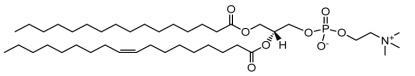
Figure 3.1 High throughput DNA barcoding can be used to test the hypothesis that cholesterol modifications influence LNP delivery *in vivo*. (A) LDL and VLDL particles share physical traits with LNPs, including composition and size. Notably, LDL and VLDL both contain unmodified cholesterol as well as modified cholesterol. (B) Cells naturally interact with (and traffic) LDL and VLDL, suggesting similar mechanisms may alter LNP targeting. (C) A diverse library of 141 LNPs was formulated using 6 cholesterol variants to test the hypothesis that cholesterol structure altered LNP delivery *in vivo*. (D) Each LNP was formulated to carry a distinct DNA barcode, before (E) stable LNPs were pooled together and administered to either WT, LDLR$^{-/-}$, or VLDLR$^{-/-}$ mice. After isolating 6 cell types from each mouse, delivery mediated by all LNPs was measured concurrently using DNA sequencing. (F) Hydrodynamic diameter, measured by DLS, for all individual LNPs included, as well as the diameter of the LNP pool after mixing. (G, H) Normalized delivery for the negative control (naked barcode) - averaged across all 18 samples - was much lower than normalized delivery for all LNPs.

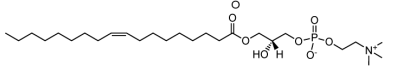
A

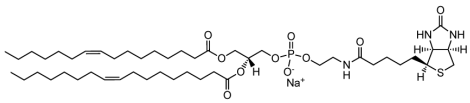
LNP #	Lipomer	Cholesterol	PEG MW	PEG Chain	Helper Lipid	Lipomer Mole %	Cholesterol Mole %	PEG Mole %	Helper Lipid Mole %
1	7C1	4B-OH Cholesterol	2000	C14	-	62	30	8	0
2	7C1	4B-OH Cholesterol	2000	C14	-	35	60	5	0
3	7C1	4B-OH Cholesterol	2000	C14	-	62	22	16	0
4	7C1	4B-OH Cholesterol	2000	C14	-	35	35	30	0
5	7C1	No Cholesterol	2000	C14	-	80	0	20	0
6	7C1	4B-OH Cholesterol	2000	C14	-	80	15	5	0
7	7C1	4B-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	40	2	8
8	7C1	4B-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	29	11	10
9	7C1	4B-OH Cholesterol	2000	C14	18:1 Lyso PC	50	29	11	10
10	7C1	4B-OH Cholesterol	2000	C14	18:1 Lyso PC	50	40	2	8
11	7C1	4B-OH Cholesterol	2000	C14	Biotinyl PE	50	40	2	8
12	7C1	4B-OH Cholesterol	2000	C14	Biotinyl PE	50	29	11	10
13	7C1	4B-OH Cholesterol	2000	C14	DC Chol-HCl	50	29	11	10
14	7C1	4B-OH Cholesterol	2000	C14	DC Chol-HCl	50	40	2	8
15	7C1	4B-OH Cholesterol	2000	C14	DMPC	50	29	11	10
16	7C1	4B-OH Cholesterol	2000	C14	DMPC	50	40	2	8
17	7C1	4B-OH Cholesterol	2000	C14	DOPC	50	29	11	10
18	7C1	4B-OH Cholesterol	2000	C14	DOPC	50	40	2	8
19	7C1	4B-OH Cholesterol	2000	C14	DOPE	50	40	2	8
20	7C1	4B-OH Cholesterol	2000	C14	DOPE	50	29	11	10
21	7C1	4B-OH Cholesterol	2000	C14	DOTAP	50	29	11	10
22	7C1	4B-OH Cholesterol	2000	C14	DOTAP	50	40	2	8
23	7C1	4B-OH Cholesterol	2000	C14	O Chems PC	50	29	11	10
24	7C1	4B-OH Cholesterol	2000	C14	O Chems PC	50	40	2	8
25	7C1	7A-OH Cholesterol	2000	C14	-	62	30	8	0
26	7C1	7A-OH Cholesterol	2000	C14	-	80	15	5	0
27	7C1	No Cholesterol	2000	C14	-	80	0	20	0
28	7C1	7A-OH Cholesterol	2000	C14	-	35	60	5	0
29	7C1	7A-OH Cholesterol	2000	C14	-	62	22	16	0
30	7C1	7A-OH Cholesterol	2000	C14	-	35	35	30	0
31	7C1	7A-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	29	11	10
32	7C1	7A-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	40	2	8
33	7C1	7A-OH Cholesterol	2000	C14	18:1 Lyso PC	50	29	11	10
34	7C1	7A-OH Cholesterol	2000	C14	18:1 Lyso PC	50	40	2	8
35	7C1	7A-OH Cholesterol	2000	C14	Biotinyl PE	50	40	2	8
36	7C1	7A-OH Cholesterol	2000	C14	DC Chol-HCl	50	29	11	10
37	7C1	7A-OH Cholesterol	2000	C14	DC Chol-HCl	50	40	2	8
38	7C1	7A-OH Cholesterol	2000	C14	DMPC	50	40	2	8
39	7C1	7A-OH Cholesterol	2000	C14	DMPC	50	29	11	10
40	7C1	7A-OH Cholesterol	2000	C14	DOPC	50	29	11	10
41	7C1	7A-OH Cholesterol	2000	C14	DOPC	50	40	2	8
42	7C1	7A-OH Cholesterol	2000	C14	DOPE	50	29	11	10
43	7C1	7A-OH Cholesterol	2000	C14	DOPE	50	40	2	8
44	7C1	7A-OH Cholesterol	2000	C14	DOTAP	50	29	11	10
45	7C1	7A-OH Cholesterol	2000	C14	O Chems PC	50	29	11	10
46	7C1	7B-OH Cholesterol	2000	C14	-	80	15	5	0
47	7C1	No Cholesterol	2000	C14	-	80	0	20	0
48	7C1	7B-OH Cholesterol	2000	C14	-	62	30	8	0
49	7C1	7B-OH Cholesterol	2000	C14	-	35	60	5	0
50	7C1	7B-OH Cholesterol	2000	C14	-	35	35	30	0
51	7C1	7B-OH Cholesterol	2000	C14	-	62	22	16	0
52	7C1	7B-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	40	2	8
53	7C1	7B-OH Cholesterol	2000	C14	16:0 - 18:1 PC	50	29	11	10
54	7C1	7B-OH Cholesterol	2000	C14	18:1 Lyso PC	50	40	2	8
55	7C1	7B-OH Cholesterol	2000	C14	18:1 Lyso PC	50	29	11	10
56	7C1	7B-OH Cholesterol	2000	C14	Biotinyl PE	50	29	11	10
57	7C1	7B-OH Cholesterol	2000	C14	Biotinyl PE	50	40	2	8
58	7C1	7B-OH Cholesterol	2000	C14	DC Chol-HCl	50	40	2	8
59	7C1	7B-OH Cholesterol	2000	C14	DC Chol-HCl	50	29	11	10
60	7C1	7B-OH Cholesterol	2000	C14	DMPC	50	40	2	8
61	7C1	7B-OH Cholesterol	2000	C14	DMPC	50	29	11	10
62	7C1	7B-OH Cholesterol	2000	C14	DOPC	50	29	11	10
63	7C1	7B-OH Cholesterol	2000	C14	DOPC	50	40	2	8
64	7C1	7B-OH Cholesterol	2000	C14	DOPE	50	40	2	8
65	7C1	7B-OH Cholesterol	2000	C14	DOPE	50	29	11	10
66	7C1	7B-OH Cholesterol	2000	C14	DOTAP	50	29	11	10
67	7C1	7B-OH Cholesterol	2000	C14	DOTAP	50	40	2	8
68	7C1	7B-OH Cholesterol	2000	C14	DSPC	50	29	11	10
69	7C1	7B-OH Cholesterol	2000	C14	DSPC	50	40	2	8
70	7C1	7B-OH Cholesterol	2000	C14	O Chems PC	50	29	11	10
71	7C1	7B-OH Cholesterol	2000	C14	O Chems PC	50	40	2	8
72	7C1	Cholesterol	2000	C14	-	62	30	8	0
73	7C1	No Cholesterol	2000	C14	-	80	0	20	0
74	7C1	Cholesterol	2000	C14	-	80	15	5	0
75	7C1	Cholesterol	2000	C14	-	35	60	5	0
76	7C1	Cholesterol	2000	C14	-	35	35	30	0
77	7C1	Cholesterol	2000	C14	-	62	22	16	0
78	7C1	Cholesterol	2000	C14	16:0 - 18:1 PC	50	29	11	10
79	7C1	Cholesterol	2000	C14	16:0 - 18:1 PC	50	40	2	8
80	7C1	Cholesterol	2000	C14	18:1 Lyso PC	50	40	2	8
81	7C1	Cholesterol	2000	C14	18:1 Lyso PC	50	29	11	10
82	7C1	Cholesterol	2000	C14	Biotinyl PE	50	29	11	10
83	7C1	Cholesterol	2000	C14	Biotinyl PE	50	40	2	8
84	7C1	Cholesterol	2000	C14	DC Chol-HCl	50	29	11	10
85	7C1	Cholesterol	2000	C14	DC Chol-HCl	50	40	2	8
86	7C1	Cholesterol	2000	C14	DMPC	50	29	11	10
87	7C1	Cholesterol	2000	C14	DMPC	50	40	2	8
88	7C1	Cholesterol	2000	C14	DOPC	50	40	2	8
89	7C1	Cholesterol	2000	C14	DOPC	50	29	11	10
90	7C1	Cholesterol	2000	C14	DOPE	50	40	2	8
91	7C1	Cholesterol	2000	C14	DOPE	50	29	11	10
92	7C1	Cholesterol	2000	C14	DOTAP	50	29	11	10
93	7C1	Cholesterol	2000	C14	DOTAP	50	40	2	8
94	7C1	Cholesterol	2000	C14	DSPC	50	40	2	8
95	7C1	Cholesterol	2000	C14	DSPC	50	29	11	10
96	7C1	Cholesterol	2000	C14	O Chems PC	50	40	2	8
97	7C1	Cholesterol	2000	C14	O Chems PC	50	29	11	10
98	7C1	Cholesteryl Oleate	2000	C14	-	62	30	8	0
99	7C1	No Cholesterol	2000	C14	-	80	0	20	0
100	7C1	Cholesteryl Oleate	2000	C14	-	35	60	5	0
101	7C1	Cholesteryl Oleate	2000	C14	-	35	35	30	0
102	7C1	Cholesteryl Oleate	2000	C14	-	80	15	5	0
103	7C1	Cholesteryl Oleate	2000	C14	-	62	22	16	0
104	7C1	Cholesteryl Oleate	2000	C14	16:0 - 18:1 PC	50	29	11	10
105	7C1	Cholesteryl Oleate	2000	C14	16:0 - 18:1 PC	50	40	2	8
106	7C1	Cholesteryl Oleate	2000	C14	18:1 Lyso PC	50	29	11	10
107	7C1	Cholesteryl Oleate	2000	C14	18:1 Lyso PC	50	40	2	8
108	7C1	Cholesteryl Oleate	2000	C14	Biotinyl PE	50	29	11	10
109	7C1	Cholesteryl Oleate	2000	C14	Biotinyl PE	50	40	2	8
110	7C1	Cholesteryl Oleate	2000	C14	DC Chol-HCl	50	40	2	8
111	7C1	Cholesteryl Oleate	2000	C14	DC Chol-HCl	50	29	11	10
112	7C1	Cholesteryl Oleate	2000	C14	DMPC	50	29	11	10
113	7C1	Cholesteryl Oleate	2000	C14	DMPC	50	40	2	8
114	7C1	Cholesteryl Oleate	2000	C14	DOPC	50	40	2	8
115	7C1	Cholesteryl Oleate	2000	C14	DOPC	50	29	11	10
116	7C1	Cholesteryl Oleate	2000	C14	DOPE	50	29	11	10
117	7C1	Cholesteryl Oleate	2000	C14	DOPE	50	40	2	8
118	7C1	Cholesteryl Oleate	2000	C14	DOTAP	50	29	11	10
119	7C1	Cholesteryl Oleate	2000	C14	DOTAP	50	40	2	8
120	7C1	Cholesteryl Oleate	2000	C14	DSPC	50	29	11	10
121	7C1	Cholesteryl Oleate	2000	C14	DSPC	50	40	2	8
122	7C1	Cholesteryl Oleate	2000	C14	O Chems PC	50	40	2	8
123	7C1	Cholesteryl Oleate	2000	C14	O Chems PC	50	29	11	10
124	7C1	Cholesteryl Stearate	2000	C14	-	62	22	16	0
125	7C1	No Cholesterol	2000	C14	-	80	0	20	0
126	7C1	Cholesteryl Stearate	2000	C14	-	80	15	5	0
127	7C1	Cholesteryl Stearate	2000	C14	-	62	30	8	0
128	7C1	Cholesteryl Stearate	2000	C14	-	35	35	30	0
129	7C1	Cholesteryl Stearate	2000	C14	16:0 - 18:1 PC	50	29	11	10
130	7C1	Cholesteryl Stearate	2000	C14	16:0 - 18:1 PC	50	40	2	8
131	7C1	Cholesteryl Stearate	2000	C14	18:1 Lyso PC	50	29	11	10
132	7C1	Cholesteryl Stearate	2000	C14	DC Chol-HCl	50	40	2	8
133	7C1	Cholesteryl Stearate	2000	C14	DC Chol-HCl	50	29	11	10
134	7C1	Cholesteryl Stearate	2000	C14	DMPC	50	40	2	8
135	7C1	Cholesteryl Stearate	2000	C14	DOPC	50	29	11	10
136	7C1	Cholesteryl Stearate	2000	C14	DOPC	50	40	2	8
137	7C1	Cholesteryl Stearate	2000	C14	DOPE	50	40	2	8
138	7C1	Cholesteryl Stearate	2000	C14	DOTAP	50	29	11	10
139	7C1	Cholesteryl Stearate	2000	C14	DSPC	50	40	2	8
140	7C1	Cholesteryl Stearate	2000	C14	DSPC	50	29	11	10
141	7C1	Cholesteryl Stearate	2000	C14	O Chems PC	50	40	2	8

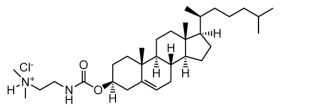


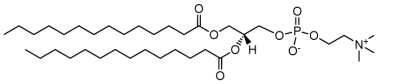


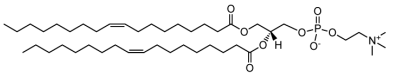


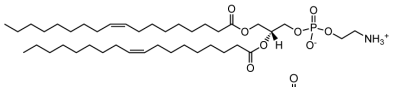


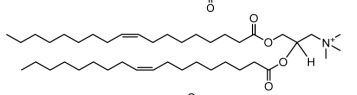


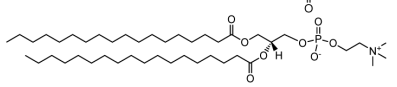


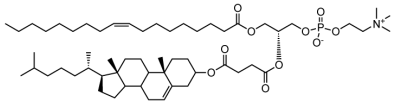


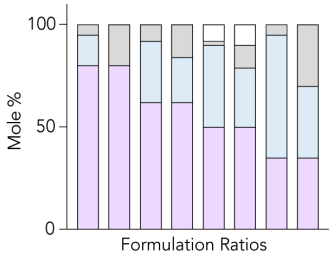




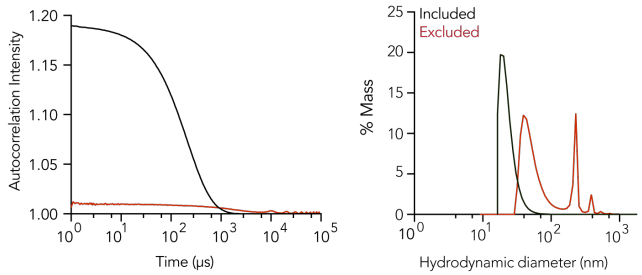




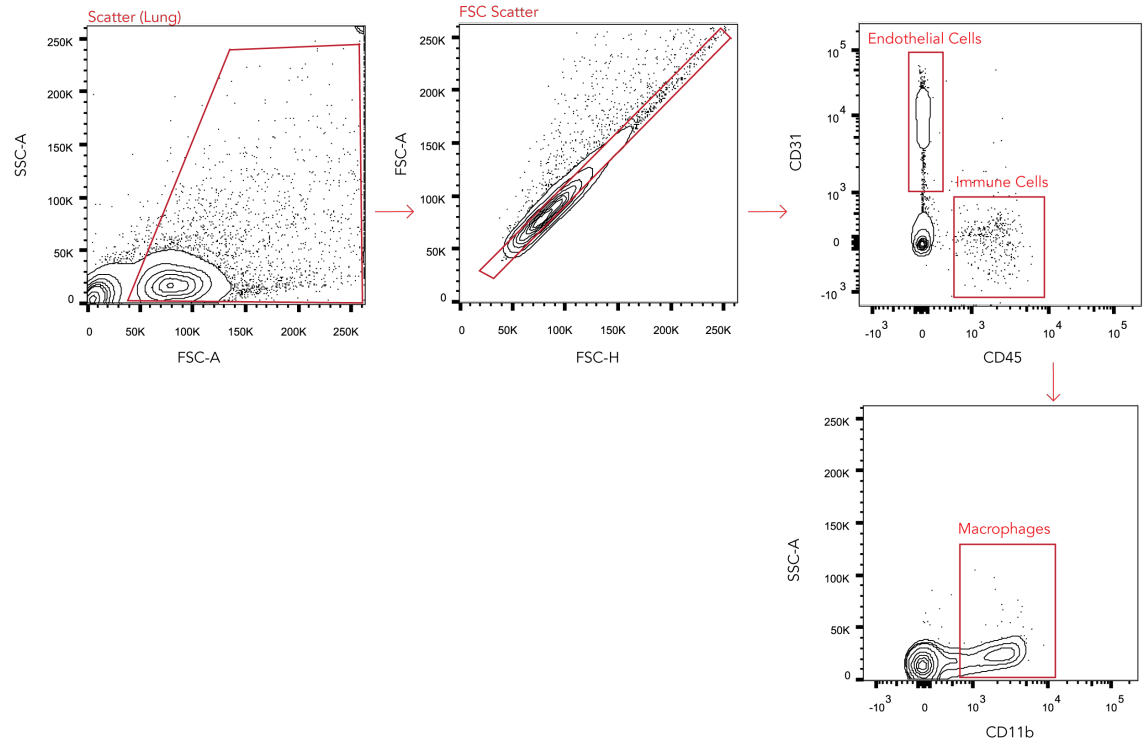




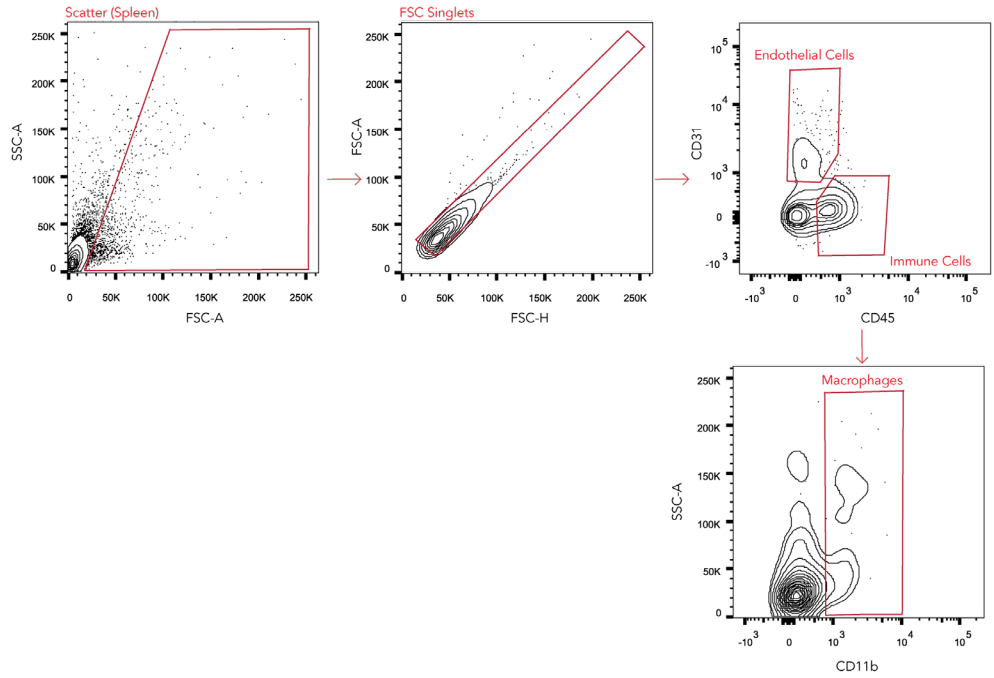
B


Autocorrelation Intensity
1.20
1.15
1.10
1.05
1.00
Time (µs)
% Mass
Included
Excluded
Hydrodynamic diameter (nm)

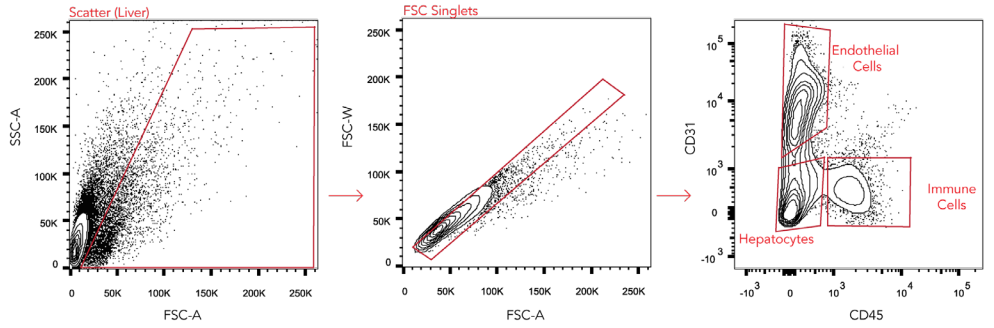
C


Scatter (Lung)
SSC-A
FSC-A
FSC Scatter
FSC-A
FSC-H
CD31
Endothelial Cells
Immune Cells
CD45
Macrophages
SSC-A
CD11b

D


Scatter (Spleen)
SSC-A
FSC-A
FSC Singlets
FSC-A
FSC-H
Endothelial Cells
Immune Cells
CD31
CD45
Macrophages
SSC-A
CD11b

E


Scatter (Liver)
SSC-A
FSC-A
FSC Singlets
FSC-W
FSC-A
Endothelial Cells
Immune Cells
Hepatocytes
CD31
CD45

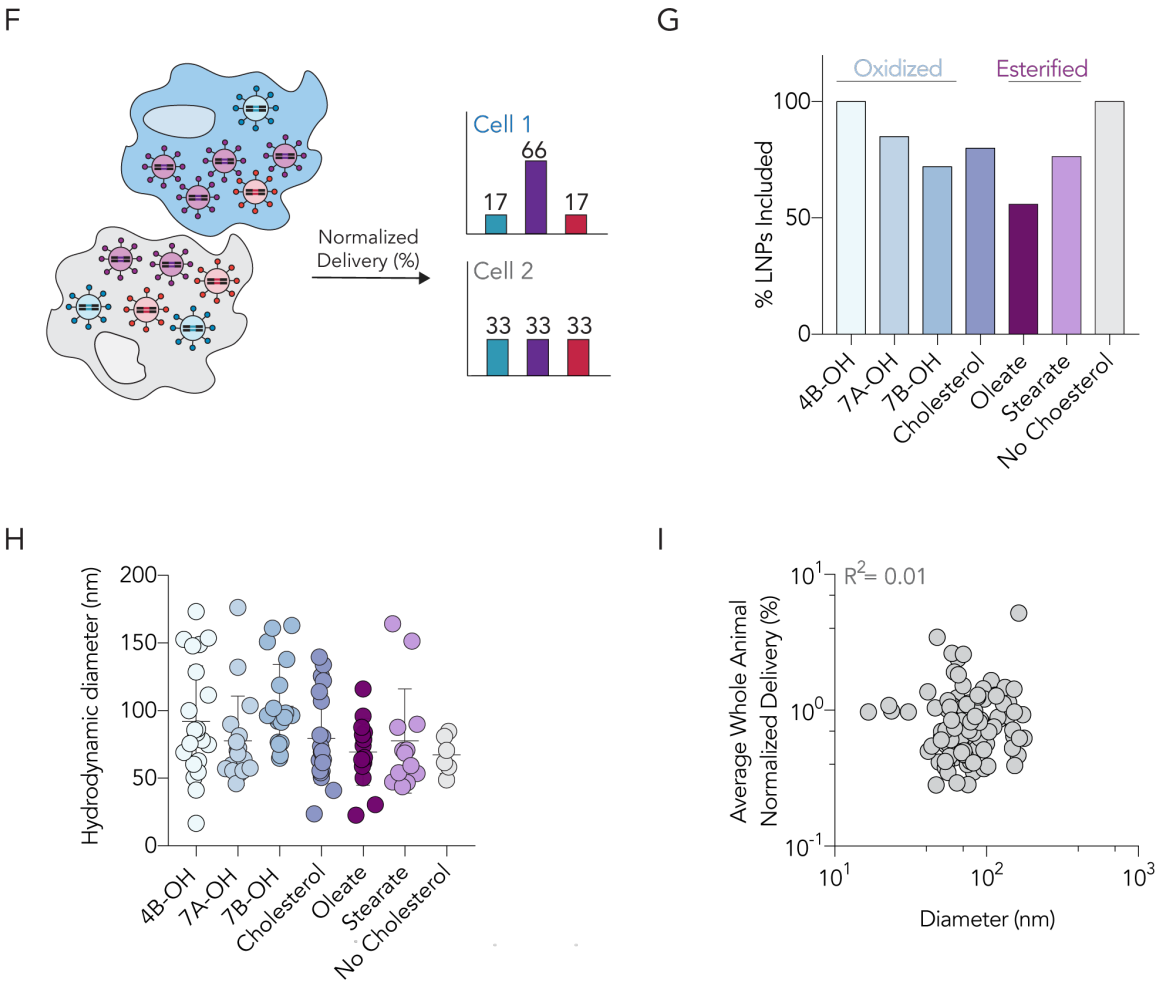


Figure 3.2 (A) A library of 141 LNPs was formulated to test 6 cholesterol variants using 8 different molar ratios. (B) LNPs were pooled based on whether or not they met our inclusion criteria: autocorrelation curves with 1 inflection point and hydrodynamic diameters between 20 nm and 200 nm, based on DLS. Representative FACS gates for sorting of lung (C) and spleen (D) endothelial cells and macrophages, as well as liver (E) endothelial cells and hepatocytes. (F) Data normalization after sequencing of individual cell types yields relative distribution of different LNPs to cell types and tissues. (G) More than 50% of LNPs of every cholesterol type met our inclusion criteria and were pooled and analyzed. (H) Diameter distributions for all LNPs, subdivided by cholesterol type. (I) Diameter plotted against distribution across all WT cell types in all tissues.

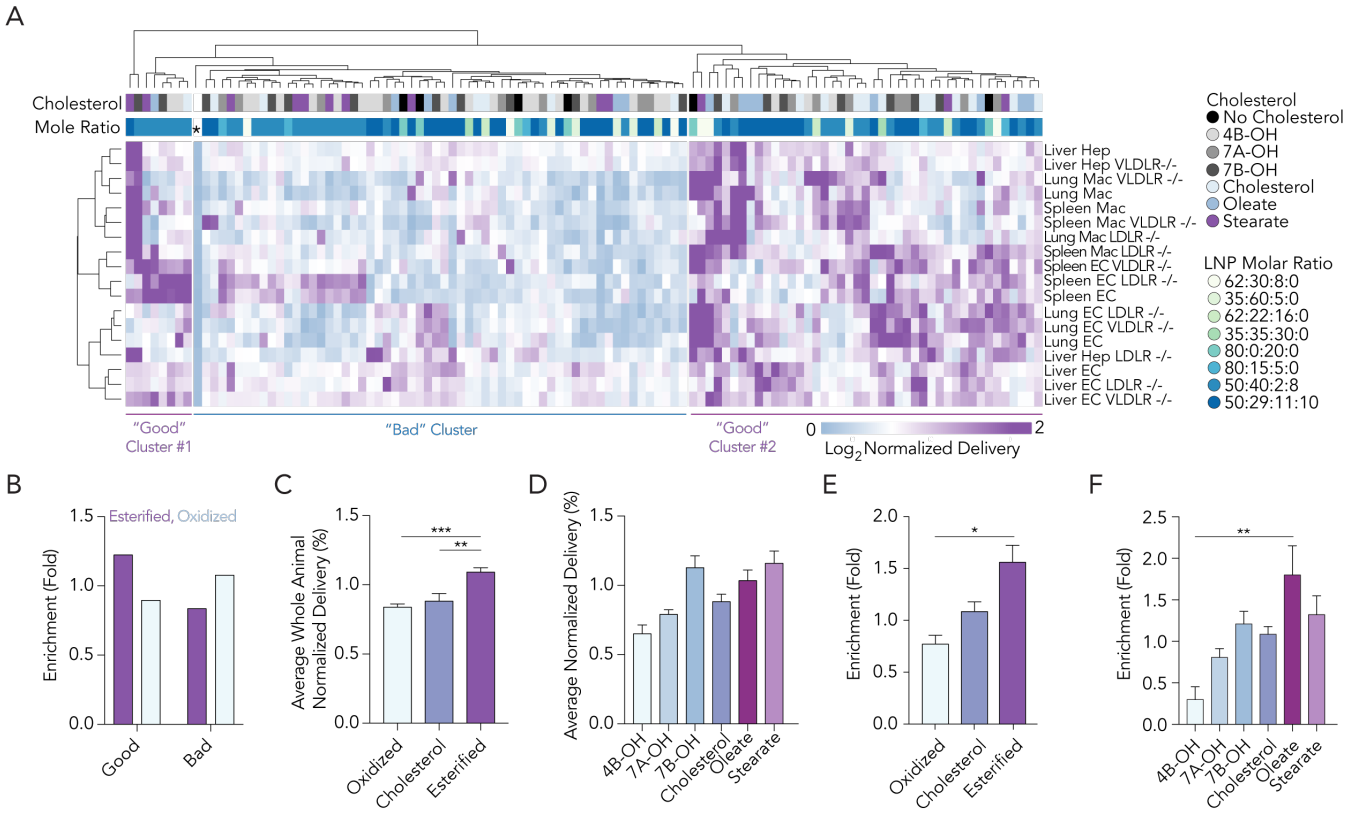


Figure 3.3 LNPs made with esterified cholesterol outperform LNPs made with oxidized cholesterol *in vivo*. (A) A nanoparticle targeting heatmap depicting normalized delivery generated by unbiased Euclidean clustering. This dataset contains nearly 2,000 *in vivo* drug delivery data points. The negative control (*) performed worse than all LNPs. LNP delivery was divided into 3 horizontal clusters. Lung, liver, and spleen endothelial cells (ECs), lung and spleen macrophages (Macs) and liver hepatocytes (Hep) are clustered vertically. (B) Enrichment of esterified and oxidized cholesterols in the left-/right-most (good) clusters and center-most (bad) cluster. LNPs made with esterified and oxidized cholesterol were more likely to be found in good and bad clusters, respectively. (C, D) Normalized delivery for all LNPs in WT mice, subdivided by the cholesterol type. (E, F) Enrichment in the top 15% of LNPs, subdivided by the cholesterol type. *p>0.0332, **p<0.0021, ***p<0.0002, 1-way ANOVA.

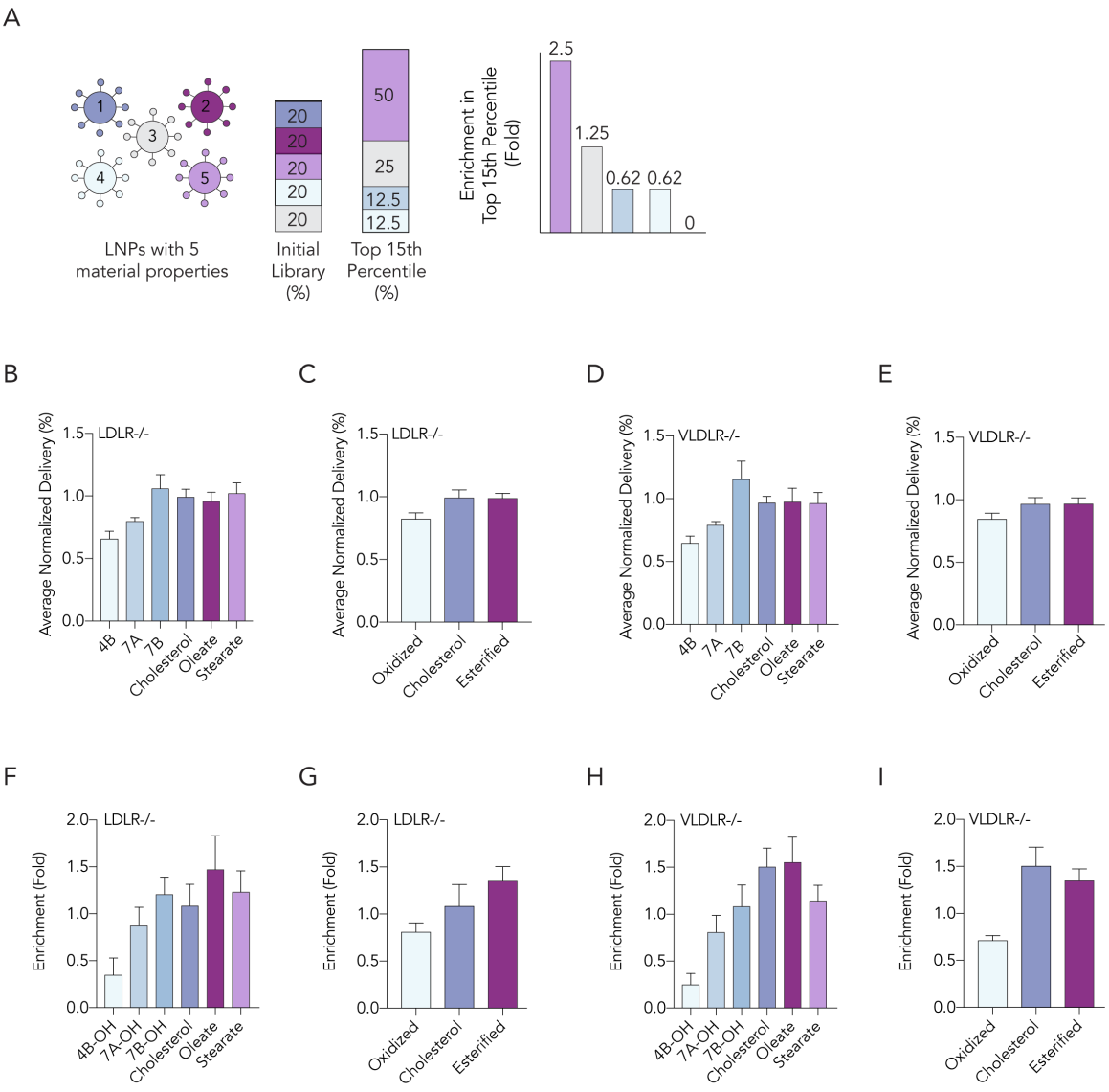


Figure 3.4 Cholesterol structure can alter nanoparticle delivery *in vivo.* (A) Schematic (not real data; cartoon data simplified for this example) showing enrichment for 5 LNP properties; initially, each LNP property is represented 20% of the time in all LNPs. However, property 5 is represented 50% of the time in the top 15% of LNPs. We would consider this a 2.5-fold enrichment in the top 15%. Normalized delivery for all LNPs in (B, C) LDLR$^{-/-}$ and (D,E) VLDLR$^{-/-}$ mice, subdivided by cholesterol type. Enrichment in the top 15% of LNPs in (F, G) LDLR$^{-/-}$ and (H,I) VLDLR$^{-/-}$ mice, subdivided by cholesterol type.

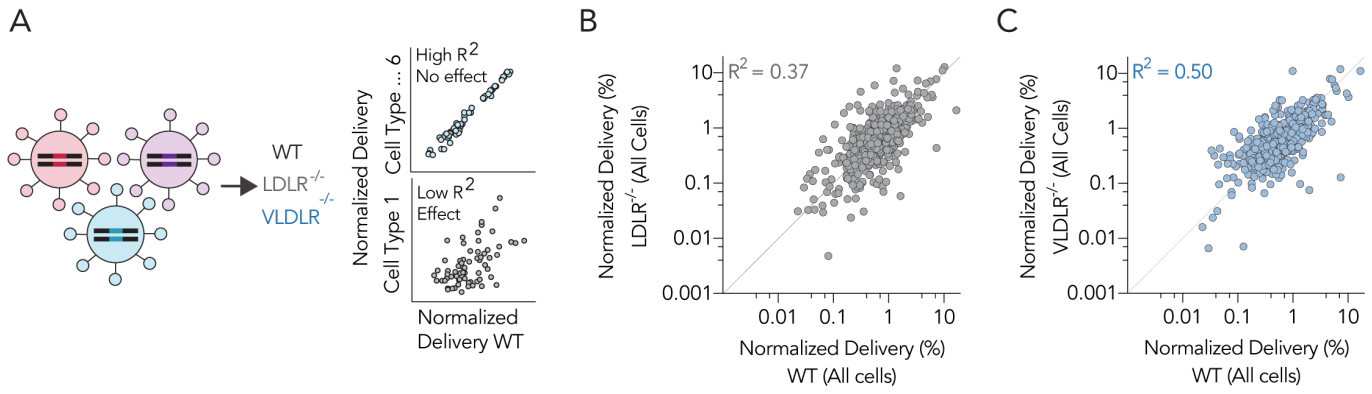


Figure 3.5 LDLR and VLDLR affect LNP *in vivo* delivery globally. (A) To quantify the extent to which LDLR and VLDLR influenced LNP delivery, we quantified the correlation between delivery for all LNPs in all 6 cell types (>650 data points per mouse model). Normalized delivery in (B) LDLR$^{-/-}$ and (C) VLDLR$^{-/-}$ knockout mice plotted against normalized delivery in WT mice. The R^2 values strongly suggest that both genes affect LNP targeting *in vivo*, and that LDLR affects delivery slightly more than VLDLR.

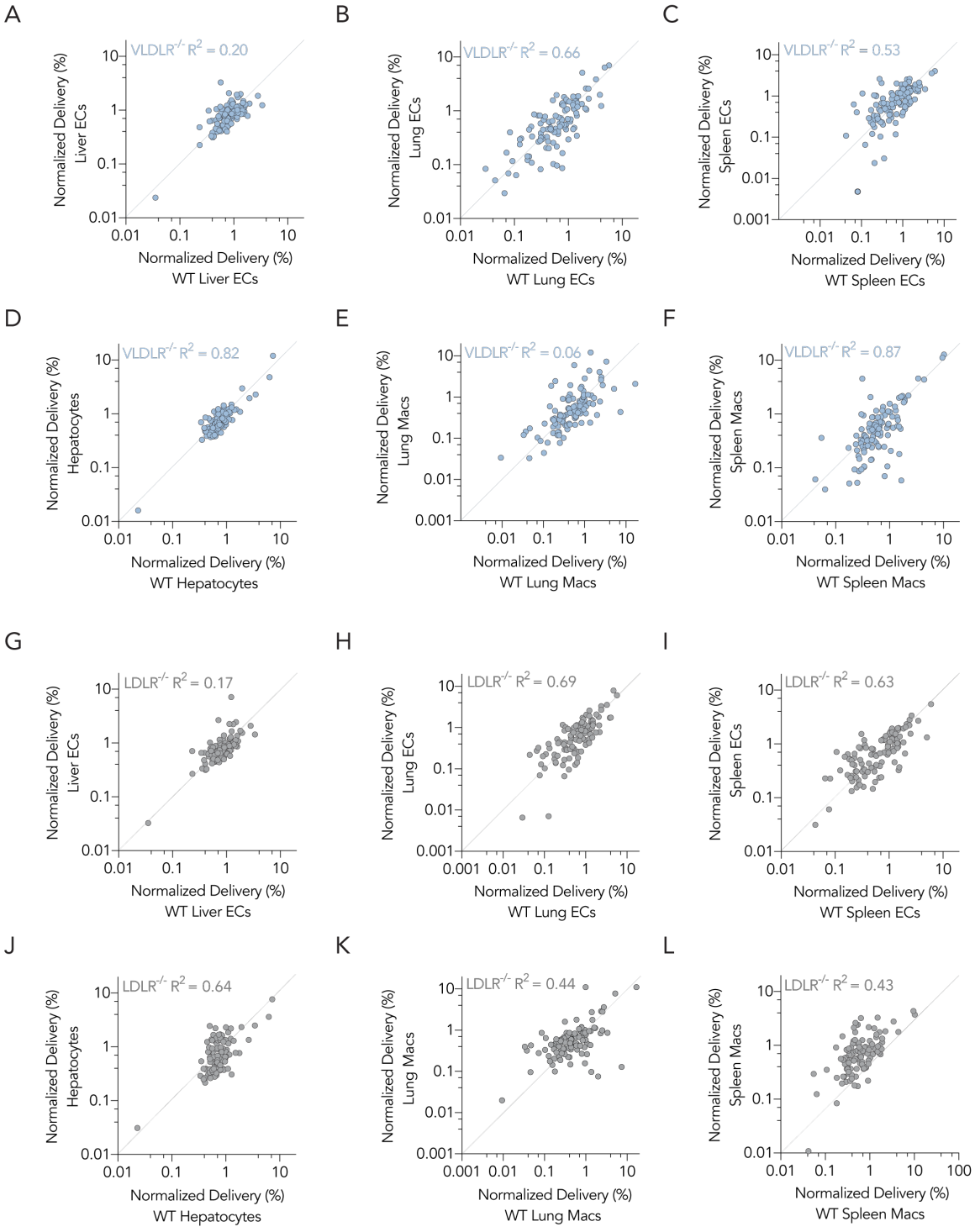


Figure 3.6 Normalized delivery in (A-F) VLDLR$^{-/-}$ mice and (G-L) LDLR$^{-/-}$ mice plotted against normalized delivery in WT mice for endothelial cells (ECs), macrophages (Macs), and Hepatocytes.

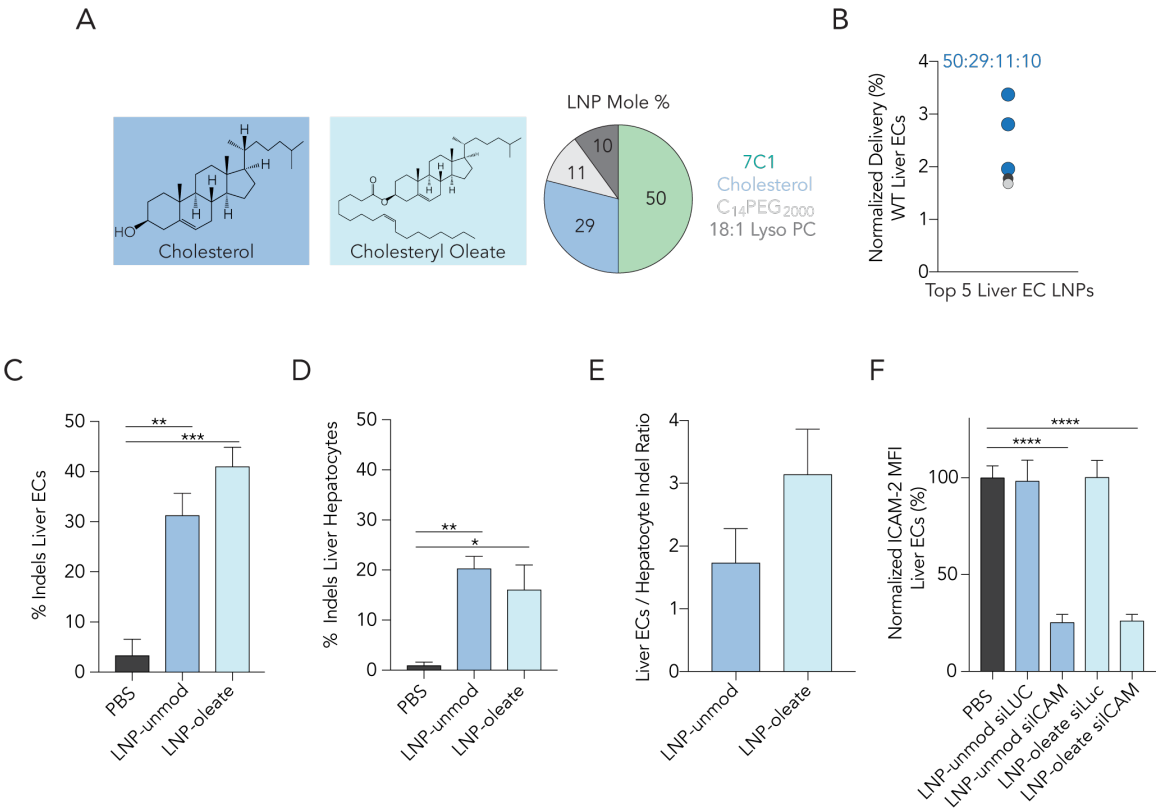


Figure 3.7 LNPs formulated with cholesteryl oleate deliver therapeutic RNAs as – or more – efficiently than LNPs formulated with unmodified cholesterol. (A, B) Based on the DNA barcoding screen, LNPs with a 50:29:11:10 molar ratio of 7C1: cholesterol: $C_{14}PEG_{2000}$: 18:1 Lyso PC were highly enriched in hepatic endothelial cells. We formulated 2 LNPs with this molar ratio; LNP-oleate contained cholesteryl oleate, where LNP-unmod contained unmodified cholesterol, the current gold standard in the field. (C-D) Indel percentage from (C) hepatic endothelial cells and (D) hepatocytes isolated 5 days after a single intravenous injection of either LNP-oleate or LNP-unmod. (E) Interestingly, LNP-oleate delivery led to 3X more editing in hepatic ECs relative to hepatocytes. (F) ICAM-2 MFI 3 days after an injection of PBS (control) siRNA targeting Luciferase (control) or siICAM-2. Robust ICAM-2 protein silencing was observed in siICAM-2 treated mice, but not siLuc treated mice. * p< 0.0332, **p<0.0021, ***p<0.0002, ****p<0.0001, 1-way ANOVA.

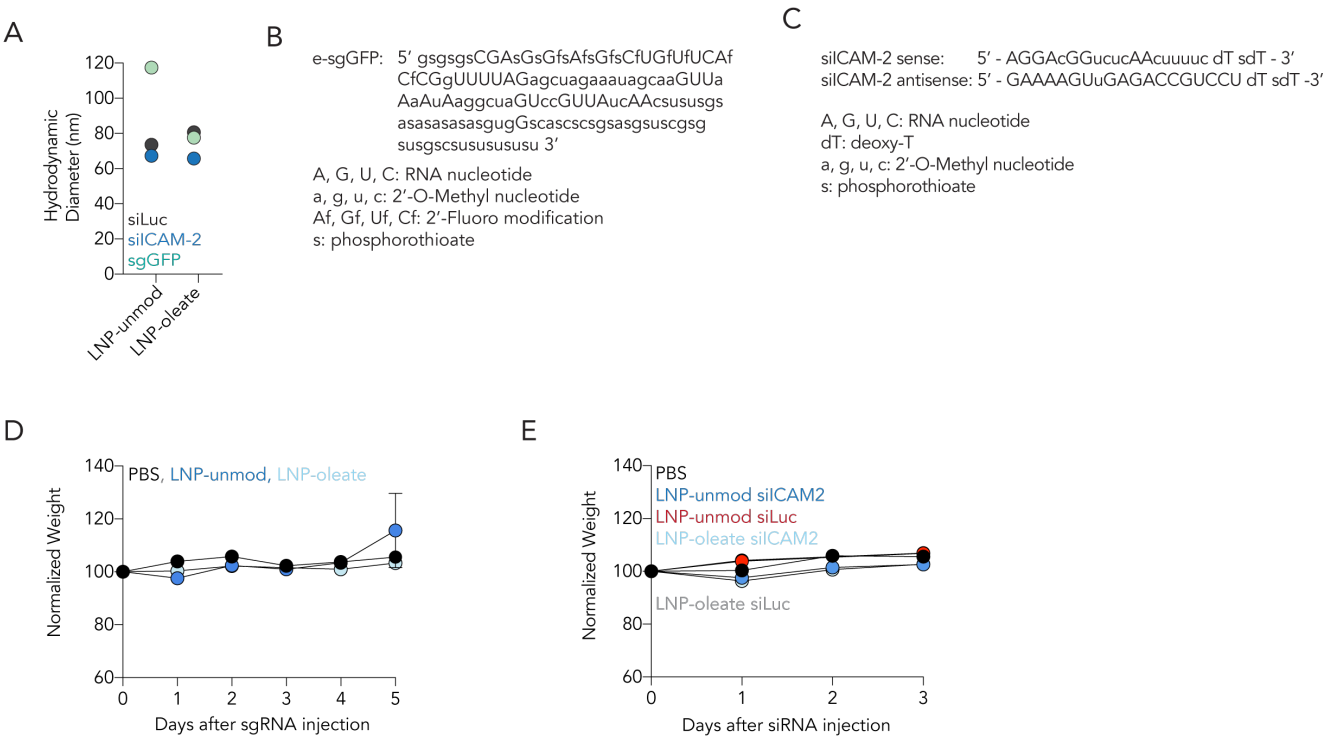


Figure 3.8 (A) LNP-unmod and LNP-oleate formed small, stable LNPs when formulated to carry (B) modified sgRNA targeting the GFP locus or (C) modified siRNA against ICAM-2. (D) Normalized mouse weights 0-5 days post sgRNA containing LNP-unmod and LNP-oleate injection. (E) Normalized mouse weights for 0, 1, and 2 days post siRNA containing LNP-unmod and LNP-oleate injection. Mouse weights are always normalized to mouse weight immediately before injection.

Table 1 Calculation showing the # of *in vivo* data points.

# of LNPs Tested + Naked BC	# of Cell Types Tested	# of Mouse Models	Total Data Points
111	6	3	1998

Table 2 Significance data for Figure 3.3D.

WT Mice (Normalized Delivery) Fig 2D			
Comparison	Significant?	Summary	Adjusted P Value
7A vs. 4B	No	ns	0.6595
7B vs. 4B	Yes	**	0.0013
Cholesterol vs. 4B	No	ns	0.1688
Oleate vs. 4B	Yes	**	0.0013
Stearate vs. 4B	Yes	***	0.0005
7B vs. 7A	Yes	*	0.0375
Cholesterol vs. 7A	No	ns	0.7502
Oleate vs. 7A	Yes	*	0.0375
Stearate vs. 7A	Yes	*	0.0156
Cholesterol vs. 7B	No	ns	0.27
Oleate vs. 7B	No	ns	0.9965
Stearate vs. 7B	No	ns	0.9755
Oleate vs. Cholesterol	No	ns	0.27
Stearate vs. Cholesterol	No	ns	0.1603
Stearate vs. Oleate	No	ns	0.9755

Table 3 The strain, sex, and age of each mouse as well as the number of mice used in each experiment and the type of procedure done on each group of mice.

Strain	Sex	Procedure	# of mice (N)	Age
C57BL/6	F	IV Injection (Inj.) LNPs	4	5-8 weeks
VLDLR-/-	F	IV Inj. LNPs	3	5-8 weeks
LDLR-/-	F	IV Inj. LNPs	3	5-8 weeks
Strain	Sex	Procedure	N	Age
C57BL/6	F	IV Injection (Inj.) PBS	4	5-8 weeks
C57BL/6	F	IV Inj. LNP-unmod siICAM	4	5-8 weeks
C57BL/6	F	IV Inj. LNP-oleate siICAM	4	5-8 weeks
C57BL/6	F	IV Inj. LNP-unmod siLuc	4	5-8 weeks
C57BL/6	F	IV Inj. LNP-oleate siLuc	4	5-8 weeks
Strain	Sex	Procedure	N	Age
Cas9-GFP	F	IV Injection (Inj.)PBS	4	5-8 weeks
Cas9-GFP	F	IV Inj. sgGFP LNP-unmod	4	5-8 weeks
Cas9-GFP	F	IV Inj. sgGFP LNP-oleate	4	5-8 weeks

CHAPTER 4. LIPID NANOPARTICLES CONTAINING OXIDIZED CHOLESTEROL EFFICIENTLY DELIVER MRNA TO THE LIVER MICROENVIRONMENT

The work presented here is an excerpt from Paunovska K, Da Silva Sanchez AJ, Sago CD, Lokugamage MP, Islam F, Kalathoor S, Krupczak BR, Dahlman JE (2019). "Nanoparticles Containing Oxidized Cholesterol Deliver mRNA to the Liver Microenvironment at Clinically Relevant Doses." *Advanced Materials.*

4.1 Introduction

There have been several advances in local and systemic mRNA delivery[199], especially to hepatocytes[192, 193]. However, low dose mRNA delivery to other cell types within the liver microenvironment remains challenging. Using a high-throughput DNA barcode screen named JORDAN, we observed that *in vitro* nanoparticle delivery can be a poor predictor of *in vivo* delivery[8]. Since *in vitro* assays typically lack a complete immune system[200], variable blood flow[99], heterogeneous vasculature, off-target cells, and other factors that influence delivery *in vivo*[98], we reasoned that an ideal nanoparticle experiment would study LNPs directly *in vivo.* We therefore developed FIND[9] to quantify how over 100 LNPs deliver functional mRNA into the cytoplasm of target cells in a single mouse.

LNPs are formulated by mixing nucleic acids with ionizable lipids, PEG-lipids, phospholipids, and cholesterol. The diversity of available chemistries enables scientists to evaluate how LNP structure affects delivery. However, these studies are mostly performed *in vitro*[71, 139] and use LNPs formulated with unmodified cholesterol. Several lines of evidence led us to hypothesize that cholesterol chemical structure affected LNP mRNA delivery. First, cholesterol structure is modified via oxidation; this requires specific

enzymes, demonstrating that cells regulate cholesterol structure[201, 202]. Second, oxidized cholesterols signal differently than unmodified cholesterols, altering interactions with membrane and nuclear receptors that regulate gene expression[203]. Third, cholesterol is trafficked in lipoproteins using forward and reverse transport[204] and trafficking to endothelial cells, hepatocytes, and macrophages may change with cholesterol structure *in vivo*[165]. Fourth, LNPs and lipoproteins have similar size and composition. Finally, using JORDAN, we found that modified cholesterols alter LNP biodistribution[78], however this study measured LNP biodistribution; it did not measure functional mRNA delivery. Given that less than 4% of exogenously delivered RNA escapes endosomes[88] and endosomal escape changes with cell type[205], it is difficult to predict functional mRNA delivery using biodistribution.

To test the hypothesis that cholesterol structure affects LNP delivery of mRNA, we created a library of LNPs with nine cholesterol variants. We chose variants that were (i) produced naturally or synthetically and (ii) had oxidative modifications on sterol ring B, the hydrocarbon tail attached to sterol ring D, or both. We formulated LNPs to carry DNA barcodes and Cre mRNA[9], systemically injected them into Cre reporter mice, and analyzed the efficiency with which they delivered Cre mRNA in 28 cells types *in vivo*. We identified an LNP that delivers mRNA to Kupffer cells and hepatic endothelial cells five-fold more potently than to hepatocytes. Notably, the lead LNPs performed as predicted by the FIND screen, suggesting that FIND can rapidly identify LNPs for mRNA delivery.

4.2 Results

4.2.1 Rationally designed nanoparticles can be screened for mRNA delivery

Unmodified cholesterol is acted upon by enzymes to form variants that are side-chain or ring oxidized (**Fig. 4.1A**). To investigate whether these modifications altered LNP targeting, we formulated 125 FIND LNPs using microfluidics (**Fig. 4.1B**). To minimize variation from components other than cholesterol, LNPs were made of the ionizable lipid-like material cKK-E12[71], two well validated PEG-lipids, the phospholipid DOPE, and one of nine different cholesterol variants (**Fig. 4.1B**).

We formed LNPs by mixing the contents together in a microfluidic device[142]. Each LNP carried a unique DNA barcode, which acted as a tag for that LNP, as well as Cre mRNA, which signals functional mRNA delivery (**Fig. 4.1C**). Stable LNPs were pooled together (**Fig. 4.1D**) and administered to Ai14 mice at a total nucleic acid dose of 1.0 mg/kg. Ai14 mice contain a LSL-tdTomato construct under the control of a CAG promoter; as a result, cells in Ai14 mice become tdTomato$^+$ if: (1) Cre mRNA is delivered into the cytoplasm, (2) Cre mRNA is translated into Cre protein, (3) Cre protein translocates from the cytoplasm into the nucleus, and (4) Cre protein edits the genome by removing the 'Stop' between Lox sites. Therefore, by isolating tdTomato$^+$ cell types using FACS and using NGS to quantify barcodes within them, FIND identifies LNPs located in cells where functional mRNA delivery occurred (**Fig. 4.1E**). We quantified NGS sequencing data as 'normalized delivery', analogous to counts per million in RNA-seq[9] (**Fig. 4.1E**).

As a quality control, we analyzed the size of each individual LNP. LNPs were only pooled if their hydrodynamic diameter was between 20 and 215 nm and their autocorrelation curve contained one inflection point (**Fig. 4.2A, B**). Eighty-six of the 125 LNPs we formulated met these criteria and were pooled. As a control, we compared the diameters of all 86 LNPs to the diameter of the pooled LNP solution, and found they were

similar (**Fig. 4.3C**). This suggests that the pooled LNPs did not aggregate after mixing; we have previously observed that pooled LNPs are stable[9]. We added a naked DNA barcode as a negative control[8] since naked DNA does not readily enter cells. After isolating cells and performing NGS, we found that - as expected - the naked DNA was delivered into cells less frequently than all the DNA barcodes delivered by LNPs (**Fig. 4.3D, E**).Mice were sacrificed 72 hours after we injected the pooled LNPs[9]; this timepoint allows cells to express tdTomato after Cre mRNA delivery (**Fig. 4.3A**).

We isolated the liver, spleen, heart, kidney, pancreas, lung, and bone marrow. Using FACS, we then isolated 28 different tdTomato$^+$ cell types (**Fig. 4.3A**). Comparing the percentage of tdTomato$^+$ cell types in different organs, cells in the liver tended to be targeted more than cells in other organs (**Fig. 4.3B**). The organ with the second-highest percentage of tdTomato$^+$ cells was the spleen; the remaining five organs had negligible delivery (**Fig. 4.2C**). We then clustered the barcode sequencing data using an unbiased Euclidean algorithm. This bioinformatics technique is regularly applied to gene expression data[189] and can analyze nanoparticle barcoding data[206]. Euclidean clustering revealed that the 4 liver cell types tended to 'cluster' together more closely than they did to splenic cell types (**Fig. 4.3F**).

4.2.2 Modifying cholesterol can alter nanoparticle functional delivery in vivo

Unexpectedly, we found that the percentage of tdTomato$^+$ hepatic endothelial cells, hepatic immune cells, and Kupffer cells was much higher than the percentage of tdTomato$^+$ hepatocytes (**Fig. 4.3B**). This result was surprising, however, the stability of the pooled LNPs, the poor performance of the naked barcode, the high percentage of tdTomato$^+$ cells

in the liver relative to other organs, and the Euclidean clustering convinced us the data were reliable enough to test our hypothesis: LNP cholesterol structure alters mRNA delivery *in vivo.* To quantify the impact of cholesterol structure on overall splenic and hepatic delivery, we calculated the normalized delivery across all 8 cell types we sequenced (4 in spleen and 4 in liver) (**Fig. 4.4A**). We noted that side-chain oxidized cholesterol variants tended to enhance delivery relative to the other cholesterol variants (**Fig. 4.4B, C**). In particular, LNPs formulated with 25-hydroxycholesterol (25-OH) and 20α-hydroxycholesterol (20α-OH) resulted in higher normalized delivery across all 8 cell types (**Fig. 4.4D**). To complement these calculations, we assessed which cholesterol variants were enriched in the top 10% of LNPs (**Fig. 4.4E**). We then calculated the enrichment in the bottom 10% of LNPs[69] and subtracted it from enrichment in the top 10% (**Fig. 4.4E,F**). This identifies how likely it is that a given cholesterol variant is found in the best- and worst-performing LNPs. Enrichment calculations are detailed in **Fig. 4.2A**. As an additional control, we performed the same two analyses – normalized delivery across all 8 cell types, and enrichment – for cholesterol mole percentage, ionizable lipid mole percentage, and phospholipid mole percentage (**Fig. 4.4B-D**). We did not observe any significant trends.

These data suggested that LNP cholesterol chemical composition was an important factor in LNP targeting. However, they did not take LNP size into account. We previously found no relationship between nanoparticle size and delivery for hydrodynamic diameters between ~20 and ~200 nm[26,34]. To investigate how size altered delivery in this experiment, we calculated whether normalized delivery varied with LNP size for all LNPs (**Fig. 4.4E**), LNPs with tail oxidized cholesterols (**Fig. 4.4F**), and LNPs with ring oxidized cholesterols (**Fig. 4.4G**); we found no relationship. We then calculated the percentage of formulated

LNPs that met our inclusion criteria as a function of cholesterol structure and the average size of stable LNPs based on cholesterol structure and found no significant differences (**Fig. 4.4H, I**). We performed the same analyses as a function of cholesterol mole percentage and reached the same conclusions (**Fig. 4.4J, K**). Thus, we did not find evidence that size affected LNP delivery between 20 and 220 nm. However, we cannot exclude the possibility that size could affect delivery if very small or very large nanoparticles are analyzed.

4.2.3 LNPs formulated with 20a-OH deliver mRNA at clinically relevant doses

Like all high-throughput screens, the utility of FIND is governed by its ability to make predictions that can be verified. To this end, we formulated our top 3 LNP candidates identified by FIND with Cre mRNA. Of the 3 LNPs formulated, one LNP – which contained 25-OH cholesterol - did not formulate consistently and was excluded. The remaining two LNPs – which contained 20α-OH cholesterol (**Fig. 4.5G, H, Fig. 4.6F, G**) – formed stable LNPs, and were administered to Ai14 mice at a total mRNA dose of 0.25 mg/kg. Encouragingly, we found that both LNPs recapitulated the results from the FIND screen (**Fig. 4.5I-L**). Both LNP1 and LNP2 robustly targeted cells in the microenvironment after a 0.25 mg/kg injection; as predicted by the screen, hepatocytes were targeted far less efficiently. LNP1 did not cause mouse weight loss in any experiment; LNP2 did cause mice to lose weight at 0.25 mg/kg (**Fig. 4.8B**). Encouraged by robust delivery at 0.25 mg/kg, we injected LNP1 and LNP2 at a dose of 0.05 mg/kg.

Once again, we found robust delivery (**Fig. 4.5I-L**). Whereas LNP1 and LNP2 both performed equally well at 0.25 mg/kg by saturating our Cre-based system, LNP1 outperformed LNP2 at the lower dose of 0.05 mg/kg. We then investigated whether LNP1

formulated with unmodified cholesterol performed as effectively as LNP1 formulated with 20α-OH **(Fig. 4.5M)**. Three days after injecting mice with 0.05 mg/kg Cre mRNA, we found that 20α-OH improved delivery to the liver, across all hepatic cell types **(Fig. 4.5N-Q)**. As an additional control, we assessed the potency of LNP1 compared to a previously reported LNP with potency at 0.2 mg/kg **(Fig. 4.5M)**; LNP1 outperformed this positive control LNP **(Fig. 4.5N-Q)** between 14- and 77-fold within the liver microenvironment. We then used QUANT[207], a highly sensitive biodistribution system. to assess LNP biodistribution, for LNP1, LNP2, and LNP1-Chol. We did not find a difference in LNP biodistribution within the liver microenvironment **(Fig. 4.6D, E)**.

4.3 Discussion

Systemic delivery of therapeutic RNA to hepatocytes has led to an FDA approved drug[208]. Delivery to non-hepatocytes has remained more challenging. Thus, an unbiased, high throughput method to study how LNPs deliver RNA *in vivo* could accelerate the discovery of nanoparticles with new tropisms. Here we report that FIND can predict delivery of LNPs to the liver microenvironment. If this holds true for other cell types, then FIND could reduce the time and resources needed to identify clinically relevant LNPs, relative to traditional *in vitro* screening. Future studies will also be needed to identify therapeutic mRNA that – when delivered to Kupffer cells – will alleviate disease.

We found that LNPs containing oxidized cholesterol can deliver mRNA to cells in the liver microenvironment more potently than to hepatocytes. Notably, robust delivery occurred at 0.05 mg/kg, which is below the dose regime used for siRNA therapies in humans[208]. Given the importance of liver endothelial cells[196] and Kupffer cells[209] in

disease, these data suggest that additional advances in delivery could eventually result in protein replacement therapies within the liver microenvironment. We observed that oxidative modifications made to the hydrocarbon tail were more well tolerated than those made to the B cholesterol ring.

These results need to be repeated with more ionizable lipids before they can be considered a general statement. However, in the context of these experiments, cholesterol structure impacts functional delivery. We find it feasible that different protein coronas may adsorb onto LNPs based on cholesterol structure, thereby altering targeting. In this paper, one limitation is that we did not identify the mechanism; although future work is needed to substantiate this hypothesis, we believe these data are an exciting first step to understanding how cholesterol structure can be rationally altered to change LNP tropism.

4.4 Materials and Methods

Synthesis of CKK-E12. Compound **1** (20 g, 41.9 mmol) was added into a 100 ml flask and trifluoroacetic acid (42 mL) was added slowly at 0 ºC, then stirred at room temperature for 30 min. The solvent was evaporated under reduced pressure and then the crude product dissolved in DMF (5 mL) was added dropwise to pyridine (300 mL) at 0 ºC. The reaction mixture was stirred at room temperature overnight. The solvents were evaporated under reduced pressure to afford a white solid and washed with ethyl acetate to give pure compound **2** (8.4 g, 13.04 mmol, 31% yield) (**Fig. 4.6B**). To a solution of compound **2** in acetic acid/CH_2Cl_2 (150/150 mL) was added Pd/C (10 wt. %, 3.0 g). The black suspension was degassed for 5 min with hydrogen and stirred at room temperature under hydrogen atmosphere overnight. The reaction mixture was filtered by celite and washed with MeOH.

The combined filtrates were concentrated to obtain a crude yellow viscous oil. The oil was solidified by adding ethyl acetate and washed with ethyl acetate to yield compound **3** (**Fig. 4.6C**). To a solution of compound **3** (84 mg, 0.22 mmol) and 1,2-epoxydodecane (247 mg, 1.34 mmol) in EtOH (2 mL) was added triethylamine (0.12 mL, 0.88 mmol). After stirring for 30 min at room temperature, the reaction mixture was then irradiated in the microwave reactor at 150 ºC for 5 h. Purification of the crude residue via flash column chromatography (**Fig. 4.6C**). After purifying the version with four alkyl tails using flash column chromatography, we confirmed its chemical structure using ^{1}H-NMR (**Fig. 4.6D**).

Nanoparticle Formulation. Nanoparticles were formulated in a microfluidic device by mixing Cre mRNA, DNA, the ionizable lipid, PEG, and cholesterol as previously described[142]. Nanoparticles were made with variable mole ratios of these constituents. The nucleic acid (*e.g.* DNA barcode, mRNA) was diluted in 10 mM citrate buffer (Teknova) and loaded into a syringe (Hamilton Company). The materials making up the nanoparticle (CKK-E12, cholesterol, PEG, DOPE) were diluted in ethanol, and loaded into a second syringe. The citrate phase and ethanol phase were mixed together in a microfluidic device using syringe pumps.

DNA barcoding. Each chemically distinct LNP was formulated to carry its own distinct DNA barcode. For example, LNP1 carried Cre mRNA and DNA barcode 1, whereas the chemically distinct LNP2 carried Cre mRNA and DNA barcode 2. The DNA barcodes were designed rationally with universal primer sites and a specific 8 nucleotide barcode sequence, similar to what we previously described[7]. To ensure equal amplification of each sequence, we included universal forward and reverse primer regions on all barcodes. An 8nt sequence can generate over 4^{8} (65,536) distinct barcodes.

Nanoparticle Characterization. LNP hydrodynamic diameter was measured using DLS. LNPs were diluted in sterile 1X PBS to a concentration of ~0.06 μg/mL, and analyzed. LNPs were included if they met 3 criteria: diameter >20 nm, diameter <215 nm, and autocorrelation function with only 1 inflection point. Particles that met these criteria were pooled and dialyzed in 1X phosphate buffered saline (PBS, Invitrogen), and sterile filtered with a 0.22 μm filter.

Animal Experiments. All animal experiments were performed in accordance with the Georgia Institute of Technology's IACUC. C57BL/6J (#000664) and Ai14 LSL-Tomato (#007914) mice were purchased from the Jackson Laboratory. In all experiments, mice were aged 5-8 weeks, and N = 3-4 mice per group were injected intravenously *via* the lateral tail vein. The nanoparticle concentration was determined using NanoDrop (Thermo Scientific). Weights for all mice are included in **Fig. 4.7A-D**.

Cell Isolation & Staining. In all cases, mice were sacrificed 3 days after administration of LNPs and immediately perfused with 20 mL of 1X PBS through the right atrium. Organs were isolated immediately following perfusion. Tissues were cut, and then placed in a digestive enzyme solution with Collagenase Type I (Sigma Aldrich), Collagenase XI (Sigma Aldrich), and Hyaluronidase (Sigma Aldrich) at 37ºC and 750 rpm for 45 minutes. The digestive enzyme for heart and spleen included Collagenase IV (Sigma Aldrich)[76]. Digested tissues were passed through a 70 μm filter and red blood cells were lysed. Cells were stained to identify specific cell populations and sorted using a BD FACSFusion cell sorter. Antibody clones used for staining were: anti-CD31 (390, BioLegend), anti-CD45.2 (104, BioLegend), anti-CD11b (M1/70, BioLegend), anti-CD68 (FA-11, BioLegend), anti-CD3 (17A2, BioLegend), anti-CD19 (6D5, BioLegend), anti-CD34 (SA376A4,

BioLegend). Representative gating strategies for liver and spleen cell populations are included in **Fig. 4.8A-D**.

PCR Amplification. All samples were amplified and prepared for sequencing using a nested PCR protocol. More specifically, 1 μL of each primer (10 uM Reverse/Forward) were added to 5 μL of Kapa HiFi 2X master mix, 2 μL sterile H_2O, and 1 μL DNA template. The second PCR added Nextera XT chemistry, indices, and i5/i7 adapter regions and used the product from 'PCR 1' as template.

Deep Sequencing. Illumina deep sequencing was performed on Illumina MiniSeq™ using standard protocols suggested by Illumina. The sequencing was conducted in the Georgia Tech Molecular Evolution core.

Data Normalization. Counts for each particle were normalized to the barcoded LNP mixture injected into mice, as we previously described[9].

Data Analysis. Sequencing results were processed using a custom python-based tool to extract raw barcode counts for each tissue. These raw counts were then normalized with an R script prior to further analysis. Statistical analyses were done using GraphPad Prism 7. Data is plotted as mean ± standard error mean unless otherwise stated.

Data Access. All data are available upon request to james.dahlman@bme.gatech.edu or dahlmanlab.org.

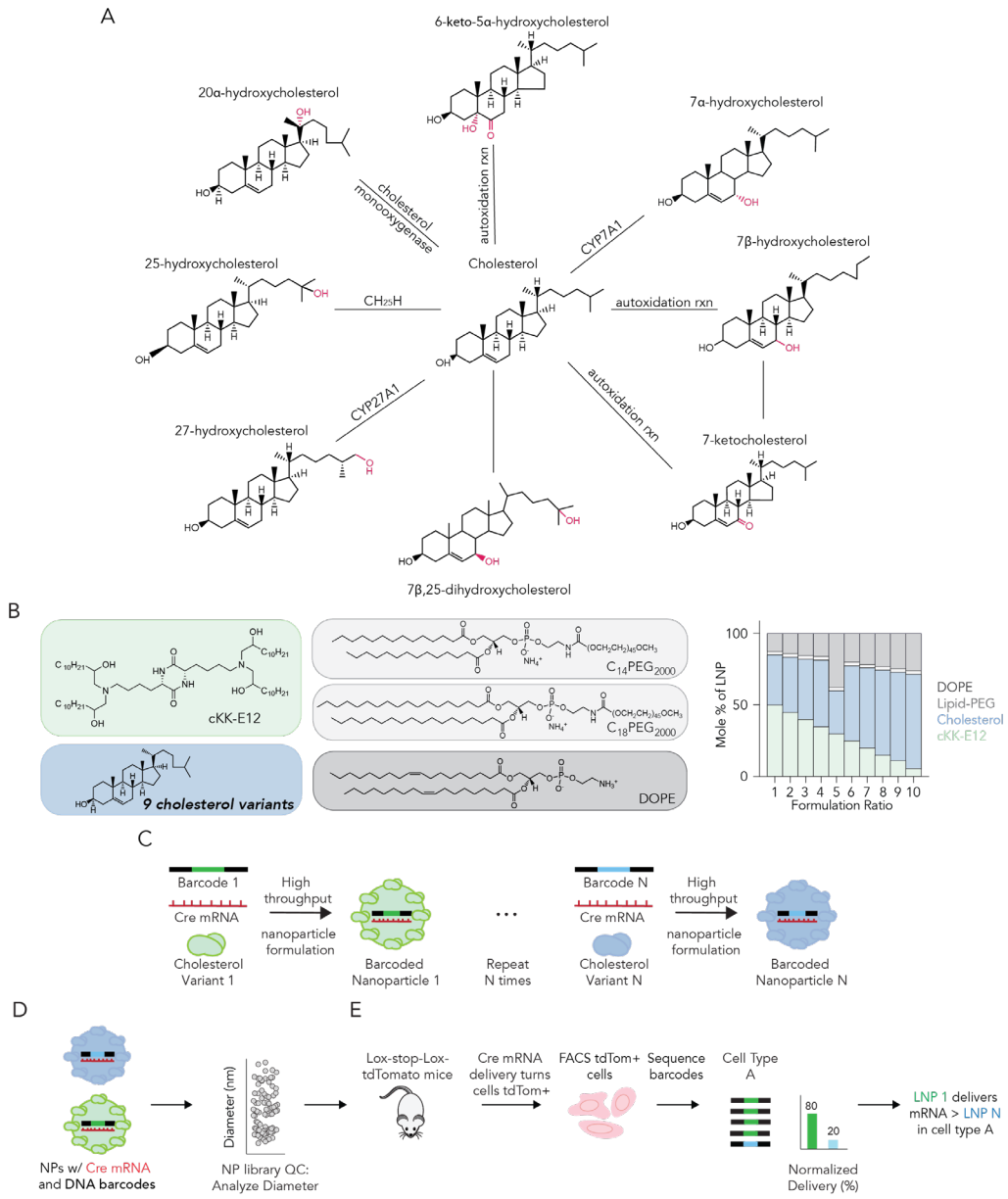


Figure 4.1 Synthesizing a library of rationally designed nanoparticles containing oxidized cholesterols. (A) Cholesterol can be modified to form sterol variants that differentially act upon a number of biological pathways. Here we investigate whether cholesterol variants affected delivery. (B) We formulated 125 nanoparticles by combining the ionizable material cKK-E12, PEG-lipids, DOPE, and 9 different cholesterol variants. (C) These distinct nanoparticles were then barcoded and screened for functional mRNA delivery using FIND. (D) After formulation, barcoded nanoparticles are put through DLS-based quality control, pooled, and (E) intravenously administered to Ai14 mice. If Cre mRNA is delivered into the cytoplasm and translated into Cre protein that edits the genome, the cells become tdTomato+. tdTomato+ cells are isolated using FACS and barcodes are sequenced to identify nanoparticles that co-localize with cells transfected with Cre *in vivo.* This enables us to study how many LNPs deliver mRNA in a single mouse.

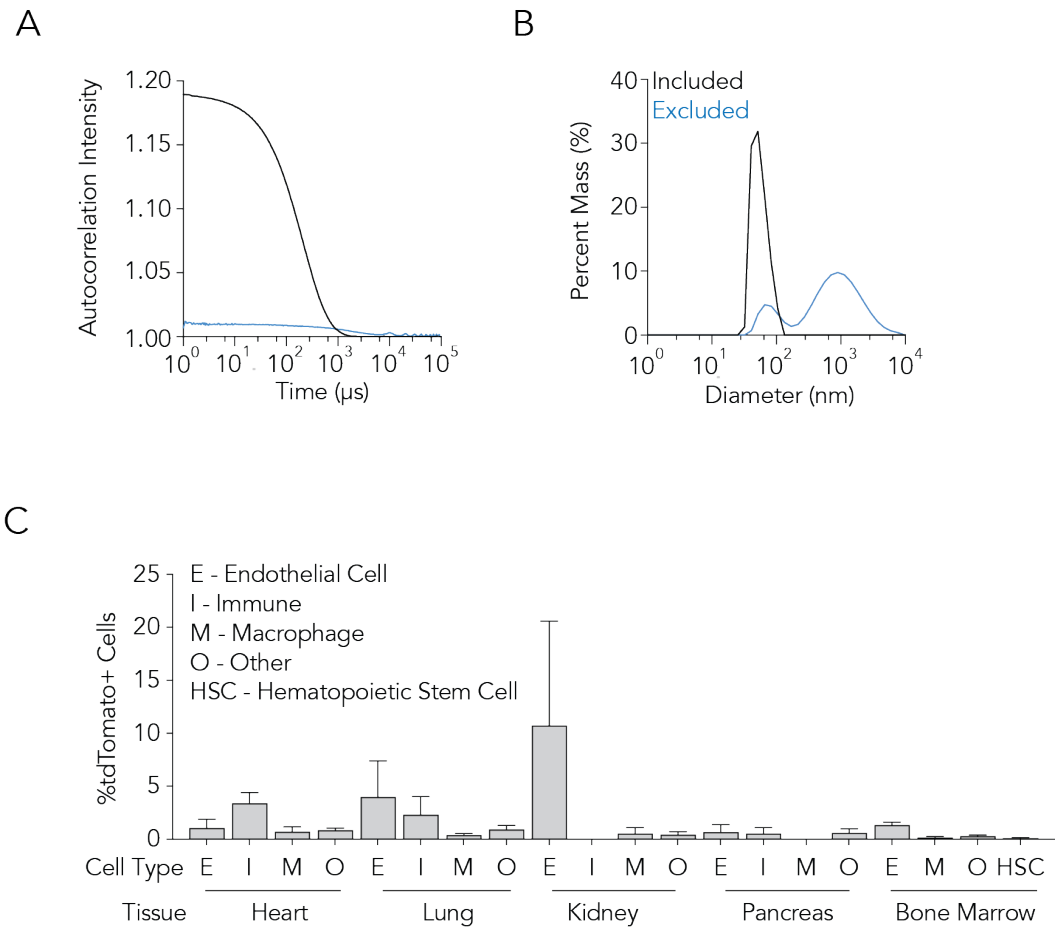
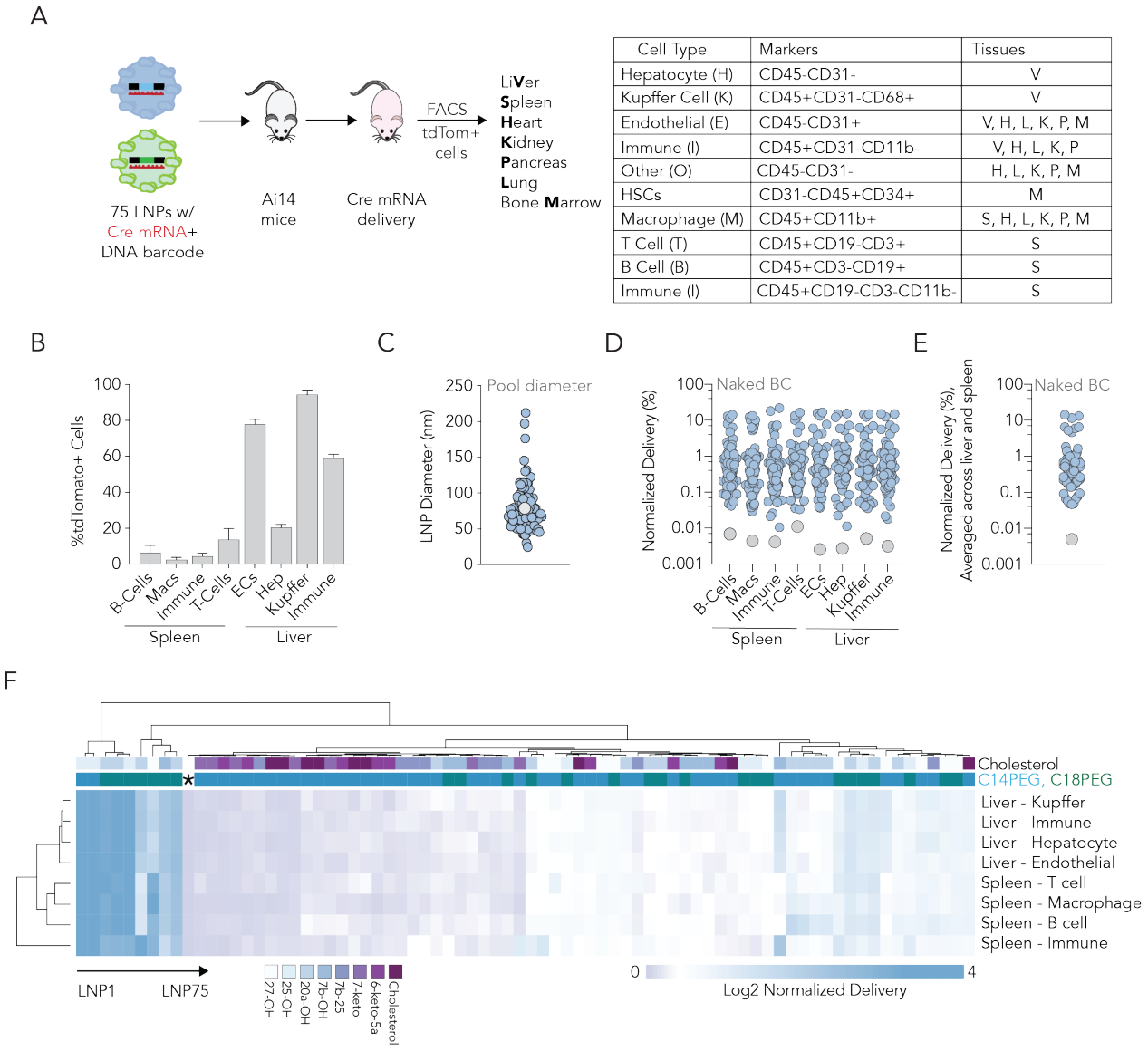
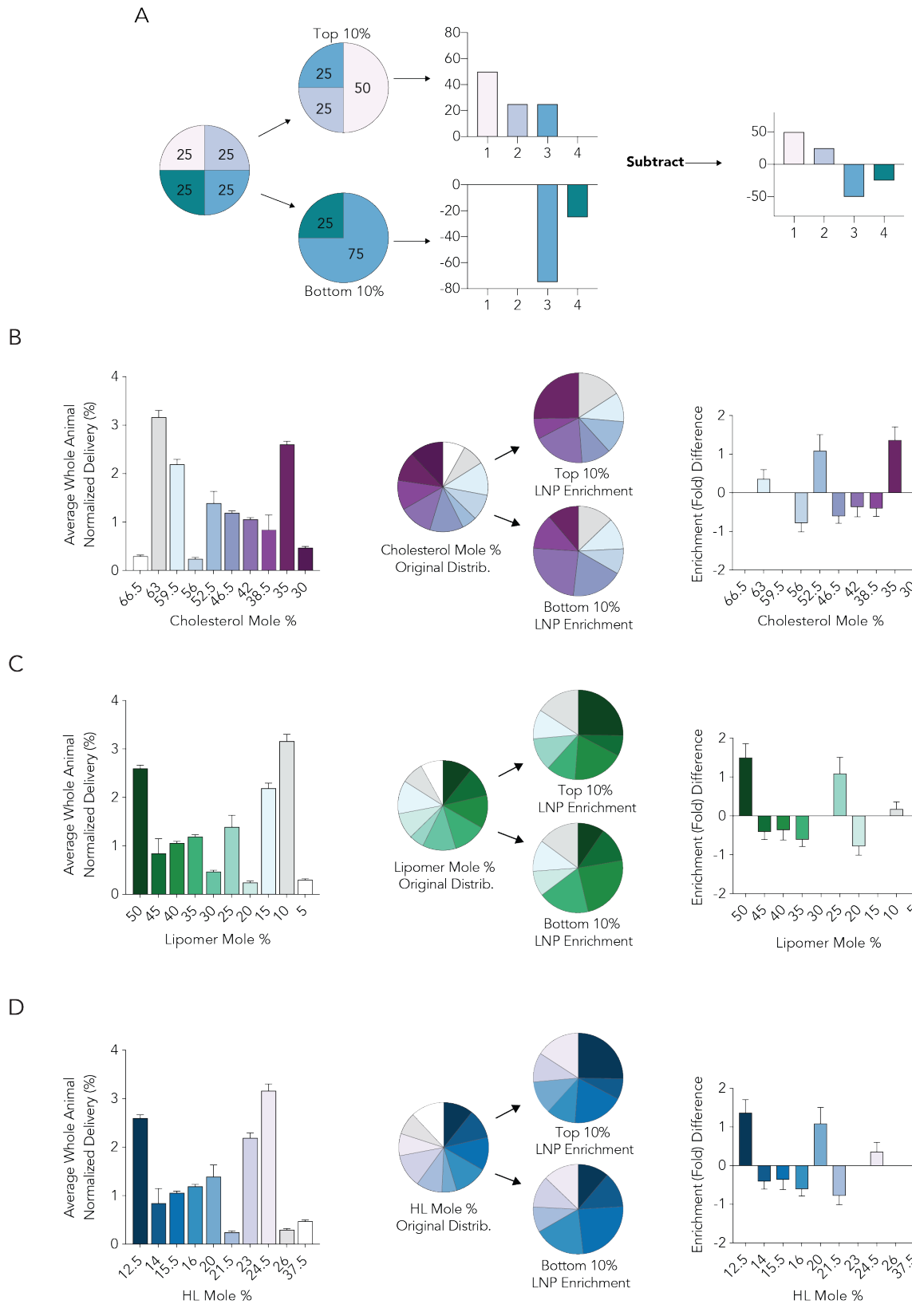


Figure 4.2 (A) Representative autocorrelative curve and (B) diameter distribution for included and excluded LNPs. (C) %tdTomato+ cells for cell types isolated from heart, lung, kidney, pancreas, and bone marrow.



Cell Type	Markers	Tissues
Hepatocyte (H)	CD45-CD31-	V
Kupffer Cell (K)	CD45+CD31-CD68+	V
Endothelial (E)	CD45-CD31+	V, H, L, K, P, M
Immune (I)	CD45+CD31-CD11b-	V, H, L, K, P
Other (O)	CD45-CD31-	H, L, K, P, M
HSCs	CD31-CD45+CD34+	M
Macrophage (M)	CD45+CD11b+	S, H, L, K, P, M
T Cell (T)	CD45+CD19-CD3+	S
B Cell (B)	CD45+CD3-CD19+	S
Immune (I)	CD45+CD19-CD3-CD11b-	S

Figure 4.3 Quantifying mRNA delivery mediated by 75 nanoparticles to 28 cell types *in vivo*. (A) After injecting mice with pooled nanoparticles carrying mRNA and DNA barcodes, we isolated 28 cell types from 7 tissues using FACS. (B) Functional Cre mRNA delivery, quantified as %tdTomato+ cells, in 8 cell types from the liver and spleen. (C) Hydrodynamic diameter of all administered LNPs; the diameter of the LNP pool is similar. (D) Normalized delivery of all LNPs, for all 8 cell types individually as well as (E) averaged for all 8 cell types. The naked barcode – negative control – falls out of every cell type. (F) Unbiased Euclidean clustering of 8 different cell types in 2 different tissues based on LNP delivery (%) clusters cell types according to tissue. The naked barcode is designated by an asterisk.


A
Top 10%
25
25
50
25
25
25
25
25
25
75
Bottom 10%
Subtract
B
Average Whole Animal Normalized Delivery (%)
Cholesterol Mole %
Cholesterol Mole % Original Distrib.
Top 10% LNP Enrichment
Bottom 10% LNP Enrichment
Enrichment (Fold) Difference
C
Lipomer Mole %
Lipomer Mole % Original Distrib.
D
HL Mole %
HL Mole % Original Distrib.

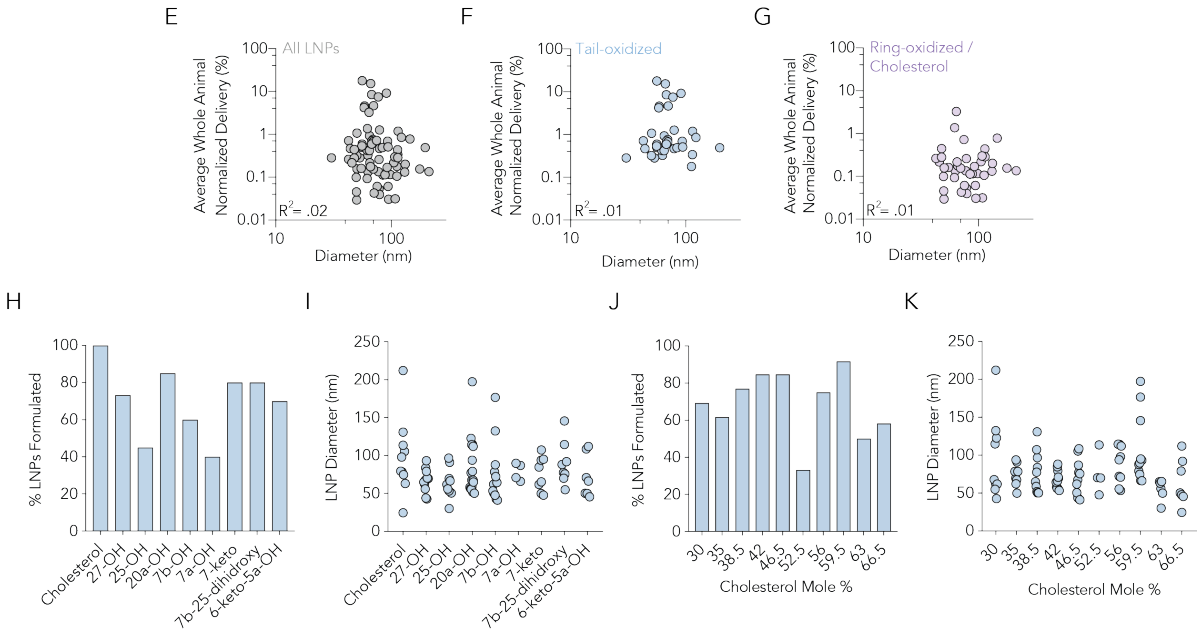
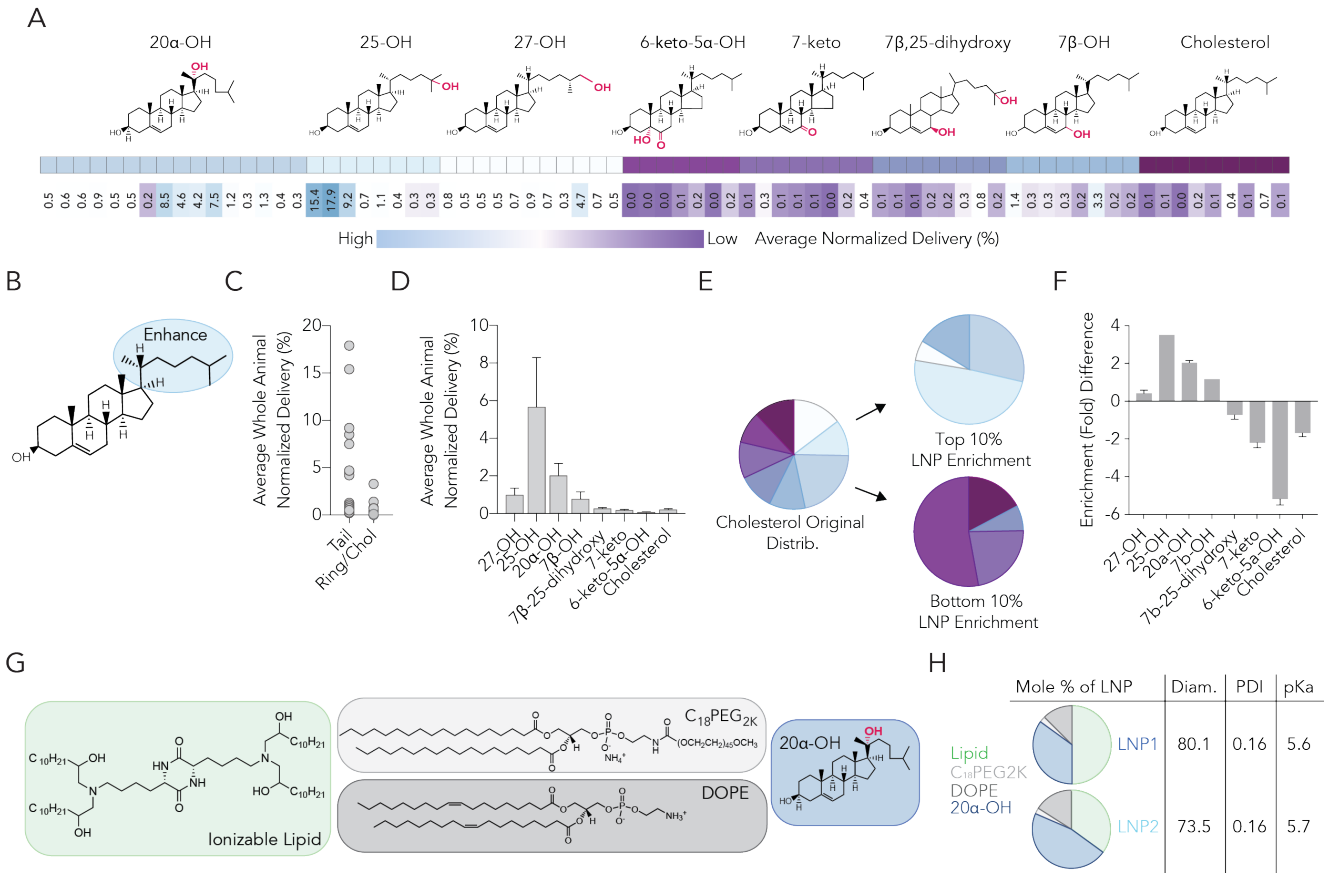


Figure 4.4 (A) Enrichment calculations detailed for obtaining enrichment (fold) difference values. Average whole animal normalized delivery (%), enrichment in the top and bottom 10% of LNPs, and fold difference in enrichment calculated for (B) cholesterol mole %, (C) ionizable lipid mole %, and (D) phospholipid mole %. LNP diameter plotted against average whole animal normalized delivery (%) for (E) all LNPs, (F) LNPs containing only tail-modified cholesterols, and (G) LNPs containing ring-oxidized cholesterols. LNPs formulated (%) shown based on (H) cholesterol variant used and (J) cholesterol mole %. LNP diameter shown based on (I) cholesterol variant used and (K) cholesterol mole %.


A
20α-OH
25-OH
27-OH
6-keto-5α-OH
7-keto
7β,25-dihydroxy
7β-OH
Cholesterol
High
Low
Average Normalized Delivery (%)
B
Enhance
C
Average Whole Animal Normalized Delivery (%)
Tail
Ring/Chol
D
Average Whole Animal Normalized Delivery (%)
27-OH
25-OH
20α-OH
7β-OH
7β-25-dihydroxy
7-keto
6-keto-5α-OH
Cholesterol
E
Cholesterol Original Distrib.
Top 10% LNP Enrichment
Bottom 10% LNP Enrichment
F
Enrichment (Fold) Difference
27-OH
25-OH
20α-OH
7b-OH
7b-25-dihydroxy
7-keto
6-keto-5α-OH
Cholesterol
G
Ionizable Lipid
C18PEG2K
DOPE
20α-OH
H
Lipid
C18PEG2K
DOPE
20α-OH
Mole % of LNP
Diam.
PDI
pKa
LNP1
80.1
0.16
5.6
LNP2
73.5
0.16
5.7

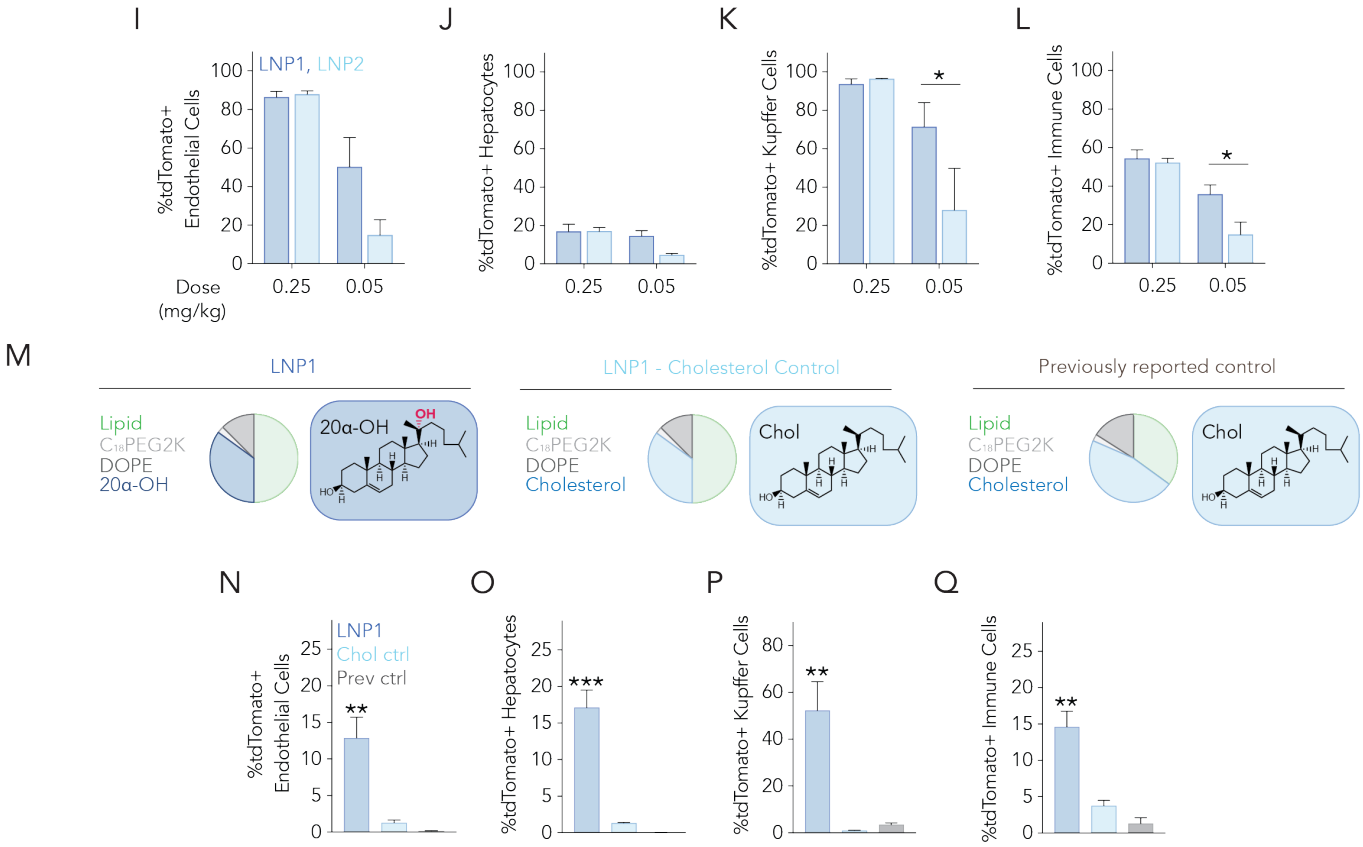
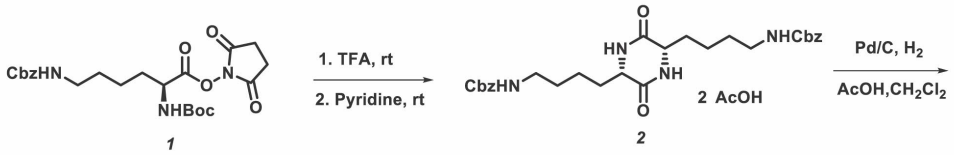
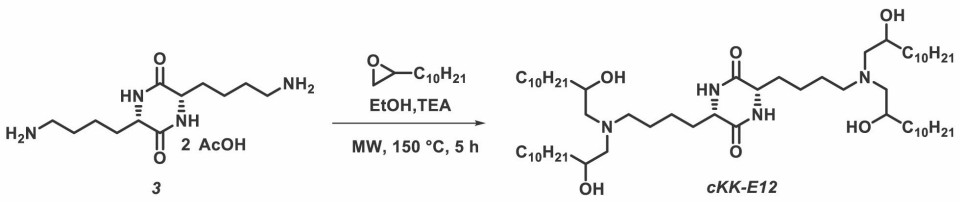


Figure 4.5 Modified cholesterols can alter nanoparticle delivery *in vivo.* (A) Normalized delivery averaged across all 8 liver and splenic cell types for each LNP, sorted by cholesterol type. (B,C) We observed that cholesterols modified on the cholesterol tail tended to enhance delivery more so than modifications made to the B ring. (D) Normalized delivery averaged across all 8 liver and splenic cell types suggested that 25-OH and 20α-OH cholesterols promote delivery. (E) Enrichment of cholesterol variants in the top 10% and bottom 10% of nanoparticles. (F) Fold enrichment in the top 10% of LNPs, calculated by subtracting enrichment in the bottom 10% of LNPs from enrichment in the top 10% of LNPs. (G) Based on these results we selected LNP1 and LNP2; they contain cKK-E12, DOPE, C18PEG2k, and 20α-OH. (H) LNP mole %, diameter (nm), polydispersity index (PDI), and pKa. (I-L) As predicted by the FIND screen, LNP1 and LNP2 delivered Cre mRNA to (I) endothelial cells, (J) hepatocytes, (K) Kupffer cells, and (L) immune cells, at 0.25 mg/kg. LNP1 delivered Cre mRNA more efficiently than LNP2 at 0.05 mg/kg. (M) The efficiency of LNP1 carrying Cre mRNA was compared to a control LNP1 that contained cholesterol and a previously reported control LNP, all 3 of which carried Cre mRNA. %tdTomato+ cells with each LNP are shown for (N) endothelial cells, (O) hepatocytes, (P) Kupffer cells, and (Q) immune cells. P-values are represented as p-value < 0.0332 (*), <0.0021 (**), <0.0002 (***).

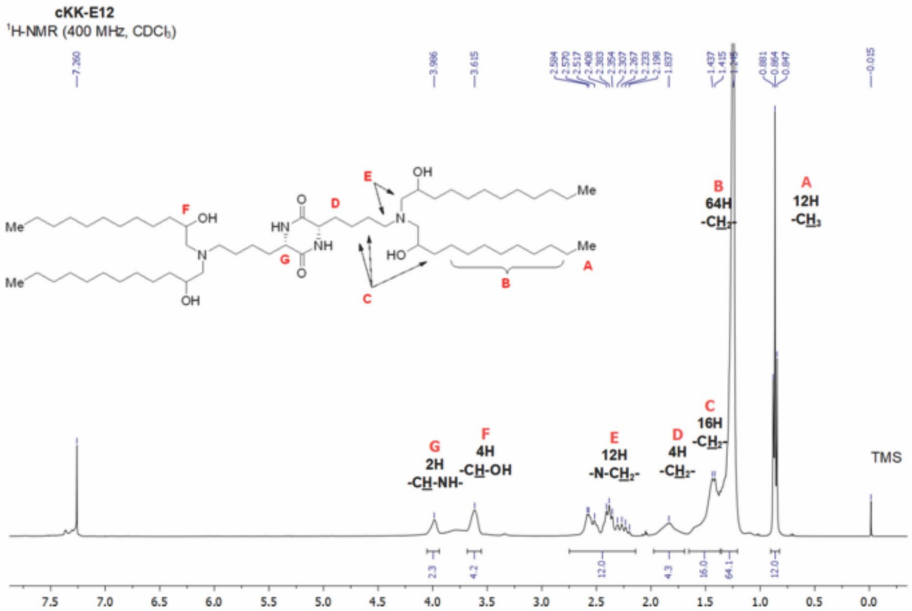
A



B



C



1H NMR (400 MHz, CDCl3) δ 0.87 (t, J = 6.8 Hz, 12 H, CH3), 1.21-1.35 (m, 64 H, CH2), 1.37-1.65 (m, 16 H, CH2), 1.71-1.95 (m, 4 H, CH2), 2.19-2.66 (br, 12 H, NCH2), 3.62 (m, 4 H, CHOH), 3.99 (m, 2 H, COCH).; HRMS (ESI, m/z) calcd for C60H121N4O6 [M + H]+: 993.9281, found 993.9269.

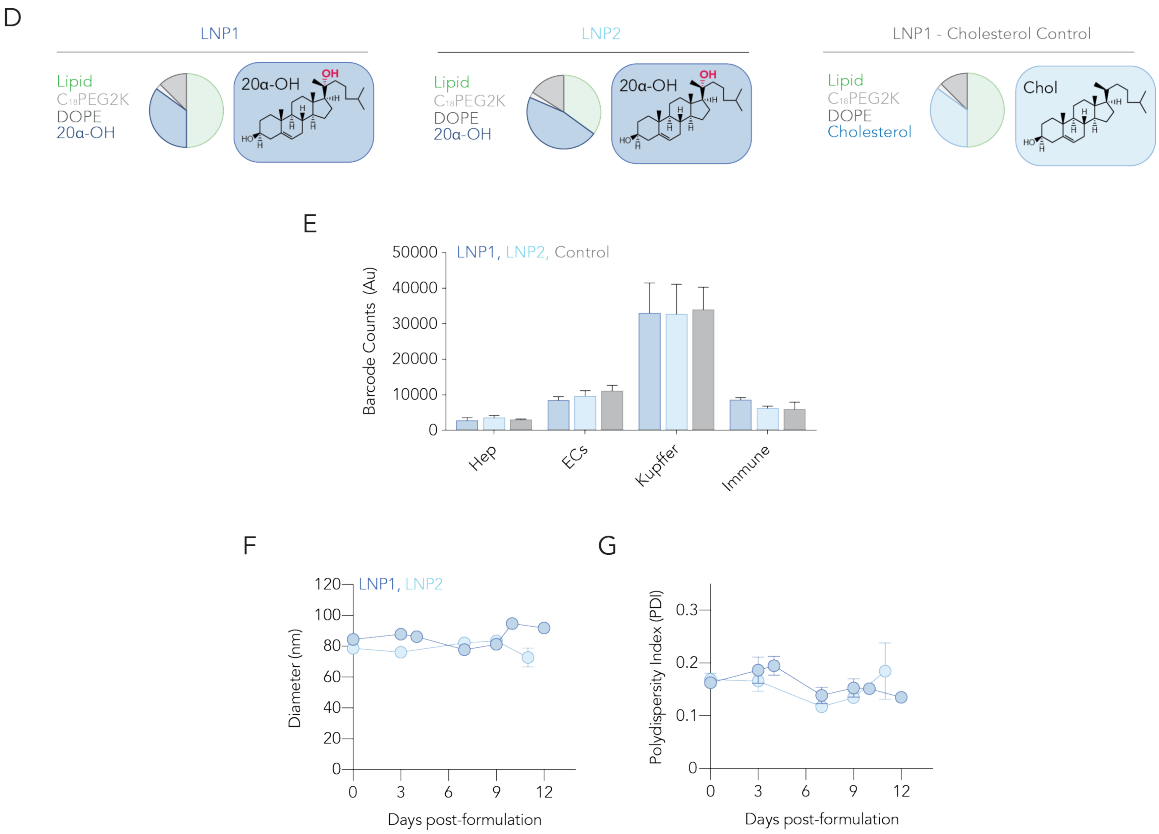


Figure 4.6 (A-B) CKK-E12 synthesis. (C) ^{1}H-NMR spectra confirming CKK-E12 structure. (D) LNP formulation ratios and components for LNP1, LNP2, and LNP1 – Cholesterol control. (E) LNP biodistribution, shown as barcode counts (Au), for LNP1, LNP2, and LNP1 – Cholesterol control. LNP (F) diameter (nm) and (G) polydispersity index (PDI) shown for LNP1 and LNP2 over time.

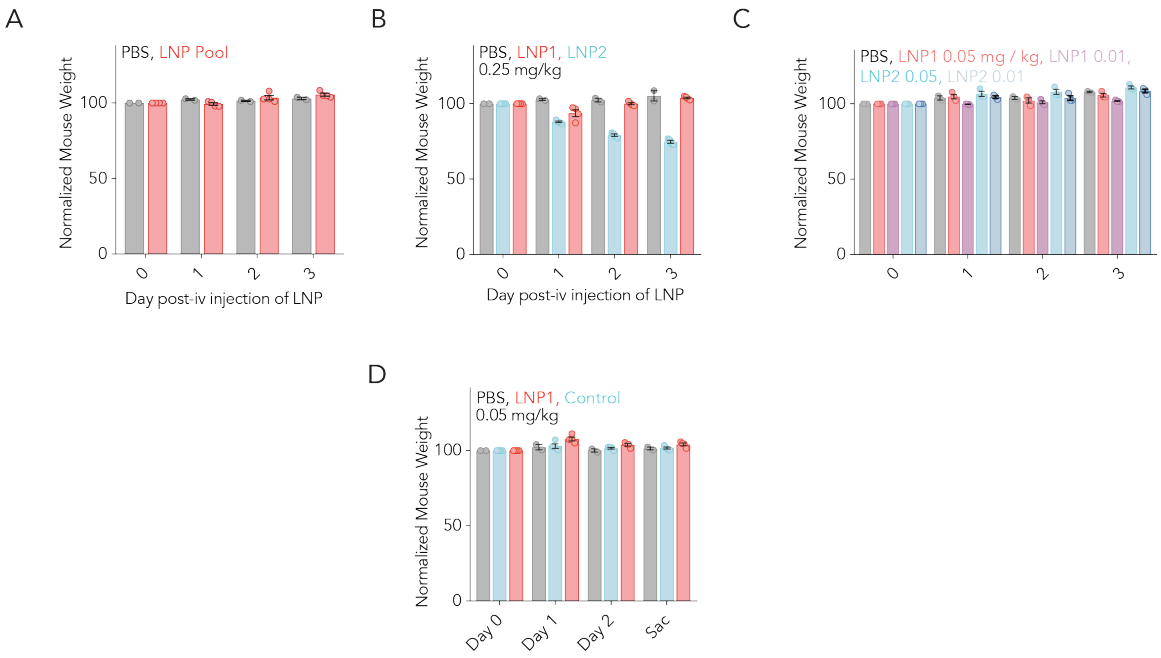
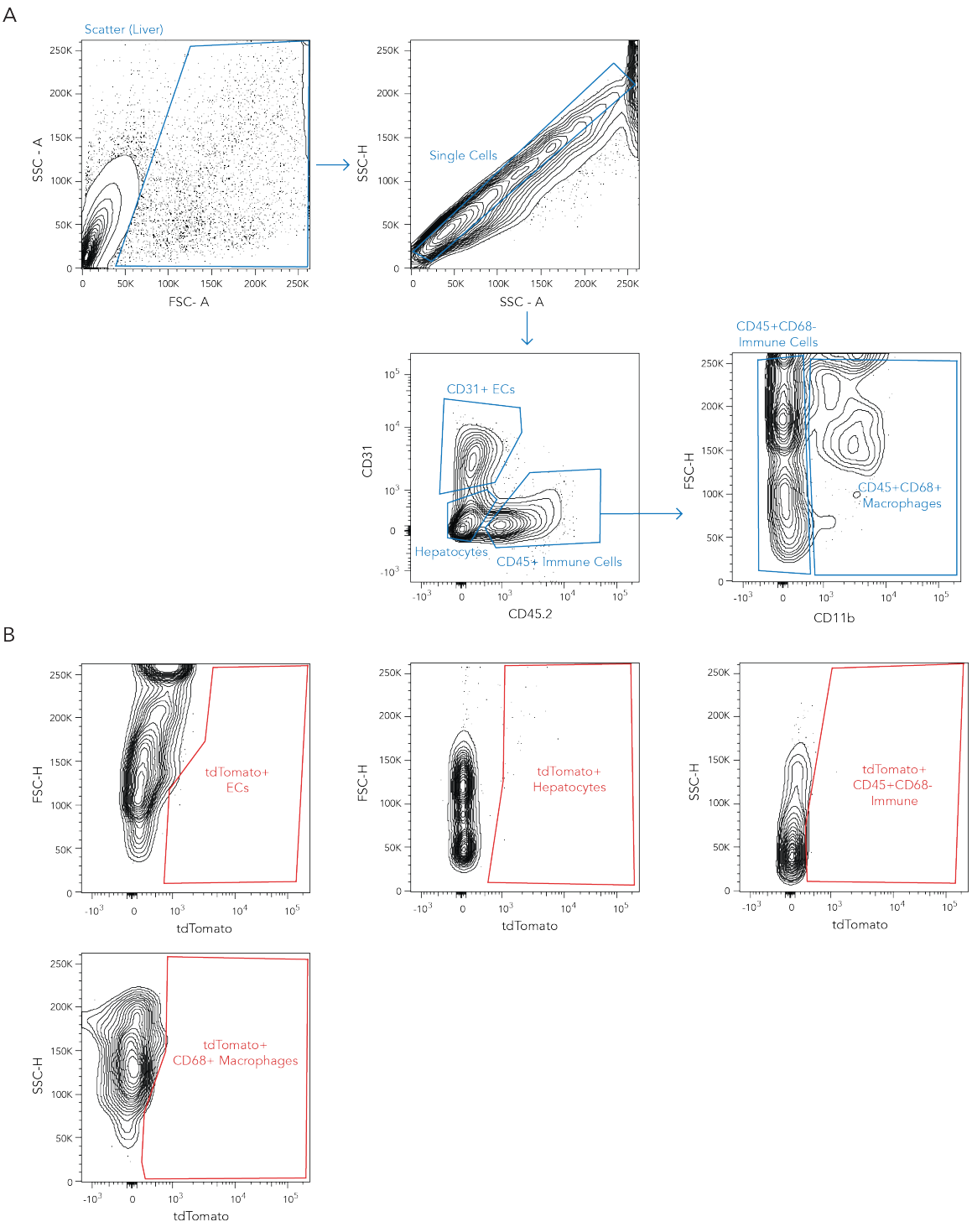


Figure 4.7 Mouse weights shown for mice injected with (A) 1 mg/kg LNP pool, (B) LNP1 and LNP2 containing Cre mRNA at a dose of 0.25 mg/kg, (C) LNP1 and LNP2 containing Cre mRNA at a dose of 0.05 mg/kg, (D) LNP1 and Control LNP containing Cre mRNA at a dose of 0.05 mg/kg.


A
Scatter (Liver)
SSC - A
FSC- A
Single Cells
SSC-H
SSC - A
CD31+ ECs
Hepatocytes
CD45+ Immune Cells
CD31
CD45.2
CD45+CD68-
Immune Cells
CD45+CD68+
Macrophages
FSC-H
CD11b
B
tdTomato+
ECs
FSC-H
tdTomato
tdTomato+
Hepatocytes
tdTomato+
CD45+CD68-
Immune
SSC-H
tdTomato+
CD68+ Macrophages

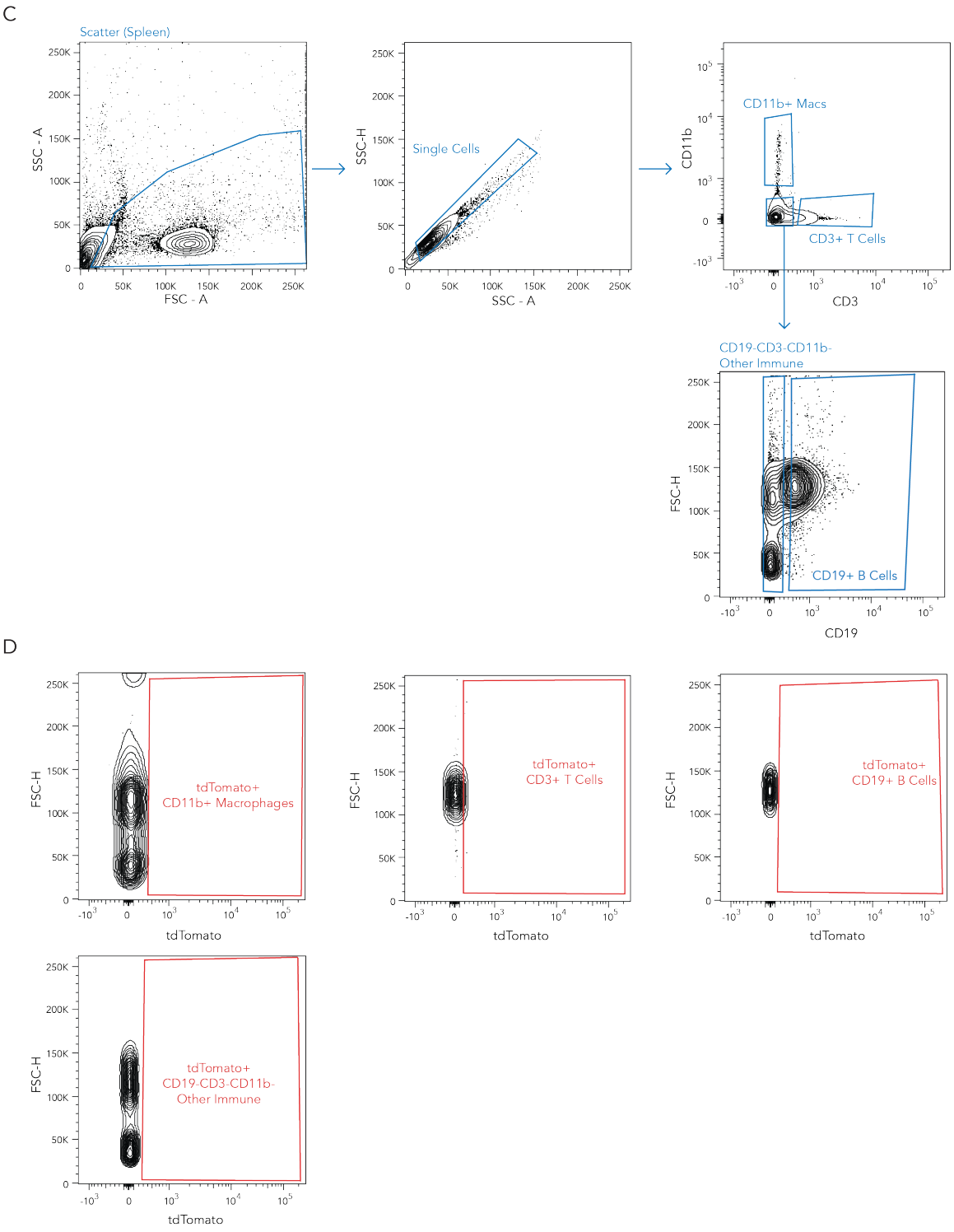


Figure 4.8 Representative gating strategies for FACS for cell types in the (A,B) liver, and (C,D) spleen.

CHAPTER 5. SMALL MOLECULES CAN BE USED TO MANIPULATE MRNA DELIVERY MEDIATED BY LIPID NANOPARTICLES

The work presented here is an excerpt from Paunovska* K, Da Silva Sanchez* AJ, Foster MT, Loughrey D, Blanchard EL, Islam FZ, Gan Z, Mantalaris A, Santangelo PJ, Dahlman JE (2020). "Increased PIP3 Activity Blocks Nanoparticle mRNA Delivery." *Under Review.*

5.1 Introduction

Nanoparticle-mediated mRNA delivery has the potential to express any gene, making this approach a promising way to treat disease. Nanoparticle delivery is a multi-step process governed by interactions between synthetic materials and the body. As a result, understanding the biological pathways that affect nanoparticles enable scientists to design effective drug delivery systems. For example, nanoparticles can interact with proteins expressed in the serum and on the cell surface. Serum proteins adsorbed onto nanoparticles can promote, or alternatively, block, interactions between nanoparticles and cell surface receptors[5, 106, 111]. In one case, the expression of apolipoprotein E was necessary for ionizable lipid nanoparticles (LNPs) to deliver siRNA to hepatocytes *in vivo*; this was mediated by the low-density lipoprotein receptor, which is expressed on the cell surface[84]. Understanding this biological mechanism of action helped lead to the development of an FDA approved LNP[1]. It is similarly accepted that once a nanoparticle reaches a cell, the route by which it accesses the cytoplasm impacts its activity *in vitro* and *in vivo*. For example, studies manipulating endocytosis with small molecules[125, 210, 211], siRNA[88, 124], CRISPR-Cas9[125], or knockout mice[78, 207] have revealed that endocytosis, and subsequent endosomal escape, is carefully regulated. In one representative example, researchers found

that siRNA-containing LNPs engage both clathrin-mediated endocytosis and macropinocytosis to deliver their cargo to the cytoplasm. Furthermore, the bioavailability of the siRNA was limited by the escape of the drug from hybrid endosomes, which exhibited early and late endosomal characteristics[88]. The researchers also characterized the escape window for siRNAs and estimate that only 1-2% of administered siRNAs make it to the cytosol of a target cell. In another example, researchers found that less than 1% of mRNAs delivered by LNPs made it to the cytosol of a cell and that mRNA endosomal escape was highly dependent on the type of LNP used to deliver the mRNA[212].

In this study, we sought to answer whether cell metabolism alters nanoparticle delivery. We focused on this question for four reasons. First, it has immediate clinical relevance; nanomedicines are currently administered to patients with disorders characterized by strong metabolic phenotypes, including cancer[213]. Second, literature suggests metabolism could affect some steps of the drug delivery process. Specifically, to achieve cytoplasmic mRNA delivery, a nanoparticle first interacts with serum proteins and the cell surface. Metabolism influences how cells interact with lipoproteins[214], which are naturally occurring nanomaterials that can have a similar chemical structure to LNPs[78, 197]. After a nanoparticle reaches the cell it can enter, and with less frequency exit, an endosome; metabolism affects endocytosis pathways important for nanomedicine[215]. Third, mRNA that enters the cytoplasm must be translated into protein; cell metabolism affects mRNA translation and degradation[205]. Finally, recent evidence implicates mTOR, a canonical metabolic pathway, as a mediator and player in both antisense oligonucleotide (ASO) activity[216] and nanoparticle delivery via to-be-determined mechanisms[125]. In the first example, authors found that small-molecule inhibition of mTOR increased ASO activity

in vivo. In the second example, authors inactivated genes related to endocytosis using CRISPR. They found knocking out Rab7a, which is necessary for late endosomal trafficking, reduced delivery, whereas knocking out Rab4a and Rab5a, which is necessary for endosomal recycling and early endosomal trafficking respectively, did not. The authors reasoned that halting endosomal maturation by deleting Rab7a blocked the metabolic gene mTORC1, which is expressed on the lysosomal surface, from initiating mRNA translation. To verify this, the authors activated mTORC1 and observed increased protein expression.

These lines of evidence led us to reason that we could manipulate metabolism with small molecules to improve LNP delivery. Specifically, we hypothesized that it was possible to metabolically reprogram cells, so more mRNA was translated once it reached the cytoplasm. To achieve this goal, we chose the bioactive lipid PIP3. PIP3, a membrane phospholipid created by the phosphorylation of PIP2, mediated by PI3K, signals via interactions with proteins containing pleckstrin homology domains at the plasma membrane[217]. Specifically, PIP3 binds to PDK1, initiating the protein kinase to phosphorylate and activate Akt. Phosphorylation of Akt leads to inhibition of the TSC complex and downstream activation of the GTPase Rheb, which directly stimulates mTORC1 kinase activity[218, 219]. Increased PIP3 concentrations upregulate clathrin- and dynamin-mediated endocytosis of EGFR[220] and sort endosomal cargos in epithelial cells[221], suggesting PIP3 could increase endocytosis. Increased PIP3 activity also promotes cell growth via several mechanisms, including increased translation[222]. We therefore reasoned treating cells with PIP3 and mRNA containing LNPs would transiently upregulate translation, thereby increasing the 'effective potency' of the LNPs. However, the data we generated did not support this hypothesis. Instead, we found the opposite: PIP3 potently

blocked mRNA delivery mediated by three clinically relevant (FDA approved or licensed for clinical translation) LNPs (**Fig. 5.1A**).

By performing RNA-seq and metabolomics analyses of PIP3-treated cells, we identified pathways that, until now, have not been related to LNP delivery. Our analysis suggests that activating metabolism increases endogenous transcription, which we posit may reduce the effective concentration of exogenous mRNA delivered by the LNPs. These data highlight the importance of understanding the metabolic profile of target and off-target cells when designing nanomedicines.

5.2 Results

5.2.1 PIP3 treatment reduces mRNA translation in multiple cell types

To study how PIP3 affected LNP-mediated mRNA delivery, we used microfluidics[142] to formulate a clinically relevant LNP (herein named LNP1) so it carried chemically modified mRNA encoding GFP (**Fig. 5.1B and Fig. 5.2A**). LNP1 has delivered mRNA to lung endothelial cells *in vivo*[81], and is composed of the lipid cKK-E12, which delivered RNA in NHPs [71] and is being considered for clinical use. We used two cell lines: HEKs and iMAECs[154]. Cells in 24 well plates were concurrently treated with GFP mRNA and PIP3. Six or 24 hours later, we quantified mRNA delivery by measuring the percentage of GFP^+ cells using flow cytometry, using untreated cells as flow cytometry gating controls (**Fig. 5.2, B to J**).

We observed a PIP3 dose-dependent decrease in GFP fluorescence (**Fig. 5.1, C and D**); the statistical significance of these results is listed (**Fig. 5.2, K to P**). The PIP3 induced

effect was significant: at six hours, GFP expression was reduced from 85% (0 μM PIP3) to 0% (10 μM PIP3) in iMAECs and from 85% (0 μM PIP3) to 20% (10 μM PIP3) in HEKs (**Fig. 5.1C**). We repeated the experiment using L2K and observed a similar PIP3 dose-dependent reduction in GFP expression (**Fig. 5.1, E and F**). We repeated the experiment with a second clinically relevant LNP[9] (termed LNP2) that has a different chemical structure and *in vivo* tropism than LNP1 (**Fig. 5.1B**). Specifically, LNP2 delivers mRNA to splenic endothelial cells in vivo[9] and consists of the lipid 7C1[76] (**Fig. 5.2A**), which delivered RNA in NHPs[223] and is licensed for clinical development. We observed PIP3 dose-dependent inhibition of GFP fluorescence after LNP2 transfection (**Fig. 5.2 Q and R**). We then treated cells with 0, 10, or 20 μM PIP3 and concurrently administered increasing doses of GFP mRNA. GFP expression decreased with PIP3 dose and increased with GFP mRNA dose, as expected (**Fig. 5.1G and Fig. 5.2, S and T**). To understand the kinetics of this effect, we then varied the timing between PIP3 and LNP treatment. PIP3-treated cells expressed less GFP when PIP3 was administered four hours before the LNP and expressed normal levels of GFP when PIP3 was administered three hours after the LNPs (**Fig. 5.1H**). Taken together, these results led us to conclude that treating cells with PIP3 reduced the amount of mRNA translated into protein after LNP delivery *in vitro*.

5.2.2 Reduced protein expression is not caused by overt toxicity

We reasoned this surprising reduction in mRNA delivery could be due to overt cellular toxicity. We performed four assays comparing untreated cells to PIP3-treated cells: (i) MTT (six or 24 hours after PIP3 administration), (ii) LDH (24 hours after PIP3 administration), (iii) Nf-κB activation (eight or 12 hours after PIP3 administration), and (iv) cell morphology (six or 24 hours after PIP3 administration). We did not find any

evidence of toxicity in any of these assays. Specifically, we found no significant change in MTT readouts when administering PIP3 at a 10 μM or 20 μM dose (**Fig. 5.2U**) and no significant decrease in LDH readouts after PIP3 was administered to cells at 10 μM or 20 μM; the positive control LPS did reduce cell viability (**Fig. 5.2V**). Similarly, PIP3 did not increase Nf-κB activation in Nf-κB reporter cells (**Fig. 5.2W**). We then analyzed the effect of PIP3 on iMAEC cell morphology and found no difference between untreated cells and cells treated with PIP3 (**Fig. 5.1I**). Finally, we performed a literature search and found similar results; 25 μM PIP3 doses did not cause toxicity[224]. Thus, these data did not support our hypothesis that reduced protein expression was due to overt cytotoxicity.

5.2.3 Reduced protein production is not explained by cell uptake or endosomal escape

Our second hypothesis was that PIP3 blocked mRNA delivery by changing (i) particle stability, (ii) cell uptake, (iii) endosomal escape, or a combination thereof (**Fig. 5.3A**). To study (i) particle stability, we quantified the hydrodynamic diameter of LNP1 in PBS supplemented with mouse serum or iMAEC media using DLS. Adding PIP3 did not change LNP diameter (**Fig. 5.4A**). To study (ii) cell uptake, we formulated LNP1 carrying GFP mRNA so it contained the phospholipid DOTAP labeled with a fluorescent probe (NBD-DOTAP) (λ_{abs} 460nm; λ_{em} 535nm). Separately, we formulated LNP2 carrying mRNA so it contained Alexa Fluor 647-tagged DOPE (AF647-DOPE) (λ_{ex} 643 nm; λ_{em} 662.0 nm). We did not change formulation ratios; we replaced normal DOTAP and DOPE with fluorescent DOTAP and DOPE. We administered the LNPs to HEKs and iMAECs at a dose of 50 ng / well for LNP1 and 150 ng / well for LNP2, and either did (or did not) add PIP3. We then isolated cells and quantified cellular fluorescence. As a negative control, we used untreated cells. PIP3 did not impact LNP1 uptake in iMAECs (**Fig. 5.3B**)

but did impact LNP1 uptake in HEKs at early timepoints; there was a 6% increase in LNP uptake in PIP3 treated HEKs at both 0.5 and 1 hour (**Fig. 5.3C**). PIP3 increased LNP2 uptake in iMAECs by 52% at two hours, 56% at six hours, and 29% at 24 hours (**Fig. 5.3D**) and decreased uptake in HEKs by 72% at two hours (**Fig. 5.3E**). These data suggested that PIP3 may alter nanoparticle uptake; however, the effects were not conserved across LNPs or cell type. More importantly, changes to LNP uptake were far less substantial than the (almost total) reduction in GFP MFI. We therefore concluded these effects on LNP uptake were not sufficient to explain the consistent and robust decrease in protein production.

We then investigated whether PIP3 reduced canonical endocytosis using two pathways: clathrin- and caveolin-mediated endocytosis. We quantified the uptake of transferrin and cholera toxin B, which are ligands for clathrin- and caveolin-mediated endocytosis, respectively[225, 226]. We treated cells with PIP3 and either transferrin or cholera toxin B that were fluorescently tagged. Compared to cells that were not pre-treated with PIP3, transferrin and cholera toxin B uptake was unchanged in iMAECs and HEKs (**Fig. 5.4 B and C**). These data led us to conclude that PIP3 did not substantially impact canonical clathrin- or caveolin-mediated endocytosis.

We then studied (iii) whether PIP3 reduced LNP endosomal escape. We focused on LNP2, since its uptake was impacted more by PIP3 than LNP1. We treated cells with LNP2 formulated with fluorescently labeled DOPE. Thirty minutes and six hours after delivery, cells were fixed and stained with 1º antibodies against Rab7, EEA1, and CD63 which mark different stages of endosomal maturation[227]. We used a 2º antibody to stain all endosomes (**Fig. 5.3F**) and quantified both the line profiles (**Fig. 5.3, G and H**) as well as the M1 and M2 coefficients of LNPs with endosomes for 30 cells per treatment condition

as previously described[227] (**Fig. 5.3, I and J**). The M1 coefficient quantifies the fraction of LNP signal that is colocalized with endosomal signal, while the M2 coefficient quantifies the fraction of endosomal signal that is colocalized with LNP signal. Thus, a lower M1 coefficient indicates greater endosomal escape of LNP in the cell at that timepoint. Cells treated with both LNP and PIP3 had 49% and 27% lower M1 coefficients than cells treated with LNP only at 30 minutes and 6 hours, respectively (**Fig. 5.3I**). Cells treated with both LNP and PIP3 also had a 7.6-fold reduction in M2 coefficient at 6 hours compared to those treated with LNP only (**Fig. 5.3J**). The decrease in the M1 coefficients after PIP3 treatment indicates that, at the tested timepoints, PIP3 increased the endosomal escape of LNPs. However, these data do not explain the absence of GFP expression in PIP3-treated cells. Although we cannot discard the potential interplay of endosomal trafficking with the metabolic alterations caused by PIP3 treatment, these results suggest that endosomal escape does not explain the observed effect.

5.2.4 RNA-seq and metabolomics suggest novel pathways could influence LNP delivery

The data above led us to conclude that our second hypothesis did not explain the robust reduction in mRNA delivery. We therefore tested a third hypothesis: that cell metabolism altered LNP delivery (**Fig. 5.3A**). To test this, we performed RNA-seq analysis of gene expression in iMAECs that were not treated with PIP3, as well as cells treated with 10 μM PIP3 for six or 24 hours. We generated a total of 530 million clean reads from 12 cDNA libraries using Illumina Nextseq, an average of 114.4x fold coverage of the coding region of the genome. We then performed quality controls to ensure it was appropriate to analyze these data. First, we found more than 90% of the clean reads had quality scores

greater than Q30. Second, more than 95% of the total reads were uniquely mapped to the reference genome.

We generated two volcano plots (six and 24 hours) to analyze the adjusted p-value as well as the fold change of each transcript, relative to the untreated cells. Specifically, we set a $p<0.05$ and fold change >1.5 as thresholds for differentially expressed genes. These data were interesting for several reasons. First, the volcano plots showed that the metabolic state of the cell was, generally, more perturbed six hours after PIP3 treatment (**Fig. 5.5A**), relative to 24 hours (**Fig. 5.5B**). When we analyzed the differentially expressed genes at the six-hour timepoint, we found seven upregulated and 11 downregulated genes following PIP3 exposure. At the 24 timepoint, zero were upregulated and six were downregulated. Second, when we analyzed these genes using traditional Pubmed searches (**Fig. 5.5, C and D, and Fig. 5.6, A to D**), we found that, as expected, many of these genes had annotated roles in cell metabolism. At six and 24 hours, 56% and 50%, respectively, of the differentially regulated genes had clear connections to mTOR, indicating that the observed transcriptional response is in line with canonical PI3K-Akt-mTOR signaling. Interestingly, almost none of these genes had currently annotated roles in endocytosis, transcription, translation, or LNP delivery. We believe these genes are interesting and novel candidates for future LNP delivery studies.

Given that most of our identified genes were not related to pathways that traditionally alter nanoparticle delivery (e.g., endocytosis), we performed a less biased analysis of the genes using the KEGG analytical tool. In the six-hour dataset, we found that, as expected, the PI3K-Akt pathway was affected by PIP3 treatment. Interestingly, signaling pathways related to immune response (IL-17), proliferation/migration (MAPK),

endocrine signaling, and differentiation regulated by immune response were also altered. We were unable to find previous studies relating these genes to LNP delivery *in vivo*. However, it was notable that scRNA-seq recently revealed that IL-17 regulates the immune response to tissue engineering constructs[228]. Our data support this IL-17 result and suggest that this pathway may regulate the biological response to synthetic materials. It was also recently shown that IL-17 production could be positively regulated through induction of RAR-related orphan receptor gamma (RORγ) and aryl hydrocarbon receptor (AHR), both of which are mediated by the mTOR pathway[229, 230]. Studies have also identified mTOR signaling as an important component of Th17 (IL-17 producing T helper) cell development and differentiation into IFN-γ producing cells[231, 232] as well as a regulator of the innate inflammatory immune response through TSC. mTOR signaling, and specifically mTORC1, is a mediator of IFN-γ production and Th17 response at sites of inflammation – indicating that mTORC1 is important in regulating metabolic heterogeneity within the Th17 population and modulating the cellular response to chronic inflammation by initiating cell differentiation[232]. Given current literature, it comes as no surprise that the IL-17 pathway is impacted by changes in mTOR signaling. Notably, twenty-four hours after PIP3 exposure, PI3K-Akt, MAPK, and IL-17 pathways were again differentially expressed.

We then performed a second, complementary unbiased analysis of our RNA-seq dataset. Specifically, we used GO enrichment (specifically, the Enrichr package) to identify cellular processes regulated by the dysregulated genes we identified. We focused on GO pathways related to cellular components, biological processes and molecular functions. In the six hour dataset, the GO categories for transcriptional repressor activity, RNA polymerase II activating transcription factor binding ($p<0.000051$, GO:0098811), the

regulation of pri-miRNA transcription from RNA polymerase II promoter ($p< 0.00013$, GO:1902895), and RNA polymerase II transcription factor complex formation ($p<0.008$, GO:0090575) were strongly regulated (**Fig. 5.6E**). After twenty-four hours, negative regulation of cell cycle ($p<0.0002$, GO:0045786), and RNA polymerase II core promoter proximal region sequence-specific DNA binding ($p<0.0025$, GO:0000978) were regulated (**Fig. 5.6F**). Taken together, our three analyses (Pubmed, KEGG, GO) suggested that a combination of cell metabolism transcription, inflammation, and cell cycle may work in concert to regulate the efficacy of mRNA delivered by LNPs.

Interestingly, endocytosis and endosomal escape-based pathways and processes were not implicated at the six or 24-hour RNA-seq timepoints with the three analyses that we preformed. Of the 18 differentially regulated genes at six hours, only one –Thbs2 – has been characterized as playing a role in endocytosis (**Fig. 5.5C**). At 24-hours, no genes are implicated in endocytosis or endosomal escape (**Fig. 5.5D**). Thbs2 has been identified as a ligand for LRP-1[233]; LRP-1 is a multifunctional endocytic receptor that is part of the LDLR family. We and others have previously shown that knocking out and / or silencing LDLR or LRP-1 impacts RNA delivery[78, 88]. However, LRP-1, LDLR, or any related proteins were not differentially regulated within our RNA-seq dataset; this data supports the conclusions we made in the previous section and suggests that alteration of endocytic pathways may not be the cause of diminished GFP protein expression after PIP3 treatment.

Based on these RNA analyses as well as the canonical role PIP3 has in regulating cell metabolism, we performed an unbiased metabolomic analysis of cells treated with 10 μM PIP3. In this case, we compared five groups of iMAECs: zero hours without PIP3, six hours with and without PIP3, and 24 hours with and without PIP3. We used five samples

per group: four for analysis and one for DNA quantification (as a control). We then performed a series of statistical analyses and dimensionality reduction techniques to ensure the data were robust; specifically, we performed PCA on cells treated with or without PIP3 for 24 hours. Control cells that were not treated with PIP3 tended to have less variability across both components while cells treated with PIP3 had little variability across PC1, which makes up for most of the variance, but displayed more variability across PC2 (**Fig. 5.7A**). Alongside PCA, we used hierarchical clustering, which confirmed that cells separated into two groups associated with (or not) treatment with PIP3. Specifically, the PIP3 treated cells formed a distinct cluster while the untreated cells clustered separately (**Fig. 5.7B**). PCA and hierarchical clustering was also performed for metabolites analyzed at six hours (**Fig. 5.8**). We performed several additional controls, which are detailed in Materials and Methods.

This analysis identified metabolites that were significantly changed in cells treated with PIP3, relative to time-matched cells that were not (**Fig. 5.7C**). We observed metabolites that were not significantly changed as well as metabolites that were significantly changed at six and 24 hours. Within the group of metabolites upregulated at both six and 24 hours, we found pathways responsible for (i) anabolic metabolism with increased glycolysis, (ii) pentose phosphate regulation, (iii) methionine-enabled methyl transferase epigenetic regulation, (iv) phospholipid/glycerolipid synthesis, and (v) nucleotide synthesis. We were especially interested in (iv) and (v), which could potentially alter the stability of (iv) LNP components or (v) mRNA, respectively. We similarly noted that hexadecanoic acid and octadecanoic acid, which contain alkyl tails similar to those in the LNPs, were upregulated at both timepoints. Within the metabolites that were

upregulated six hours following PIP3 treatment, we noted those related to cholesterol production, since these (and most reported) LNPs contain cholesterol[72, 78]. Finally, we found that several amino acids (specifically, isoleucine, alanine, ß-alanine, homoserine, and ornithine) were significantly decreased at the 24-hour timepoint in PIP3-treated cells. As detailed in the discussion, these data supported our RNA-seq analyses and led us to conclude that PIP3 metabolic reprogramming of cells blocked mRNA delivered by LNPs.

5.2.5 In vivo delivery is blocked by PIP3

All the studies described above were performed *in vitro*, which can be a poor predictor of *in vivo* delivery[8]. We therefore investigated whether PIP3 blocked LNP delivery *in vivo*. We performed these studies in Ai14 mice, which have a LSL-tdTomato construct under the control of a constitutive promoter. When Cre protein is produced, it translocates into the nucleus and excises the 'Stop' from genomic DNA; the cells then become tdTomato$^+$ (**Fig. 5.9A**). The percentage of tdTomato$^+$ cells after Cre mRNA delivery is a validated readout of nanoparticle delivery[9, 72, 81]. To quantify how PIP3 treatment altered mRNA delivery *in vivo,* we intravenously injected mice with PBS or 10 mg/kg PIP3. Immediately afterwards, we intravenously injected mice with 1 mg/kg Cre mRNA formulated inside LNP1 or LNP2. Three days later we quantified the percentage of tdTomato$^+$ cells using flow cytometry. As a control, we gated on an Ai14 mouse treated with PBS. As previously reported, LNP1 delivered Cre mRNA to lung endothelial cells (**Fig. 5.9B**) and LNP2 delivered Cre mRNA to splenic endothelial cells (**Fig. 5.9C**). LNP1 and LNP2 did not efficiently deliver mRNA to cells in the liver (**Fig. 5.10, H to K, and Fig. 5.10 S to V**). Thus, to evaluate whether PIP3 blocked mRNA translation in the liver, which is an important clinical target for RNA therapies, we formulated LNP3 (**Fig. 5.9D**),

which, as reported, delivered RNA to hepatocytes[71] (**Fig. 5.9E**). mRNA delivery mediated by LNP1, LNP2, and LNP3 was reduced robustly, between six- and 14-fold in mice treated with PIP3 (**Fig. 5.9**). Specifically, we observed a 10.4-fold reduction in tdTomato+ cells when administering LNP1 and PIP3 concurrently (**Fig. 5.9B**), a 6.2-fold reduction when administering LNP2 and PIP3 concurrently (**Fig. 5.9C**), and a 13.6-fold reduction when administering LNP3 and PIP3 concurrently (**Fig. 5.9E**). Notably, in every cell type with high levels of delivery, we observed decreases in Cre mRNA delivery, most of which were statistically significant (**Fig. 5.10**). None of the mice treated with PIP3 exhibited weight loss (**Fig. 5.12**) or changes in behavior suggesting toxicity.

5.2.6 PIP3 co-formulated within an LNP does not consistently alter protein expression

Finally, we investigated whether PIP3 blocked LNP delivery when it was administered 'intracellularly', i.e., co-formulated within an LNP (**Fig. 5.11A**). We first characterized LNP size after formulating the LNPs with PIP3; we found that LNP size did not change when PIP3 was incorporated and that PIP3-containing LNPs were stable (**Fig. 5.11B**). We repeated the experiments above, comparing LNPs formulated with and without PIP3. We found that including PIP3 did not consistently increase or decrease protein expression *in vitro* (**Fig. 5.11C**) or *in vivo* (**Fig. 5.11D-O**). We then analyzed the cellular uptake of LNPs formulated with and without PIP3 and found it to be similar (**Fig. 5.11P-S**). These data suggest that PIP3-mediated reductions in LNP efficacy occur when PIP3 is administered directly to cells, but not when administered within an LNP. The data support previous evidence suggesting that the subcellular localization of a lipid can alter its biological effect[234]. However, many more experiments are required to understand whether this is consistent and robust.

5.3 Discussion

Taken together, these data demonstrate that the bioactive lipid PIP3 reduced the efficacy with which chemically distinct LNPs delivered mRNA *in vitro* and *in vivo.* We concluded that this effect was not driven by overt toxicity or overt differences in cell uptake or endosomal escape. Instead, we found that this effect was largely driven by changes to metabolic, transcriptional, and inflammatory cell signaling.

By performing RNA-seq and metabolic analyses, we were able to generate two non-mutually exclusive hypotheses. In our first hypothesis, we propose that exogenous mRNA delivered into the cytoplasm by LNPs competes with endogenous mRNA for cellular resources that lead to protein production and protein stability. By increasing the basal metabolic activity of a cell at the time LNPs were delivered, PIP3 limited the resources available to process exogenously administered mRNA, effectively 'drowning out' LNP delivery. This hypothesis is in its early stages; however, it is supported by several lines of evidence. First, we noted that, surprisingly, several genes identified by RNA-seq regulated transcription, which could change the amount of endogenous mRNA competing with exogenous LNP-delivered mRNA. This was also supported by the metabolic data reporting increased lipid and nucleotide syntheses. Second, 24 hours after PIP3 treatment, the number of available amino acids decreased. Third, while we observed substantial reductions in protein production, we did not observe similarly substantial reductions in endocytosis or endosomal escape in the PIP3-treated group.

Our second hypothesis is that high metabolic activity in the cell leads to cell stress, and eventually catabolism; this catabolism degrades protein translated from mRNA. Six

hours after PIP3 treatment, we observed increased aerobic glycolysis, characterized by upregulated amino acid synthesis, protein synthesis, and cholesterol production. By 24 hours, we observed an ROS-driven catabolic phenotype, characterized by no increases in protein synthesis or cholesterol production, and by significant decreases in amino acids.

More specifically, we observed increased flux throughout glycolysis in the cytosol and TCA cycle within the mitochondria. Increases in glycolysis increase mitochondrial-mediated metabolic pathways, including the TCA cycle and OxPhos. We observed two lines of evidence that suggest TCA and OxPhos upregulation. First, both TCA and OxPhos utilize hexa- and octo-decanoic acids[235], two lipids that were upregulated at 6 and 24 hours. Second, upregulated OxPhos is also a metabolic phenotype for pro-survival pathway activation, which was supported by the transcriptomic changes in MAPK and IL-17 we observed[236]. These PIP3-driven increases in mitochondrial expansion and OxPhos elevated intracellular ROS, which suggests cells shifted to catabolism by 24 hours. The increase in ROS was supported by our transcriptomic data showing changes in Fos, which senses ROS stress[237]. Furthermore, the reduced aspartate and TCA cycle compounds (e.g. citrate) serve as another indication of the catabolic shift described. At 24 hours, it is feasible that catabolic metabolism led to increased autophagy, a process in which cell surface and cytosolic proteins are endocytosed and degraded to amino acids. This process may explain the drop in amino acids we observed at 24 hours. Interestingly, our transcriptomics showed significant changes in Muc5b and Fos genes at 24 hours; both genes are linked to membrane trafficking, which changes during autophagy[238].

Independent of these proposed mechanisms, we believe these data are important for pragmatic reasons. The first LNP-based RNA drug was approved by the FDA in 2018[1]; its

translation was aided by studies that identified the mechanism of action by which the LNP targeted hepatocytes[84]. Second-generation RNA therapies, which also target hepatocytes, have also shown promise in clinical trials[2]. Once again, the delivery mechanism of action is understood[239-241]. In order to continue realizing the potential of RNA drugs, it will be important to develop a more sophisticated understanding of the genes and pathways that enhance, or in this case, block RNA delivery.

Our data suggest that metabolic signaling can alter LNP delivery *in vivo* in unexpected ways. We believe these data constitute early steps toward an important goal: Exploiting natural differences in cell signaling in order to improve cell type-specific nanoparticle delivery[242]. We eventually hope to identify signaling that promotes the activity of a therapeutic RNA in a target cell type while reducing its activity in off-target cells. In addition, although we do not have direct evidence to date, these results may have implications for tumor delivery. Tumor delivery is thought to be limited, in part, by tortuous vasculature and pressure gradients that reduce efflux from the bloodstream. Our data posit a complementary mechanism: Highly metabolic tumor cells, particularly those exhibiting metabolic changes characterized as the Warburg effect, may be less likely to translate exogenous mRNA that nanoparticles deliver into the cytoplasm. This may be especially important for tumors characterized by overactive PIP3 signaling[243, 244].

It is important to acknowledge the limitations of this study. First, we were unable to identify the non-clathrin and non-caveolin pathways that were affected by PIP3. Our cell uptake data also did not elucidate the extent to which endocytosis or endosomal escape was affected by PIP3. Second, our *in vivo* studies were in mice; future work will need to evaluate whether the same results are observed in other models, most notably NHPs.

Despite these caveats, we believe these data provide evidence that cellular metabolism affects nanoparticle delivery, and more generally, that cell metabolism needs to be considered when designing RNA therapies.

5.4 Materials and Methods

Chemical Synthesis. Microwave irradiations were performed using a Biotage Initiator. TLC was performed on precoated silica Gel GF plates and visualized using $KMnO_4$ stains. ^{1}H-NMR spectra were recorded at 400 MHz (Varian) using $CDCl_3$ with TMS or residual solvent as standard. ^{13}C-NMR spectra were recorded at 100 MHz (Varian) using $CDCl_3$ with TMS or residual solvent as standard. High-resolution mass spectra (HRMS) were recorded on a time-of-flight mass spectrometer by ESI. All other chemicals were obtained from commercial sources.

Synthesis of compound **3** (**Fig. 5.13**)[71]. Compound **1** (20 g, 41.9 mmol, 1 eq.) was charged in a 100 mL flask and trifluoroacetic acid (42 mL) was slowly added at 0ºC. The reaction mixture was stirred at room temperature for 30 minutes. The solvent was evaporated under reduced pressure and the crude product dissolved in DMF (5 mL) before being added dropwise to pyridine (300 mL) at 0ºC. The reaction mixture was stirred at room temperature overnight. Pyridine was evaporated under reduced pressure and the remaining white solid was washed with EtOAc (3 x 100 mL). Intermediate **2** was used in the next step without further purification. To a degassed solution of **2** (8.4 g, 13.0 mmol, 1 eq.) in AcOH/DCM (1/1, 300 mL) was added Pd/C (10 wt. %, 3.0 g). The reaction mixture was then degassed for 5 min with H_2 and stirred at room temperature under H_2 atmosphere overnight. After completion of the reaction, the reaction mixture was filtered over a Celite

pad and washed with MeOH (500 mL). The filtrate was concentrated under reduced pressure to obtain a crude yellow viscous oil. Precipitation of the crude with EtOAc (50 mL) followed by further EtOAc washes (3 x 50 mL) afforded compound **3** (4.8 g, 98% yield) as a white solid. **^{1}H NMR** (400 MHz, D_2O) δ 1.38-1.52 (m, 4H, CH2), 1.73-1.65 (m, 4H, CH2), 1.83-1.89 (m, 4H, CH2), 2.98 (t, J = 7.6 Hz, 4H, NCH2), 4.14 (t, J = 5.2 Hz, 2H, COCH). **^{13}C NMR** (100 MHz, D_2O) δ 21.0, 23.2, 26.3, 32.8, 39.0, 54.1, 170.2. **HRMS** (ESI, m/z) $C_{12}H_{25}N_4O_2$ $[M + H]^+$ calculated 257.1972, found 257.1968.

Synthesis of compound **cKK-E12** (**Fig. 5.13**)[71]. To a solution of **3** (84 mg, 0.22 mmol, 1 eq.) and 1,2-epoxydodecane (247 mg, 1.34 mmol, 6 eq.) in EtOH (2 mL) was added triethylamine (0.12 mL, 0.88 mmol, 4 eq.) before being stirred for 30 min at room temperature. The reaction mixture was then irradiated in the microwave reactor at 150ºC for 5 h. After completion of the reaction, the reaction mixture was evaporated under reduced pressure and purified via flash column chromatography (silica gel, gradient eluent: 1-2.0 % of MeOH/DCM then 2.0-4.0 % MeOH/DCM containing 0.5 % NH_4OH) affording **cKK-E12** (148 mg, 67%) as a light-yellow oil. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 0.87 (t, J = 6.8 Hz, 12 H, CH3), 1.21-1.35 (m, 64 H, CH_2), 1.37-1.65 (m, 16 H, CH_2), 1.71-1.95 (m, 4 H, CH_2), 2.19-2.66 (br, 12 H, NCH_2), 3.62 (m, 4 H, CHOH), 3.99 (m, 2 H, COCH). **HRMS** (ESI, m/z) $C_{60}H_{121}N_4O_6$ $[M + H]^+$ calculated 993.9281, found 993.9269.

Nanoparticle Formulation. Nanoparticles were formulated in a microfluidic device as previously described[26] by mixing a nucleic acid, an ionizable lipid, PEG, cholesterol, and a phospholipid. Nanoparticles were made with variable mole ratios of these constituents. The mass ratio for all the constituents: RNA was 7.5:1 for LNP1 and 10:1 for LNP2. LNP1 consisted of the ionizable lipid CKK-E12, cholesterol (Avanti Lipids, 700000P),

C_{14}PEG2K (Avanti Lipids, 880150P), and DOTAP (Avanti Lipids, 890890P) or NBD-DOTAP (Avanti Lipids, 810890P). LNP2 consisted of the ionizable lipid 7C1, cholesterol, C_{14}PEG2K, and DOPE (Avanti Lipids, 850725P). LNP3 consisted of the ionizable lipid cKK-E12, cholesterol, C_{18}PEG2K (Avanti Lipids, 880120P), and DOPE (Avanti Lipids, 850725P) or AF647-DOPE (Millipore-Sigma, 42247). Chemically modified mRNA (GFP or Cre) was purchased from TriLink. The mRNA was diluted in 10 mM citrate buffer (Teknova) and loaded into a syringe (Hamilton Company). The materials making up the nanoparticle (cKK-E12 or 7C1, cholesterol, PEG, DOPE or DOTAP) were diluted in 100% ethanol, and loaded into a second syringe. The citrate phase and ethanol phase were mixed together in a microfluidic device using syringe pumps.

Nanoparticle Characterization. LNP hydrodynamic diameter was measured using DLS. LNPs were diluted in sterile 1X PBS to a concentration of ~0.06 μg/mL and analyzed. LNPs were used if they met 3 criteria: diameter >20 nm, diameter <200 nm, and autocorrelation function with only 1 inflection point. Particles that met these criteria were dialyzed in 1X phosphate buffered saline (PBS, Invitrogen), and sterile filtered with a 0.22 μm filter.

Animal Experiments. All animal experiments were performed in accordance with the Georgia Institute of Technology's IACUC. C57BL/6J (#000664) and Ai14 LSL-Tomato (#007914) mice were purchased from the Jackson Laboratory. In all experiments, mice were aged 5-8 weeks, and N = 3-4 mice per group were injected intravenously via the lateral tail vein.

Nanoparticle and PIP3 Dosing. Mice were injected with a total Cre mRNA (TriLink) dose of 1.0 mg/kg and a PIP3 (Cayman Chemical) dose of 10 mg/kg. RNA concentration was determined using NanoDrop (Thermo Scientific).

Cell Culture. *In vitro* experiments were performed with iMAECs (provided by Hanjoong Jo at Emory)[154], HEK cells (GenTarget), iMAECs stably transduced with CAG-SpCas9-EGFP, produced in the Dahlman Lab, or mouse aortic endothelial cells (C57E κB-GFP) (provided by Martin Schwartz at Yale). In all cases, cells were maintained at 37°C and 5% CO_2 and cultured using previously established conditions. In all cases, cell media was supplemented by penicillin-streptomycin (500 U/mL penicillin G, 0.5 mg/mL streptomycin) (PenStrep, VWR) and 10% (v/v) FBS (VWR). HEKs were passaged with DMEM F-12 50/50 (Corning). iMAECs were passaged using DMEM with 1 g/L glucose, L-glutamine, and sodium pyruvate (Corning), supplemented by 1% (v/v) MEMNEAA (Sigma Aldrich), and 25 μg/mL ECGS (Emd Millipore). C57E κB-GFP were passaged with EGM-2 media (Lonza). Unless noted, cells were seeded in a 24-well plate at a density of 40k cells/well with 500 μL media/well. 24 hours later, LNPs were added with a total RNA dose of 50, 150, or 250 ng/well. Initial experiments done with L2K used a total RNA dose of 400 ng/well.

MTT assay. Cells were seeded in a 96-well plate at a cell density of 25k cells/well with 100 μL media/well. 24 hours later, PIP3 was added at a dose of 0, 10 or 20 μM. Six or 24 hours later, media was discarded and replaced with 50 μL of serum-free media (same as described above, without FBS) and 50 μL of MTT reagent (Abcam). Cells were incubated for 3 hours at 37°C and treated with 150 μL of MTT solvent (Abcam). Plates were then

shaken for 15 minutes at 350 rpm. Formazan formation was quantified by measuring absorbance at OD = 590 nm using a BioTek Synergy HTX Multi-Mode Microplate Reader.

LDH assay. Cells were seeded in a 96-well plate at a cell density of 10k cells/well with 100 µL media/well. 24 hours later, PIP3 was added at a dose of 0, 10, or 20 µM. The positive control received 12 µL of CytoScan LDH Lysis Buffer (G Biosciences). 24 hours later, 50 µL of supernatant and 50 µL of the CytoScan LDH Assay Buffer (G Biosciences) were mixed together and incubated for 30 minutes at 37°C. 50 µL of CytoScan LDH Stop Solution (G Biosciences) was then added to the solution and absorbance was measured at OD = 490 nm using a BioTek Synergy HTX Multi-Mode Microplate Reader.

NF-κB assay. C57E cells, isolated from WT C57bl/6 aorta endothelium and immortalized with a py-MT lentiviral plasmid were used. Cells contain an NF-κB binding sequence upstream of a GFP reporter gene in their genome. The cells containing this construct constitutively express RFP and only express GFP when NF-κB is released from its receptor protein (IκB-α). Cells were seeded in a 24-well plate at a density of 40k cells/well with 500 µL media/well. 24 hours later, PIP3 was added at a dose of 0, 10, or 20 µM. The positive control received 100 ng of LPS. 8 and 12 hours later, cells were analyzed via flow cytometry using a BD Accuri C6 Benchtop Flow Cytometer (**Fig. 5.2, I to K**).

Fixed-cell staining. Cells were plated in glass-bottom 24-well plates at a density of 40k cells/well one day prior to LNP delivery. Cells were fixed 6- or 24-hours post-transfection with 4% paraformaldehyde (Electron Microscopy Sciences) for 10 min at room temperature before permeabilization with 0.2% Triton X-100 (Sigma-Aldrich) for 5 min at room temperature. To stain for endosomes, cells were first blocked for nonspecific binding

with 5% BSA (Sigma Aldrich) for 30 minutes at 37°C, then incubated with a mixture of anti-Rab7 (Invitrogen), anti-EEA1 (Invitrogen), and anti-CD63 (Abcam) for 30 minutes at 37°C. Cells were then incubated with a Alexa-Fluor 568 secondary antibody (Invitrogen) for 30 minutes at 37°C. Nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) (Life Technologies), and coverslips were placed over the cells in the dish and mounted with Prolong Gold (Life Technologies).

Microscopy. Images were acquired with a Hamamatsu Flash 4.0 v2 sCMOS camera on a PerkinElmer UltraView spinning disk confocal microscope mounted to a Zeiss Axiovert 200M body with a 63x NA 1.4 plan-apochromat objective. Images were acquired with Volocity acquisition software (PerkinElmer). All images were linearly contrast enhanced. Colocalization analysis was performed via Volocity acquisition software by calculating the M1 and M2 coefficients for 30-40 cells per condition on unenhanced images.

Metabolomics. Metabolomics was performed for five groups (0 hours no PIP3, six hours no PIP3, six hours PIP3, 24 hours no PIP3, 24 hours PIP3). Cells were seeded in 6-well plates at 500,000 cells/well and cultured for 24 hours before the start of the experiment. Four samples per time point and condition were created, with a fifth sample created for DNA quantification. Specifically, each replicate was created from two wells in different plates treated with the same condition. Cells were manually counted with a hemocytometer to check approximate number and viability. Methanol-based metabolite extraction was performed at each timepoint with, -80˚C methanol added to each sample before subsequent storage at -20˚C until analysis. Methanol based extraction preserves polar metabolites, a standard in cell-based metabolic extraction[245]. The fifth sample was treated by aspirating all PBS and adding digestion buffer consisting of 1xTris-EDTA and proteinase K. Internal

standards for each sample were added based upon cell number. Cell counting performed was performed on the fifth sample collected for each condition via the Promega Quant-iT™ PicoGreen® method utilizing pico-green dye reagent and Promega fluorescent plate reader to quantify dsDNA concentration as according to the manufacturer protocol. Internal standards of ribitol and C^{13}-glucose were added to the methanol-extracted samples in concentrations of 1 µL/1x10^6 cell and 2 µL/1x10^6 cells, respectively. and then dried in a speed vacuum system (Savant® SpeedVac SPD1030) overnight at 45˚C and 5.1 mTorr. After drying, samples consisted of only a dehumidified pellet of extracted metabolites and were stored at -80˚C. Sample derivatization procedure was performed as described previously[245]. Briefly, addition of the derivatizing agent methoxyamine (added alone for 1.5 hours) and then further addition of N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) (6 hours) was performed for samples, then placed into vials loaded to the GCMS system (Shimadzu® GCMS QP2010 ultra). No samples were left in derivatizing agents more than 22 hours, as this has been shown to degrade the metabolites present[246]. The GCMS performed 1 hour runs with a linear heating regime of the GC column from 70˚C to 310˚C.

Bioinformatics Analysis for Metabolomics. Data were collected on the GCMS post-run analysis software (Shimadzu®). Raw data containing the peak intensities were compared to a pre-existing custom-built library of metabolites containing 74 metabolites found within mammalian cell metabolism. This produced peak areas present for each of these metabolites. The peak areas of each run were normalized to area of the ribitol peak, one of the internal standards added. As per the previous section, this normalization procedure provides a method to normalize metabolite peak areas. The ribitol peak areas were

consistent within < 1% variation across all runs of each sample, an indication of run consistency and equipment accuracy. The ratio of the two C^{13}-glucose derivatives (second internal standard added) were taken for each run. For each run, if the ratio was greater than one standard deviation away from the average ratio of all runs in that sample, the run was removed. This ensured that peak area values were statistically consistent between runs. Metabolites with a coefficient of variation greater than 25% across remaining runs for each sample were also removed. The average of each metabolite across all runs in a sample was calculated to be effective metabolite presence in the original sample. The final data contained a single value for metabolite presence in each sample –four samples per condition and time point. Statistical analysis was performed on the data through dimensionality reduction techniques and significance analysis. Importantly, no additional data cleaning was performed specifically for any of these techniques. PCA and hierarchical clustering was performed in R and visualized using ggplot2 and heatplot packages. PCA considered three principal components, covering only > 50% of observed variance in the data; this indicates the data were variable in a large number of dimensions (metabolites). Identification of significantly up- and down-regulated metabolites between two groups was performed using the method of SAM as described previously[247], implemented using the SAMR package in R. Importantly, the delta value for SAM was calculated using the latest method described in the SAMR release documentation. Comparisons of metabolites between the PIP3 containing and non-PIP3 containing groups at six and 24 hours were compared and mapped to box plots and pathway maps for visualization and further analysis.

Whole-Transcriptome Sequencing. iMAECs were seeded in a 6-well plate at a seeding density of 500,000 cells/well. Cells were cultured for 24 hours as written above and then treated with 10 μM PIP3 or no PIP3. Six or 24 hours later RNA was harvested from the cells. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) following the manufacturers protocol. The quantity and quality of RNA were examined by Qubit RNA HS Assay Kit, Qubit™ 3.0 Fluorometer (Life Technologies) and Agilent 2100 Bioanalyzer (Agilent Technologies). 1μg of high-quality RNA from each sample was used for cDNA synthesis and sequencing, using the Truseq Stranded mRNA protocol from Illumina (Illumina Inc.). After generating the clusters, library sequencing was performed on an Illumina Nextseq platform, to create paired end reads with a length of 75bp.

Bioinformatics Analysis for RNA-Sequencing. The sequenced reads were trimmed and aligned to the mouse genome (mm10) using Isas analysis software 3.19.1.12, SAMtools 0.1.20, STAR aligner STAR_2.6.1a, Salmon quantification software 0.11.2 and Strelka Variant Calling software 2.9.9. Differential expression analysis was determined using DESeq2[248]. Only genes with a P adjusted value (false-discovery rate) or less than 0.05 and a fold change greater than 1.5 were included within the subsequent GO and KEGG pathway analysis. To understand the functions of the differentially expressed genes, GO functional enrichment and KEGG pathway analysis were carried out by the Enrichr web server[249] and KEGGMapper (https://www.kegg.jp/) respectively.

Cell Isolation & Staining. Mice were sacrificed 3 days after administration of LNPs and immediately perfused with 20 mL of 1X PBS through the right atrium. In all cases, the lung, liver, and spleen were isolated following perfusion. All tissues were finely minced with micro dissecting scissors, and then placed in a 1X PBS solution containing

Collagenase Type I (450 U / mL) (Sigma Aldrich), Collagenase XI (125 U / mL) (Sigma Aldrich), and Hyaluronidase (60 U / mL) (Sigma Aldrich) at 37ºC and 750 rpm for 45 minutes. The digestive enzyme for spleen included Collagenase IV (10 mg / mL) (Sigma Aldrich)[145]. Digested tissues were passed through a 70 μm filter and red blood cells were lysed with red blood cell lysis buffer (Alfa Aesar). Cells were stained to identify specific cell populations and flow cytometry was performed using a BD FacsFusion cell analyzer. Antibody clones used were anti-CD31 (390, BioLegend), anti-CD45.2 (104, BioLegend), anti-CD11b (M1/70, BioLegend), anti-CD68 (FA-11, BioLegend), anti-CD47 (miap301, BioLegend). PE conjugated anti-CD47 was used as a compensation control for tdTomato when running flow cytometry for *in vivo* experiments. Representative gating strategies for liver, lung, and spleen populations using control Ai14 mice injected with PBS are included in (**Fig. 5.14A-F**).

Statistical Analysis. Unless otherwise noted, statistical analyses were done using GraphPad Prism 8. Unless otherwise noted, data are plotted as mean ± standard error mean. As indicated in each figure caption, an unpaired t-test, 1-way ANOVA, or 2-way ANOVA was used to analyze the data.

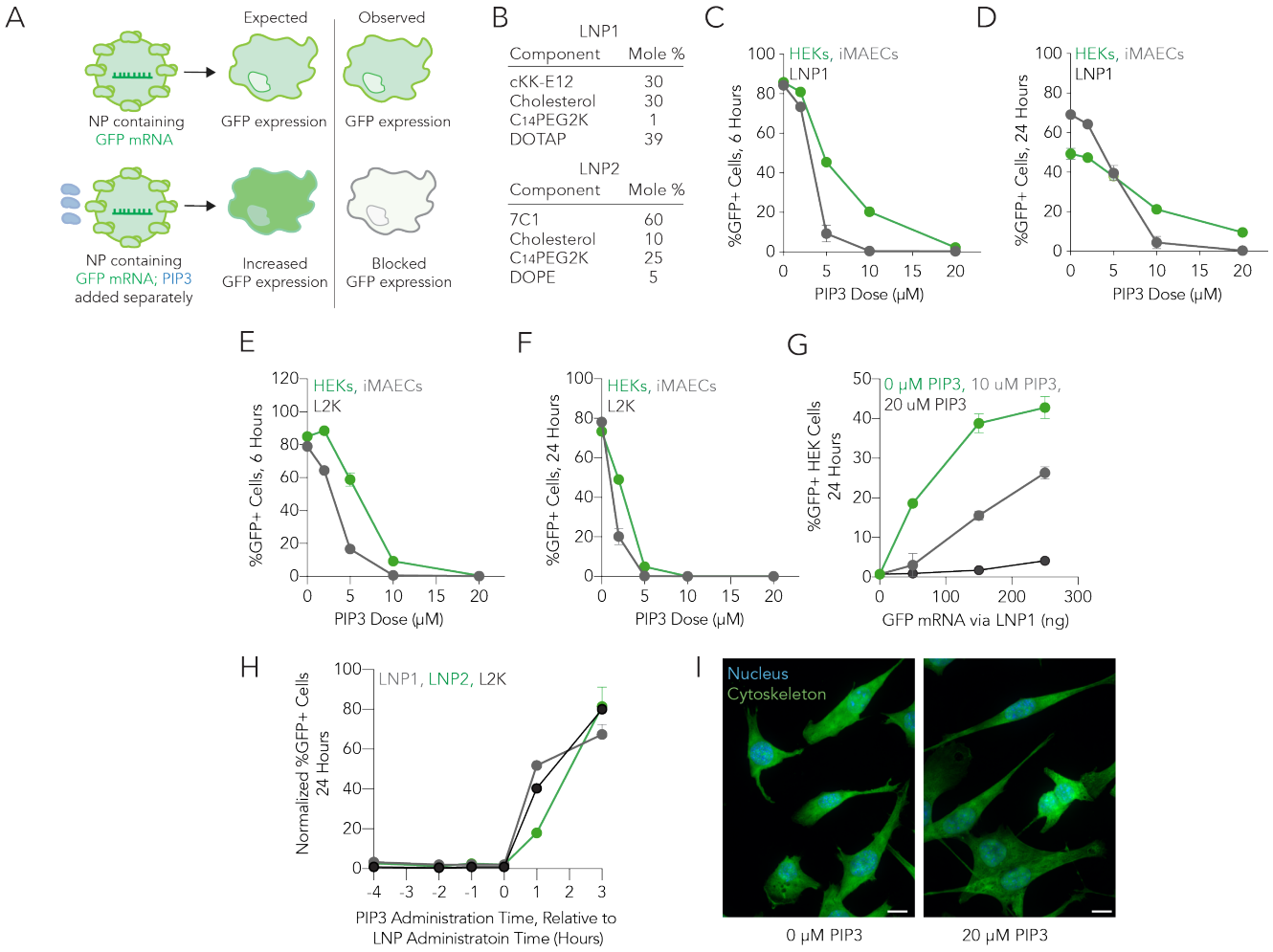
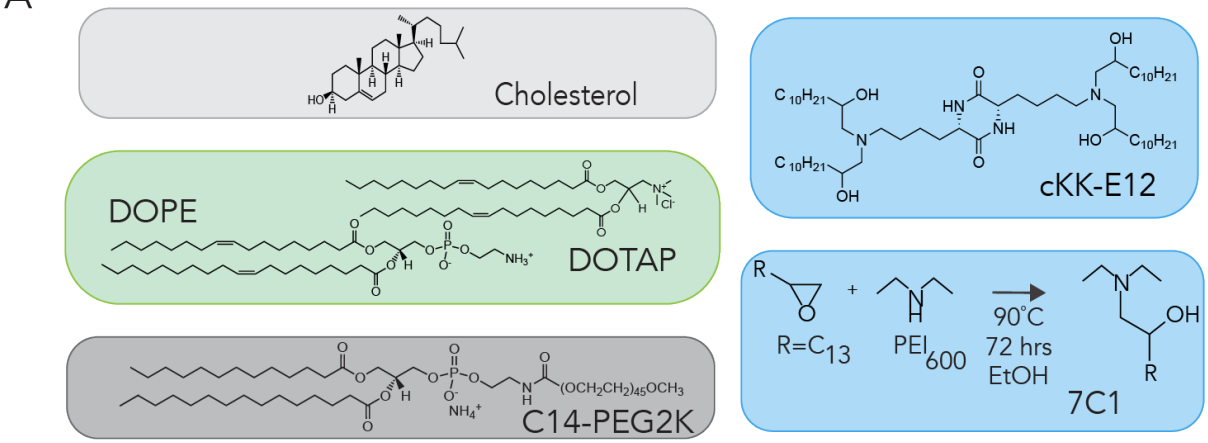
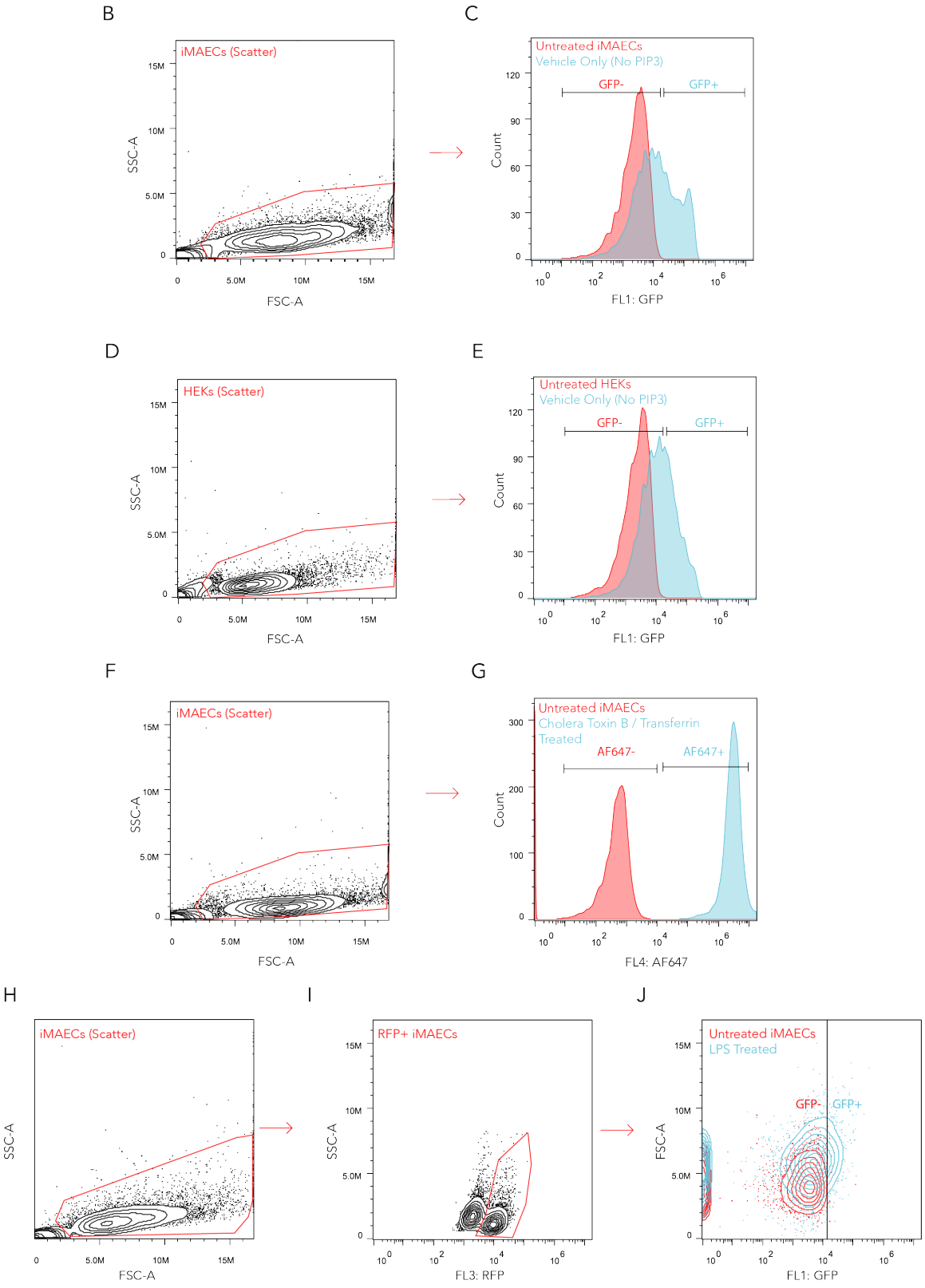


Figure 5.1 PIP3 reduces nanoparticle (NP)-mediated mRNA delivery *in vitro.* (A) (A) PIP3 blocks the functional delivery of nanoparticles carrying GFP mRNA. The data suggest this is driven in part by reduced endosomal escape. (B) LNP1 and LNP2 chemical composition. (C,D) Concurrent delivery of PIP3 and 50 ng of LNP1 carrying GFP mRNA led to decreased GFP expression in cells (C) six hours and (D) 24 hours after transfection. Reduced GFP expression was also observed at (E) six and (F) 24 hours when mRNA was carried by L2K and administered at a dose of 400 ng / well. (G) Concurrent delivery of PIP3 and varying doses of LNP1 carrying GFP mRNA shows that the percentage of GFP+ cells decreases as the amount of PIP3 increases. (H) PIP3 administered before, concurrently, or after 50 ng of LNP1, 150 ng of LNP2 and 400 ng of L2K carrying GFP mRNA leads to a decrease in GFP expression. (I) PIP3 did not lead to changes in cell morphology 24 hours after PIP3 administration. Scsale bar = 10 μm. Nuclei stained with DAPI and cytoskeleton (phalloidin) stained with GFP.

A


Cholesterol
DOPE
DOTAP
C14-PEG2K
cKK-E12
R=C13
PEI600
90˚C
72 hrs
EtOH
7C1


B
iMAECs (Scatter)
SSC-A
FSC-A
C
Untreated iMAECs
Vehicle Only (No PIP3)
GFP-
GFP+
Count
FL1: GFP
D
HEKs (Scatter)
SSC-A
FSC-A
E
Untreated HEKs
Vehicle Only (No PIP3)
GFP-
GFP+
Count
FL1: GFP
F
iMAECs (Scatter)
SSC-A
FSC-A
G
Untreated iMAECs
Cholera Toxin B / Transferrin
Treated
AF647-
AF647+
Count
FL4: AF647
H
iMAECs (Scatter)
SSC-A
FSC-A
I
RFP+ iMAECs
SSC-A
FL3: RFP
J
Untreated iMAECs
LPS Treated
GFP-
GFP+
FSC-A
FL1: GFP

K

Figure 1C	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	*	****
5 μM PIP3	****	****
10 μM PIP3	****	****
20 μM PIP3	****	****

L

Figure 1D	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	ns	ns
5 μM PIP3	****	****
10 μM PIP3	****	****
20 μM PIP3	****	****

M

Figure 1E	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	****	ns
5 μM PIP3	****	****
10 μM PIP3	****	****
20 μM PIP3	****	****

N

Figure 1F	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	****	****
5 μM PIP3	****	****
10 μM PIP3	****	****
20 μM PIP3	****	****

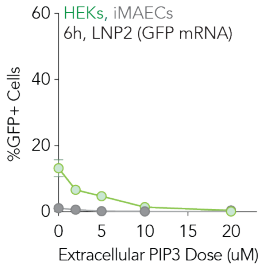
O

Figure 1G	HEKs (ng GFP mRNA)								
	50			150			250		
0	-	****	****	-	****	****	-	****	****
10	****	-	ns	****	-	****	****	-	****
20	****	ns	-	****	****	-	****	****	-
μM PIP3	0	10	20	0	10	20	0	10	20

P

Figure 1H	iMAECs		
	0 μM PIP3		
	L2K	LNP1	LNP2
-4 hrs	****	****	****
-2 hrs	****	****	****
-1 hrs	****	****	****
0 hrs	****	****	****
1 hrs	****	****	****
3 hrs	ns	ns	ns

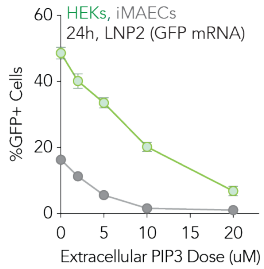
Q



Supplementary Figure 1P Statistics

	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	****	ns
5 μM PIP3	****	ns
10 μM PIP3	****	ns
20 μM PIP3	****	ns

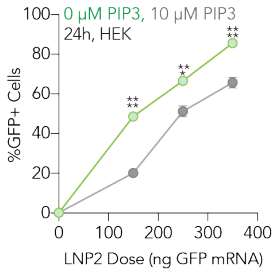
R



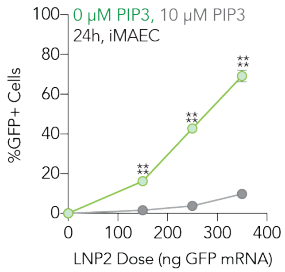
Supplementary Figure 1Q Statistics

	HEKs	iMAECs
	0 μM PIP3	
2 μM PIP3	****	**
5 μM PIP3	****	****
10 μM PIP3	****	****
20 μM PIP3	****	****

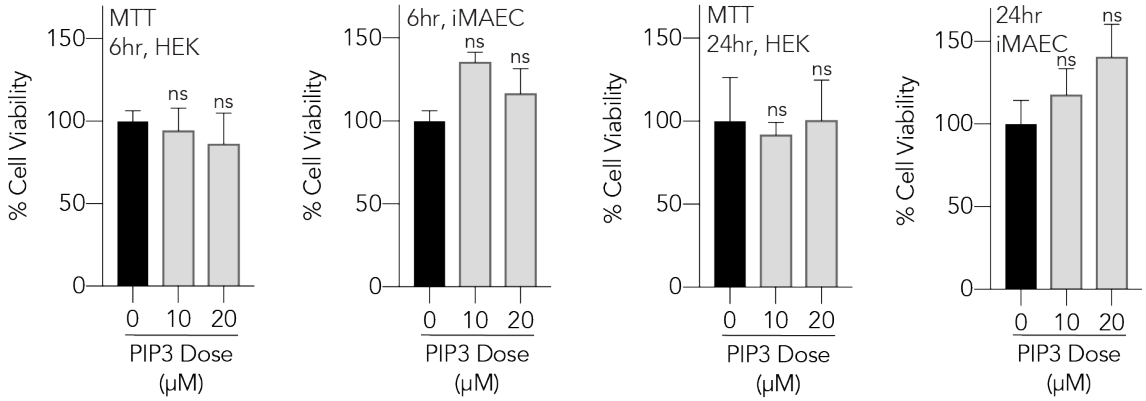
S



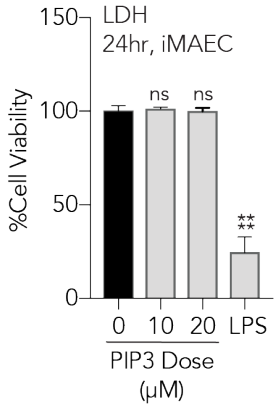
T



U


MTT
6hr, HEK
6hr, iMAEC
MTT
24hr, HEK
24hr
iMAEC
% Cell Viability
150
100
50
0
ns
0 10 20
PIP3 Dose
(μM)

V


LDH
24hr, iMAEC
%Cell Viability
150
100
50
0
ns
**
0 10 20 LPS
PIP3 Dose
(μM)

W

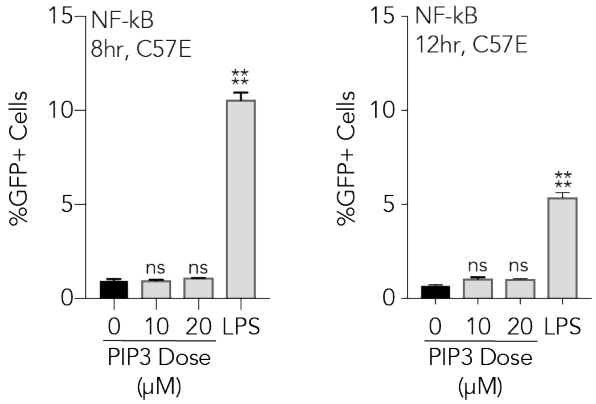

NF-kB
8hr, C57E
NF-kB
12hr, C57E
%GFP+ Cells
15
10
5
0
ns
**
0 10 20 LPS
PIP3 Dose
(μM)

Figure 5.2 LNP components and *in vitro* translation and viability experiments. (A) LNP1 and LNP2 chemical components. (B to J) Representative *in vitro* gating strategies for HEKs and iMAECs. The statistical analyses for Fig. 1, C to F are shown in (K to N) and compare the positive control (LNP only, no PIP3) to the PIP3 experimental groups using a 2-way ANOVA. Statistical analysis for Fig. 1G is shown in (O) and compares the positive control (LNP only, no PIP3) to each experimental group using a 2-way ANOVA. Statistical analysis for Fig. 1H is shown in (P) and compares the positive control (LNP only at time = 0 hrs., no PIP3) to each experimental group using a 2-way ANOVA. *p<0.032, **p<0.0021, ***p<0.0002, ****p<0.0001. Increasing doses of PIP3 led to a decrease in GFP protein expression in cells treated with LNP2 containing GFP mRNA (Q) 6 hours and (R) 24 hours after transfection. The percentage of GFP+ cells decreases with different concentrations of PIP3 despite an increase in the administered dose of GFP mRNA in (S) HEKs and (T) iMAECs. (U) PIP3 does not reduce cell viability as measured by the reduction of MTT to formazan in iMAECs or HEKs at 6 hours and 24 hours. (V) PIP3 does not reduce cell viability as measured by levels of secreted lactate dehydrogenase (LDH). (W) PIP3 does not cause inflammation as measured by levels of IκB degradation at 8 hours and 12 hours. Statistical analyses for (Q and R) compare the positive control (LNP only, 0 μM PIP3) to each experimental group (2, 5, 10, or 20 μM PIP3) using a 2-way ANOVA. Statistical analysis for (S and T) compares the positive control (LNP only, 0 μM PIP3) to the experimental group (10 μM PIP3 + LNP) for increasing doses of GFP mRNA using a 2-way ANOVA. Statistical analyses for (U) compare the negative control (untreated cells) to the experimental group (10 or 20 μM PIP3) using a Kruskal-Wallis ANOVA. Statistical analyses for (V and W) compare the negative control (untreated cells) to the experimental group (10 or 20 μM PIP3 or LPS) using a 1-way ANOVA. *p*-values are represented as p-value <0.0332 (*), ,0.0021 (**), <0.0002 (***), and <0.0001 (****).

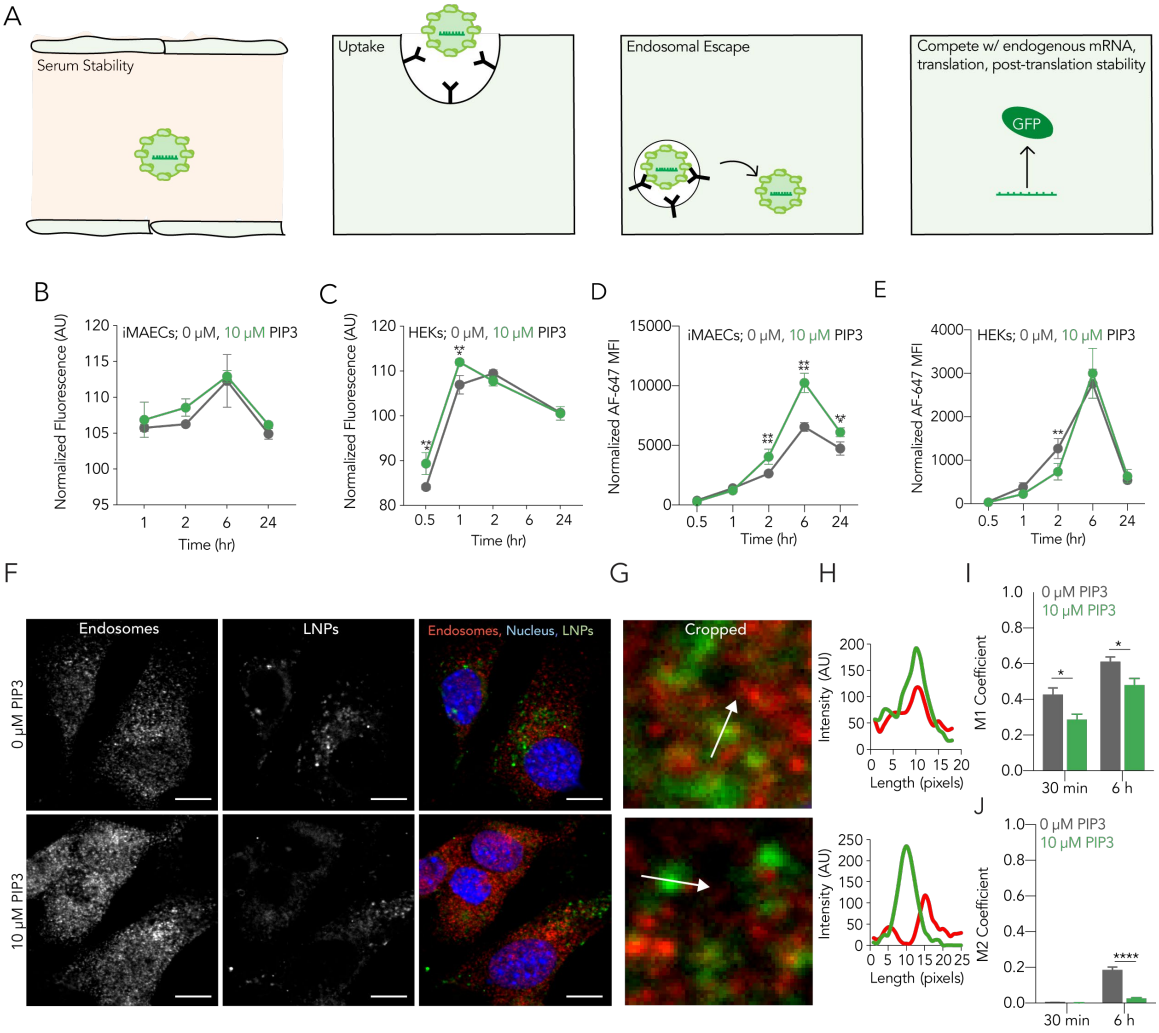


Figure 5.3 Analysis of the effects PIP3 has on LNP uptake and endosomal escape. (A) We reasoned that PIP3 could reduce LNP delivery could be driven by inhibiting serum stability, cell uptake, endosomal escape, or the metabolic state of the cell. (B-E) LNPs containing GFP mRNA were formulated with fluorescent phospholipids and administered to cells. Cell normalized fluorescence was determined relative to an untreated control at each timepoint. Normalized fluorescence did not change over time in (B) iMAECs and did change slightly in (C) HEKs after administration of LNP1 and PIP3. Similarly, PIP3 had minimal effect on LNP2 uptake, measured by mean fluorescent intensity (MFI), in either (D) iMAECs or (E) HEKs. Normalized AF-647 MFI was also determined relative to an untreated control at each timepoint. p-values are represented as p-value <0.0332 (*), <0.0021 (**), <0.0002 (***), and <0.0001 (****). (F) In cells treated with PIP3, LNPs (red) co-localize with endosomes (green) whereas the opposite is observed without PIP3. (G) Colocalization images of endosomes and LNPs can be used to draw (H) representative line profiles. (I) The M1 coefficient shows a significant difference in colocalization of LNP and endosome for cells treated with PIP3 after 30 minutes and 6 hours. (J) The M2 coefficient is low, indicating that, as expected, there are more endosomes than endosomes colocalized with LNPs. All microscopy images are shown with a 10μm scale bar. Although only a few representative cells are shown, the M1/M2 coefficient generation represents colocalization analysis of more than 30 cells per condition.

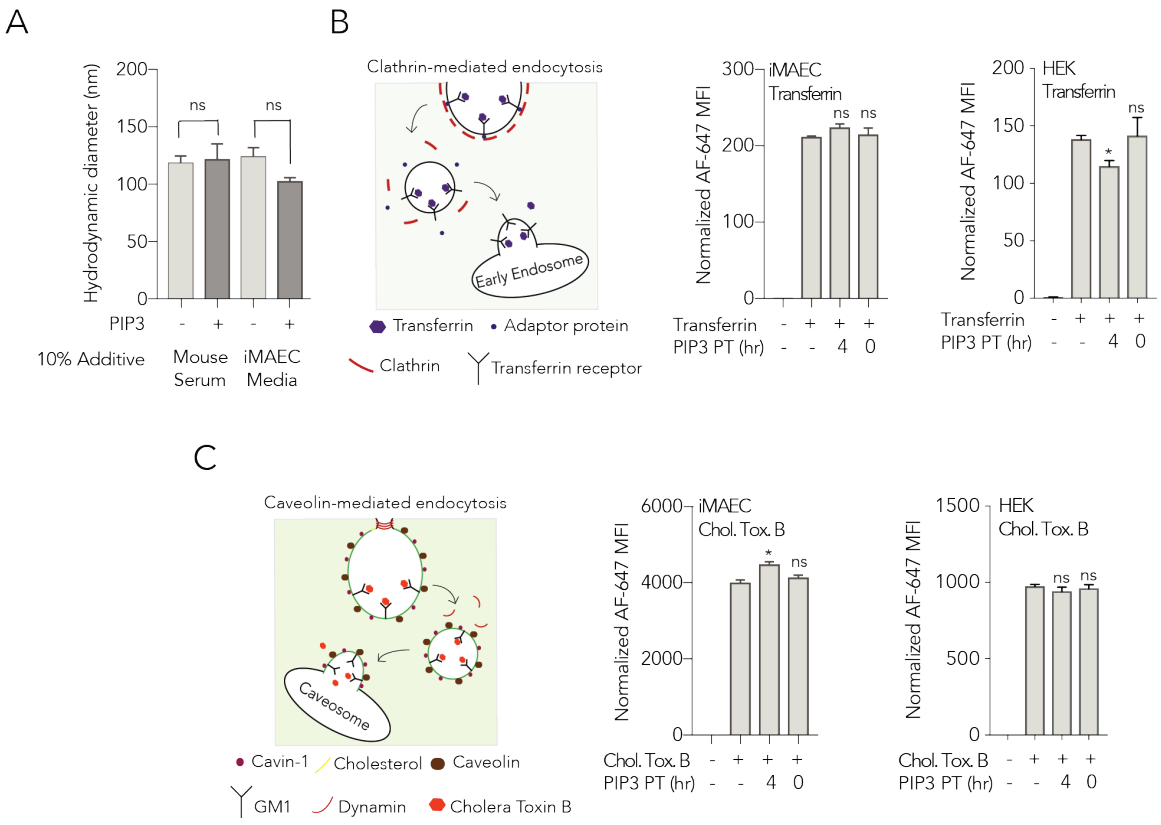


Figure 5.4 PIP3 does not affect LNP stability or canonical endocytosis pathways. (A) Addition of 10 µM does not significantly affect the hydrodynamic diameter of LNP1 suspended in PBS supplemented with mouse serum or iMAEC media. PIP3 does not alter transferrin or cholera toxin B uptake in iMAECs or HEKs. Alexa-647 fluorescence in cells was analyzed by flow cytometry after cells were treated with (B) fluorescent transferrin, a ligand endocytosed through a clathrin-dependent pathway or (C) cholera toxin B, a ligand endocytosed through a caveolin-dependent mechanism. Statistical analysis for (A) was done using a one-way ANOVA comparing the PIP3 group to the untreated group for each of the suspensions. Statistical analyses for (B and C) was done using an unpaired t-test comparing the positive control (cholera toxin B or transferrin only) to the treatment group (cholera toxin B or transferrin + PIP3) for both pre-treatment timepoints. *p*-values are represented as p-value <0.0332 (*), ,0.0021 (**), <0.0002 (***), and <0.0001 (****).

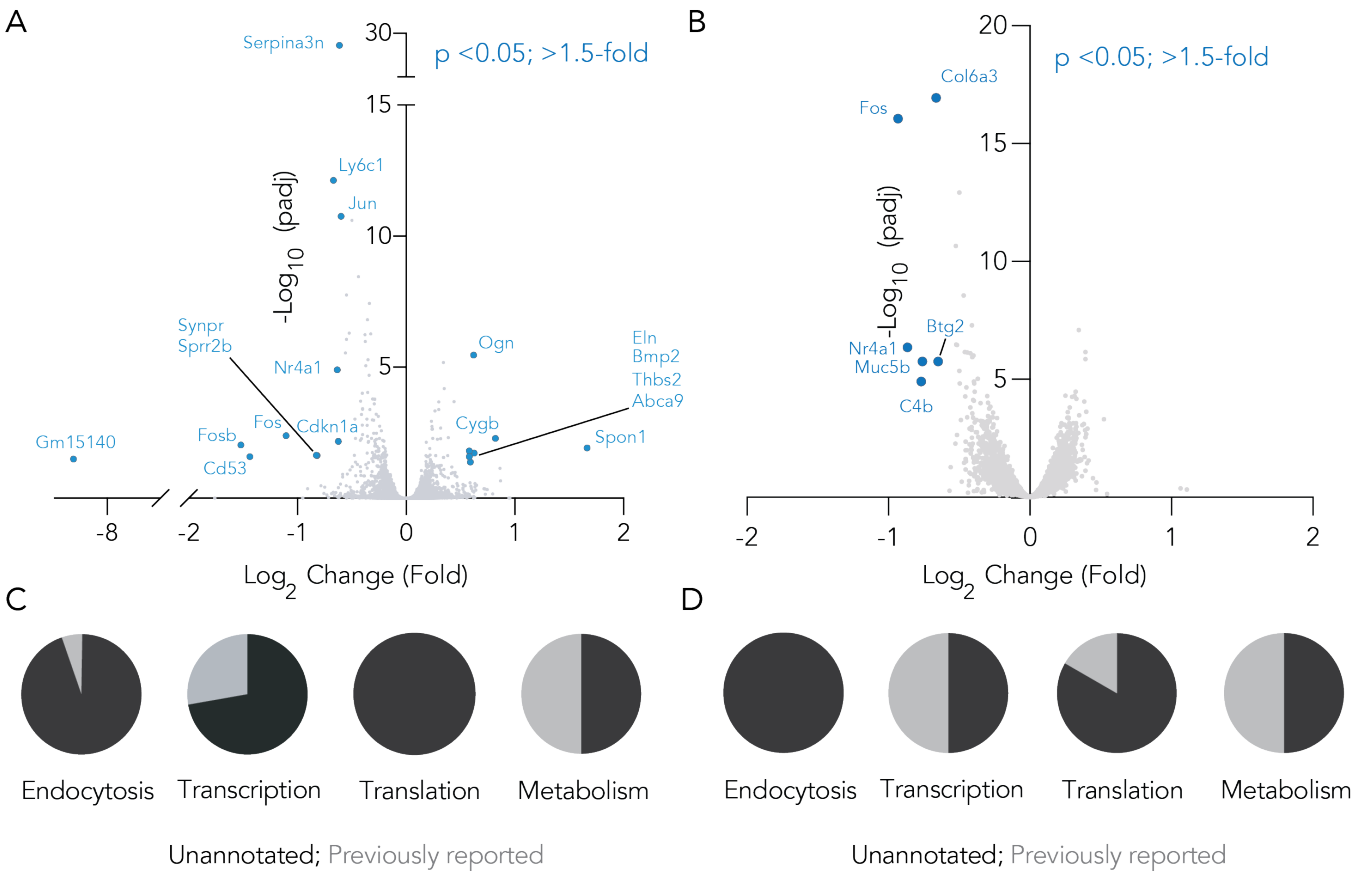
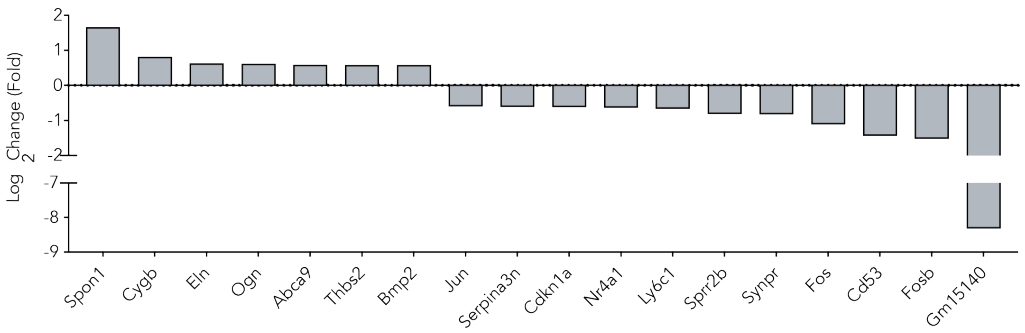


Figure 5.5 PIP3 changes the transcriptional profile of a cell. More genes are upregulated or downregulated (A) 6 hours after PIP3 administration than (B) 24 hours after PIP3 administration. Most perturbed genes were unannotated however there was a noticeable change in genes associated with metabolism both (C) 6 and (D) 24 hours after PIP3 administration.

A

Gene	Endocytosis	Transcription	Translation	Metabolism	Other
Spon1	Uncharacterized	Uncharacterized	Uncharacterized	Yes	Protein binding and processing. Cell adhesion.
Cygb	Uncharacterized	Uncharacterized	Uncharacterized	Yes	Electron transfer activity. Involved in hydrogen peroxide biosynthetic process; innate immune response; and superoxide metabolic process.
Eln	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Extracellular matrix organization; skeletal muscle tissue development; and stress fiber assembly.
Ogn	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Extracellular matrix structural constituent.
Abca9	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	ATPase-coupled transmembrane transporter activity. Lipid transporter activity.
Thbs2	Yes	Uncharacterized	Uncharacterized	Uncharacterized	Calcium ion binding activity. Heparin binding activity. Involved in negative regulation of angiogenesis.
Bmp2	Uncharacterized	Yes	Uncharacterized	Yes	Organ development. Regulation of cell differentiation. Regulation of transcription by RNA polymerase II.
Jun	Uncharacterized	Yes	Uncharacterized	Yes	Chromatin binding activity. Transcription regulatory region DNA binding activity. Animal organ development. Regulation of cell population proliferation. Regulation of protein phosphorylation.
Serpina3n	Uncharacterized	Uncharacterized	Uncharacterized	Yes	Serine-type endopeptidase inhibitor activity. Cytokine resonse. Peptide hormone response.
Cdkn1a	Uncharacterized	Uncharacterized	Uncharacterized	Yes	Cyclin binding activity. Cyclin-dependent protein serine/threonine kinase inhibitor activity. Response to radiation. Negative regulation of cell cycle. Signal transduction by p53 class mediator.
Nr4a1	Uncharacterized	Yes	Uncharacterized	Yes	DNA-binding transcription activator activity (RNA polymerase II-specific) Protein dimerization activity. Sequence-specific DNA binding activity. Cellular response to corticotropin-releasing hormone stimulus. Regulation of apoptotic process. Skeletal muscle cell differentiation.
Ly6c1	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Acetylcholine receptor activity
Sprr2b	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Structural molecule activity. Estradiol response.
Synpr	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Syntaxin-1 binding activity.
Fos	Uncharacterized	Yes	Uncharacterized	Yes	DNA-binding transcription activator activity (RNA polymerase II-specific). RNA polymerase II proximal promoter sequence-specific DNA binding activity. Chromatin binding activity. Positive regulation of osteoclast differentiation. Muscle stretch response. Skeletal muscle cell differentiation.
Cd53	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Myoblast fusion regulation.
Fosb	Uncharacterized	Yes	Uncharacterized	Yes	DNA-binding transcription activator activity (RNA polymerase II-specific). RNA polymerase II proximal promoter sequence-specific DNA binding activity. Cellular response to calcium ion. Regulation of transcription by RNA polymerase II.
Gm15140	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Spermatogenesis activity.

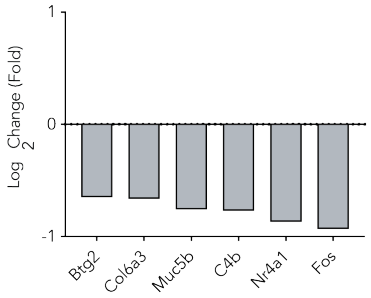
B



C

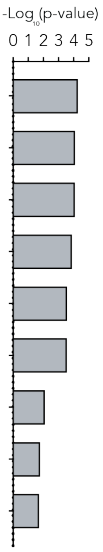
Gene	Endocytosis	Transcription	Translation	Metabolism	Other
Btg2	Uncharacterized	Yes	Yes	Yes	DNA-binding transcription repressor activity (RNA polymerase II-specific). Central nervous system development. Skeletal muscle cell differentiation.
Col6a3	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Extracellular matrix structural constituent . Growth plate cartilage chondrocyte morphogenesis activity.
Muc5b	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Macrophage activation regulation.
C4b	Uncharacterized	Uncharacterized	Uncharacterized	Uncharacterized	Carbohydrate binding activity. Complement component C1q binding activity. Complement activation. Immunoglobulin mediated immune response.
Nr4a1	Uncharacterized	Yes	Uncharacterized	Yes	DNA-binding transcription activator activity (RNA polymerase II-specific) Protein dimerization activity. Sequence-specific DNA binding activity. Cellular response to corticotropin-releasing hormone stimulus. Regulation of apoptotic process. Skeletal muscle cell differentiation.
Fos	Uncharacterized	Yes	Uncharacterized	Yes	DNA-binding transcription activator activity (RNA polymerase II-specific). RNA polymerase II proximal promoter sequence-specific DNA binding activity. Chromatin binding activity. Positive regulation of osteoclast differentiation. Muscle stretch response. Skeletal muscle cell differentiation

D



E

Domain	GO	Ontology
Molecular Function	0098811	Transcriptional repressor activity, RNA polymerase II activating transcription factor binding
Molecular Function	0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding
Molecular Function	0046982	Protein heterodimerization activity
Biological Process	1902895	Positive regulation of pri-miRNA transcription from RNA polymerase II promoter
Biological Process	0007178	Transmembrane receptor protein serine/threonine kinase signaling pathway
Biological Process	0071276	Cellular response to cadmium ion
Cellular Component	0090575	RNA polymerase II transcription factor complex
Cellular Component	0005719	Nuclear euchromatin
Cellular Component	0000228	Nuclear chromosome



F

Domain	GO	Ontology
Molecular Function	0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding
Molecular Function	0046982	Protein heterodimerization activity
Molecular Function	0035259	Glucocorticoid receptor binding
Biological Process	0045786	Negative regulation of cell cycle
Biological Process	0032490	Detection of molecule of bacterial origin
Biological Process	2000427	Positive regulation of apoptotic cell clearance
Cellular Component	0005796	Golgi lumen
Cellular Component	0090575	RNA polymerase II transcription factor complex
Cellular Component	0005788	Endoplasmic reticulum lumen

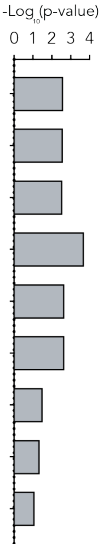
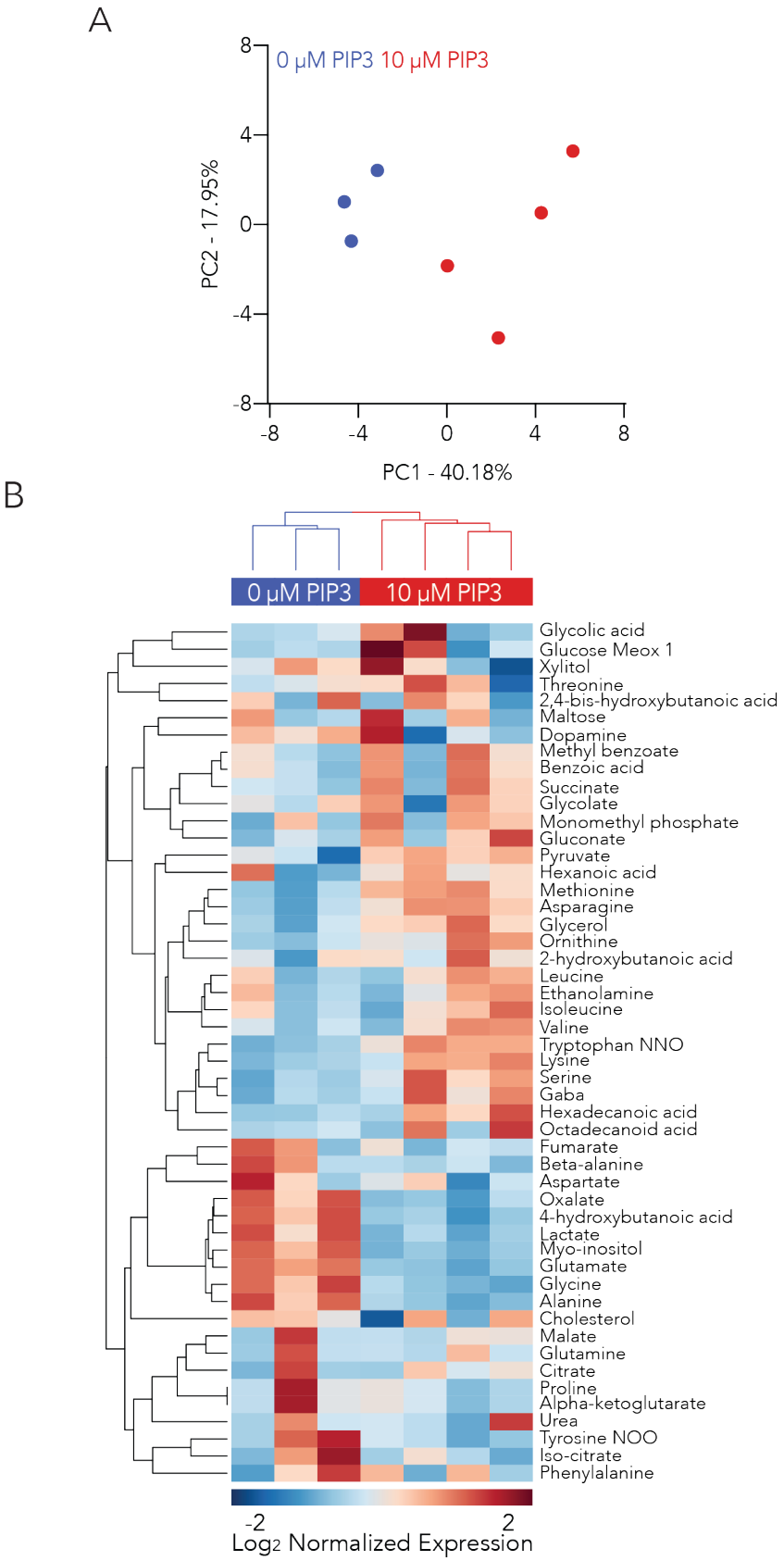


Figure 5.6 Transcriptional changes elicited by PIP3 treatment. (A) Table listing the transcripts identified as significantly differentially expressed six hours after exposure to PIP3 relative to control. Differentially expressed genes have a fold change>1.5 and p<0.05. A Pubmed search was used to identify whether the genes were associated with endocytosis, transcription, translation, metabolism or uncharacterized. A description of the known function of the gene is also included. (B) Log_2(Fold change) of significantly perturbed genes six hours after PIP3 exposure. (C) Table listing the transcripts identified as significantly differentially expressed 24 hours after exposure to PIP3 relative to control. Differentially expressed genes have a fold change>1.5 and p<0.05. A Pubmed search was used to identify whether the genes were associated with endocytosis, transcription, translation, metabolism or uncharacterized. A description of the known function of the gene is also included. (D) Log_2(Fold change) of significantly perturbed genes twenty-four hours after PIP3 exposure. (E to F) The top scoring GO pathways from each domain of the differentially expressed genes 6 (E) and 24 (F) hours after being exposed to PIP3, and the associated p-values from the Enrichr package. GO domains consist of biological processes, molecular functions, and


A
0 µM PIP3
10 µM PIP3
PC2 - 17.95%
PC1 - 40.18%
8
4
0
-4
-8
B
0 µM PIP3
10 µM PIP3
Glycolic acid
Glucose Meox 1
Xylitol
Threonine
2,4-bis-hydroxybutanoic acid
Maltose
Dopamine
Methyl benzoate
Benzoic acid
Succinate
Glycolate
Monomethyl phosphate
Gluconate
Pyruvate
Hexanoic acid
Methionine
Asparagine
Glycerol
Ornithine
2-hydroxybutanoic acid
Leucine
Ethanolamine
Isoleucine
Valine
Tryptophan NNO
Lysine
Serine
Gaba
Hexadecanoic acid
Octadecanoid acid
Fumarate
Beta-alanine
Aspartate
Oxalate
4-hydroxybutanoic acid
Lactate
Myo-inositol
Glutamate
Glycine
Alanine
Cholesterol
Malate
Glutamine
Citrate
Proline
Alpha-ketoglutarate
Urea
Tyrosine NOO
Iso-citrate
Phenylalanine
-2
2
Log2 Normalized Expression

C

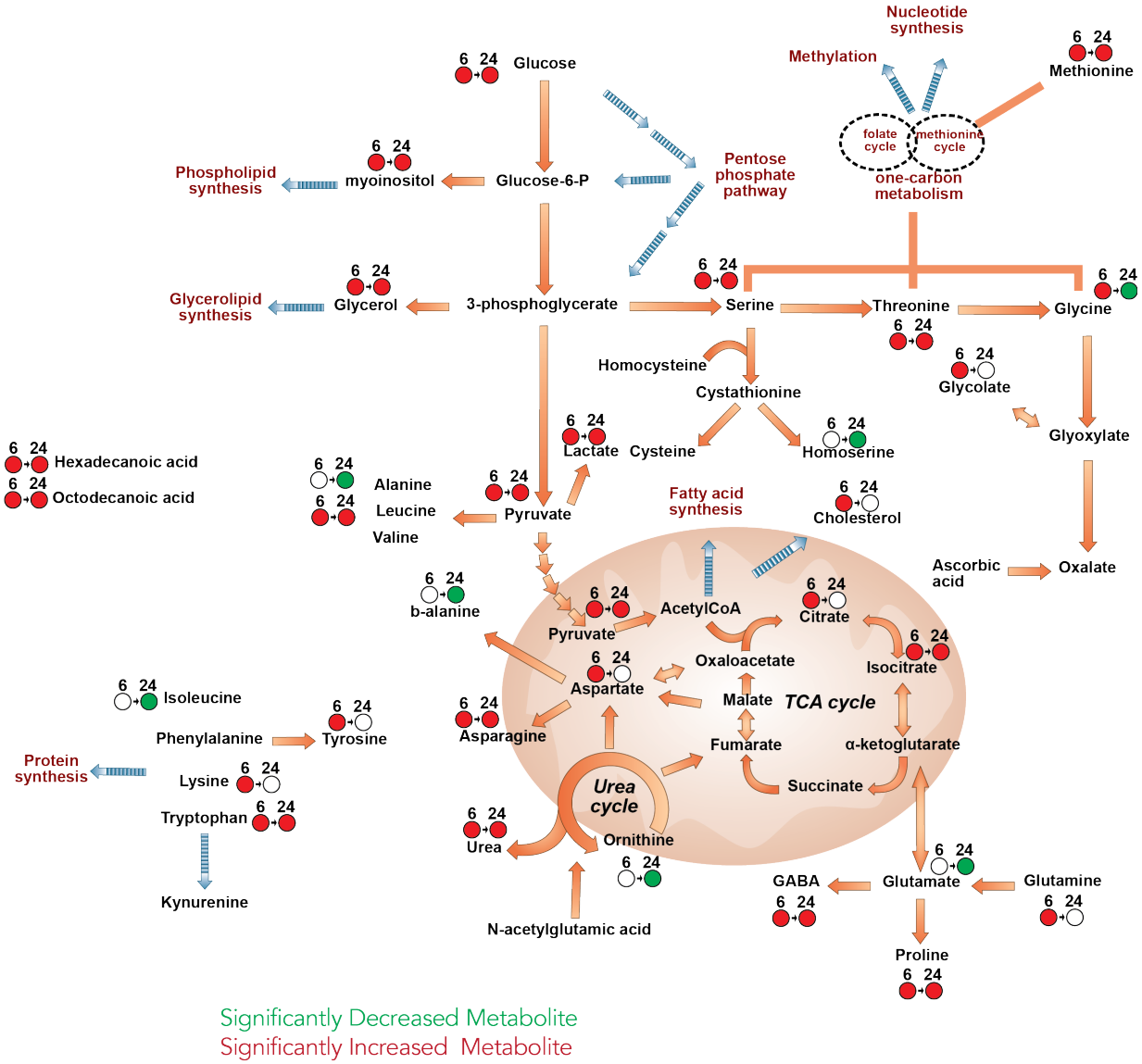


Figure 5.7 PIP3 changes the metabolic state of the cell. (A) PCA revealed distinct grouping between cells 24 hours after they were treated with PIP3 or not treated with PIP3. Similar distinctions between treated and untreated cells at 24 hours were found using (B) joint hierarchical clustering, in this case displaying metabolites across the 7 samples. Significance analysis identified metabolites which were then mapped upon a (C) metabolic pathway diagram, highlighting significantly upregulated and downregulated metabolites in the PIP3 positive group compared to the PIP3 negative control at both six and 24 hours. Common metabolic pathways in human metabolism are shown, with orange arrows representing multi-reaction steps in these pathways, maroon labels such as "protein synthesis" represent additional cellular metabolic pathways and phenotypes significantly influenced from the canonical metabolites and pathways listed.

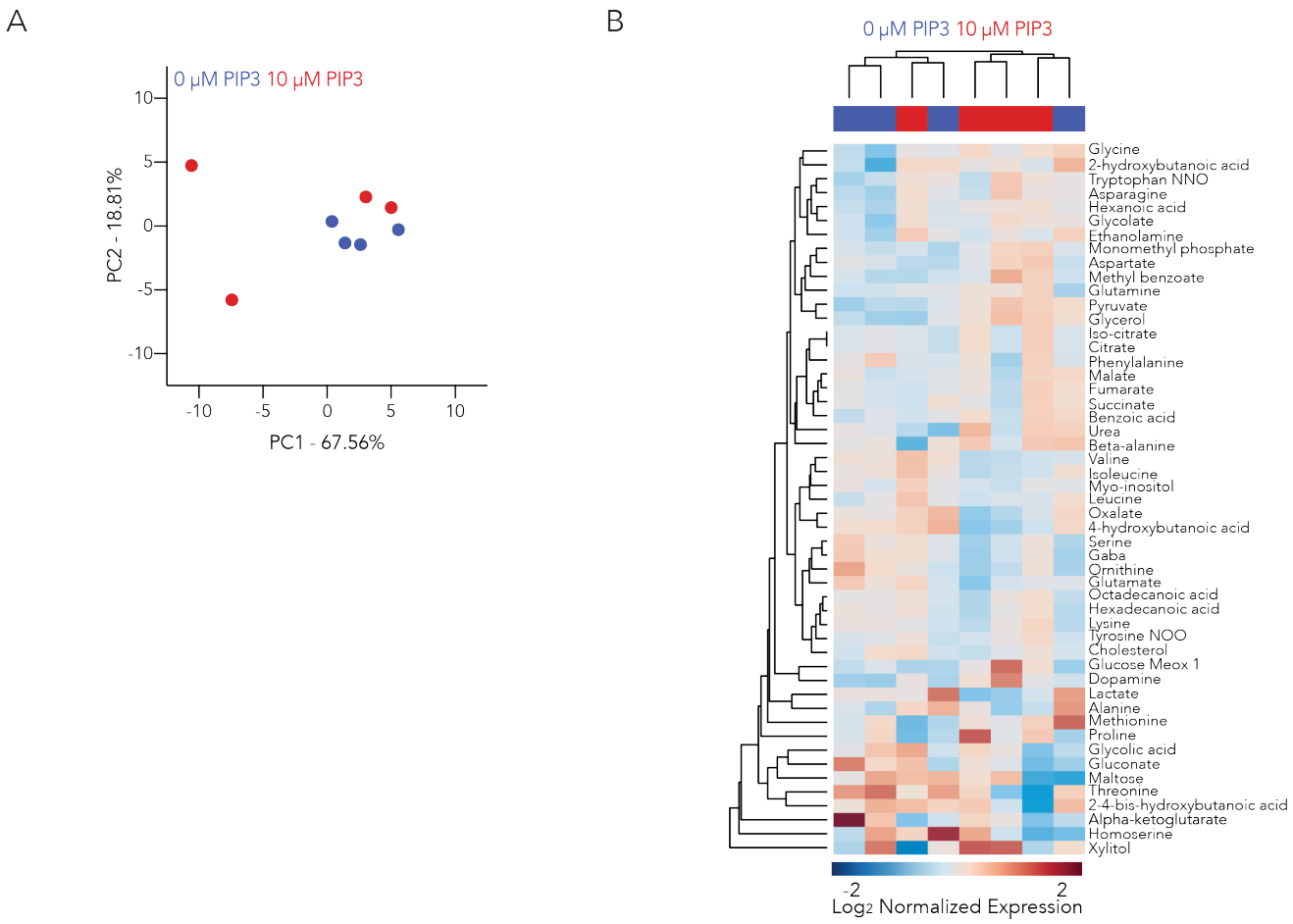


Figure 5.8 Results of metabolomic analysis for both PIP3 positive and negative cultures 6 hours after stimulation. Analysis involved statistical methods including (A) PCA and (B) joint hierarchical clustering and heat map.

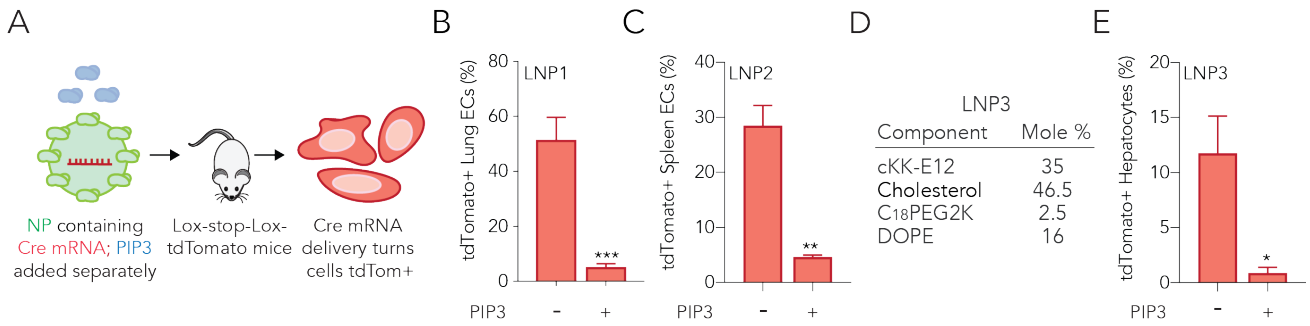


Figure 5.9 PIP3 reduces LNP mRNA delivery in vivo. (A) Ai14 mice were injected with 10 mg/kg of PIP3, and then immediately injected with LNP carrying 1 mg/kg Cre mRNA. PIP3 consistently blocks functional LNP1 delivery in (B) lung endothelial cells and functional LNP2 delivery in (C) splenic endothelial cells. (D) LNP3 was formulated to deliver Cre mRNA at 0.3 mg/kg. (E) PIP3 blocked LNP-mediated delivery of Cre mRNA to hepatocytes. Statistical analyses are done comparing the positive control (LNP only) to the treatment group (LNP + PIP3) using an unpaired t-test. p-values are represented as p-value <0.0332 (*), <0.0021 (**), <0.0002 (***), and <0.0001 (****).

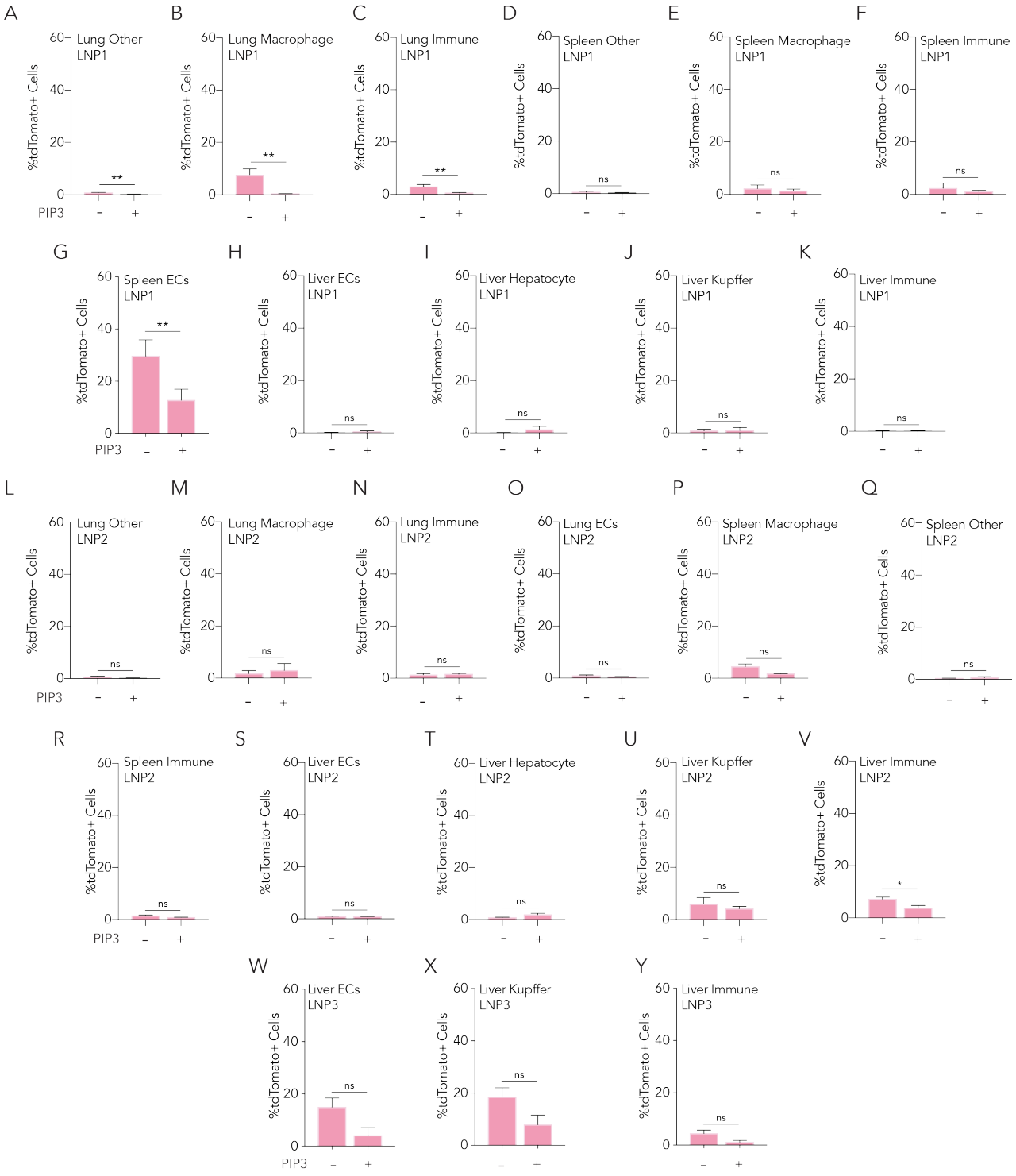

A
Lung Other LNP1
B
Lung Macrophage LNP1
C
Lung Immune LNP1
D
Spleen Other LNP1
E
Spleen Macrophage LNP1
F
Spleen Immune LNP1
G
Spleen ECs LNP1
H
Liver ECs LNP1
I
Liver Hepatocyte LNP1
J
Liver Kupffer LNP1
K
Liver Immune LNP1
L
Lung Other LNP2
M
Lung Macrophage LNP2
N
Lung Immune LNP2
O
Lung ECs LNP2
P
Spleen Macrophage LNP2
Q
Spleen Other LNP2
R
Spleen Immune LNP2
S
Liver ECs LNP2
T
Liver Hepatocyte LNP2
U
Liver Kupffer LNP2
V
Liver Immune LNP2
W
Liver ECs LNP3
X
Liver Kupffer LNP3
Y
Liver Immune LNP3
%tdTomato+ Cells
PIP3
ns

Figure 5.10 Off-target delivery of LNP-mediated mRNA delivery with PIP3 treatment. PIP3 response in blocking functional LNP1 delivery in (A) lung other cells, (B) lung macrophage cells, (C) lung immune cells, (D) splenic other cells, (E) splenic macrophages, (F) splenic immune cells, (G) splenic endothelial cells, (H) liver endothelial cells, (I) hepatocytes, (J) liver Kupffer cells, and (K) liver immune cells. PIP3 response in blocking functional LNP2 delivery in (L) lung other cells, (M) lung macrophage cells, (N) lung immune cells, (O) lung endothelial cells, (P) splenic macrophages, (Q) splenic macrophages, (R) splenic immune cells, (S) liver endothelial cells, (T) hepatocytes, (U) liver Kupffer cells, and (V) liver immune cells. PIP3 response in blocking functional LNP3 delivery in (W) liver endothelial cells, (X) liver Kupffer cells, and (Y) liver immune cells. Statistical analyses are done comparing the positive control (LNP only) to the treatment group (LNP + PIP3) using an unpaired t-test. *p*-values are represented as p-value >0.1234 (ns), <0.0332 (*), ,0.0021 (**), <0.0002 (***), and <0.0001 (****).

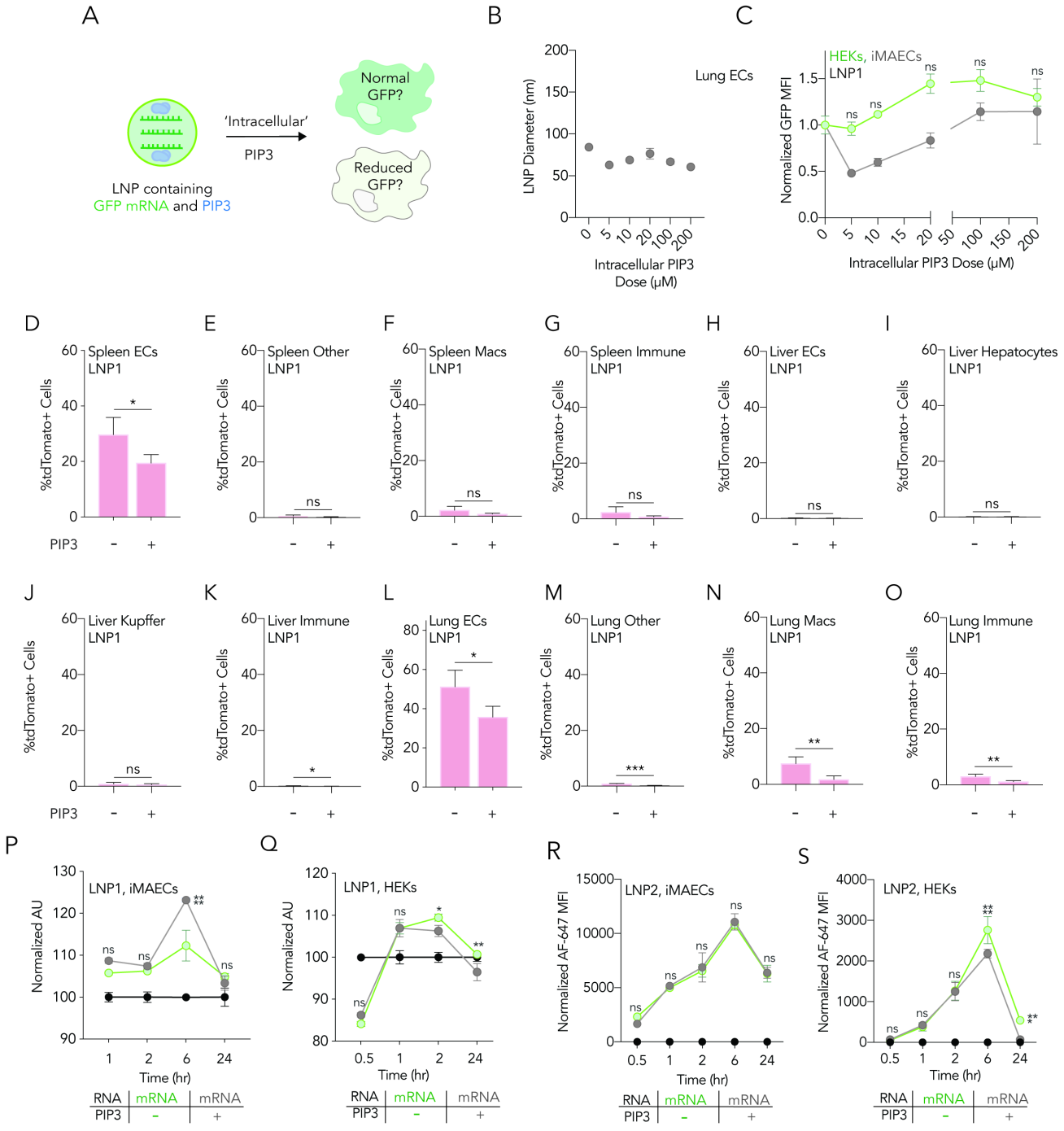
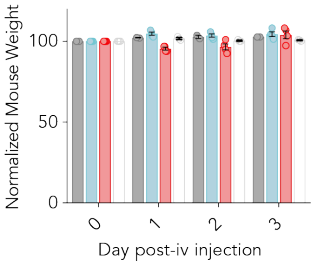

A
LNP containing
GFP mRNA and PIP3
'Intracellular'
PIP3
Normal
GFP?
Reduced
GFP?
B
Lung ECs
LNP Diameter (nm)
Intracellular PIP3
Dose (μM)
C
HEKs, iMAECs
LNP1
Normalized GFP MFI
Intracellular PIP3 Dose (μM)
D
Spleen ECs
LNP1
E
Spleen Other
LNP1
F
Spleen Macs
LNP1
G
Spleen Immune
LNP1
H
Liver ECs
LNP1
I
Liver Hepatocytes
LNP1
J
Liver Kupffer
LNP1
K
Liver Immune
LNP1
L
Lung ECs
LNP1
M
Lung Other
LNP1
N
Lung Macs
LNP1
O
Lung Immune
LNP1
%tdTomato+ Cells
PIP3
P
LNP1, iMAECs
Normalized AU
Time (hr)
Q
LNP1, HEKs
R
LNP2, iMAECs
Normalized AF-647 MFI
S
LNP2, HEKs
RNA
PIP3
mRNA

Figure 5.11 (A) PIP3 delivered inside of an LNP does not block delivery. (B) LNP diameter does not significantly change with varying PIP3 concentrations. (C) GFP mRNA and different doses of PIP3 delivered with LNP1 have opposite effects in iMAECs and HEKs. PIP3 formulated into LNP1 does not consistently impact tdTomato+ cells in (D) spleen endothelial cells, (E) spleen other cells, (F) spleen macrophages, (G) spleen immune cells, (H) liver endothelial cells, (I) hepatocytes, (J) liver Kupffer cells, (K) liver immune cells, (L) lung endothelial cells, (M) lung other cells, (N) lung macrophages, or (O) lung immune cells. PIP3 does not block uptake of fluorescent LNP1 containing mRNA in (P) iMAECs and (Q) HEKs. PIP3 has the same effect on fluorescent LNP2 containing mRNA in (R) iMAECs and (S) HEKs. Statistical analysis for (C) compares the positive control (LNP only, 0 µM PIP3) to each experimental group (2, 5, 10, 20, 100 or 200 µM PIP3) using a 2-way ANOVA. Statistical analyses for (D-O) are done comparing the positive control (LNP only) to the treatment group (LNP + PIP3) using an unpaired t-test. Statistical analyses for (P-S) compare the positive control (LNP only) to the treatment group (10 µM PIP3 + LNP) at each timepoint using a 2-way ANOVA. *p*-values are represented as p-value >0.1234 (ns), <0.0332 (*), ,0.0021 (**), <0.0002 (***), and <0.0001 (****).

A

LNP1					
Strain	Procedure	Day 0 (g)	Day 1 (g)	Day 2 (g)	Day 3 (g)
Ai14	PBS	29.8	30.4	30.6	30.6
Ai14	PBS	30.2	30.9	31	31
Ai14	LNP1	15.9	15.4	15.5	17
Ai14	LNP1	17.9	16.8	17	18.5
Ai14	LNP1	19.5	18.9	19.7	21.1
Ai14	LNP1	18.5	17.4	17.1	18
Ai14	LNP1 + 10μM PIP3 (Intracellular)	20.2	21.4	21.1	21.1
Ai14	LNP1 + 10μM PIP3 (Intracellular)	21.1	21.8	21.4	21.8
Ai14	LNP1 + 10μM PIP3 (Intracellular)	21.6	22.2	22.1	21.9
Ai14	LNP1 + 10μM PIP3 (Intracellular)	20.5	21.9	21.7	22.2
Ai14	LNP1 + 10mg/kg PIP3 (Extracellular)	27.1	27.3	27.1	27.3
Ai14	LNP1 + 10mg/kg PIP3 (Extracellular)	27.2	28	27.5	27.5
Ai14	LNP1 + 10mg/kg PIP3 (Extracellular)	26.3	26.7	26.3	26.3

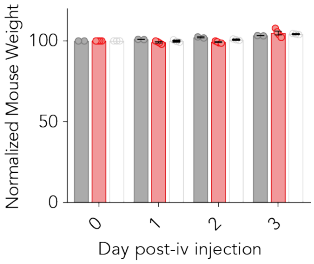
B



C

LNP2					
Strain	Procedure	Day 0 (g)	Day 1 (g)	Day 2(g)	W (72 hrs) (g)
Ai14	PBS	27.4	27.7	27.9	28.3
Ai14	PBS	23.7	23.9	24.3	24.5
Ai14	LNP2	26	25.5	25.7	27
Ai14	LNP2	25.3	25	25.3	27.3
Ai14	LNP2	22.2	22.1	22.1	22.7
Ai14	LNP2	23.7	23.7	23.4	25
Ai14	LNP2 + 10mg/kg PIP3 (Extracellular)	25.6	25.4	25.9	26.8
Ai14	LNP2 + 10mg/kg PIP3 (Extracellular)	25.8	26	26	26.8
Ai14	LNP2 + 10mg/kg PIP3 (Extracellular)	25.9	25.8	26.2	27
Ai14	LNP2 + 10mg/kg PIP3 (Extracellular)	26.1	25.9	26.2	27.2

D



E

LNP3					
Strain	Procedure	Day 0 (g)	Day 1 (g)	Day 2(g)	W (72 hrs) (g)
Ai14	PBS	15.3	16.9	16.4	16.7
Ai14	PBS	18.5	19.7	19	19.1
Ai14	LNP3	18.1	19	18.3	18.8
Ai14	LNP3	18.8	19.5	19.1	19.6
Ai14	LNP3	19.4	19.9	19.5	20
Ai14	LNP3	24.7	23.8	23.4	23.9
Ai14	LNP3 + 10mg/kg PIP3 (Extracellular)	19.6	19.7	19.8	20
Ai14	LNP3 + 10mg/kg PIP3 (Extracellular)	15.3	15.1	15.6	16.3
Ai14	LNP3 + 10mg/kg PIP3 (Extracellular)	18.8	19.3	18.9	19.6

F

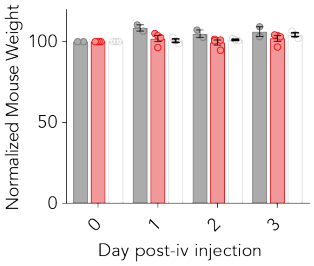


Figure 5.12 Mouse weights for *in vivo* experiments. (A) Mouse weights for LNP1 *in vivo* experiments from day of injection to day of sacrifice. (B) Normalized mouse weights for each LNP1 *in vivo* experiment compared to PBS injected mice. (C) Mouse weights for LNP2 *in vivo* experiments from day of injection to day of sacrifice. (D) Normalized mouse weights for each LNP2 *in vivo* experiment compared to PBS injected mice. (E) Mouse weights for LNP3 *in vivo* experiments from day of injection to day of sacrifice. (F) Normalized mouse weights for each LNP3 *in vivo* experiment compared to PBS injected mice.

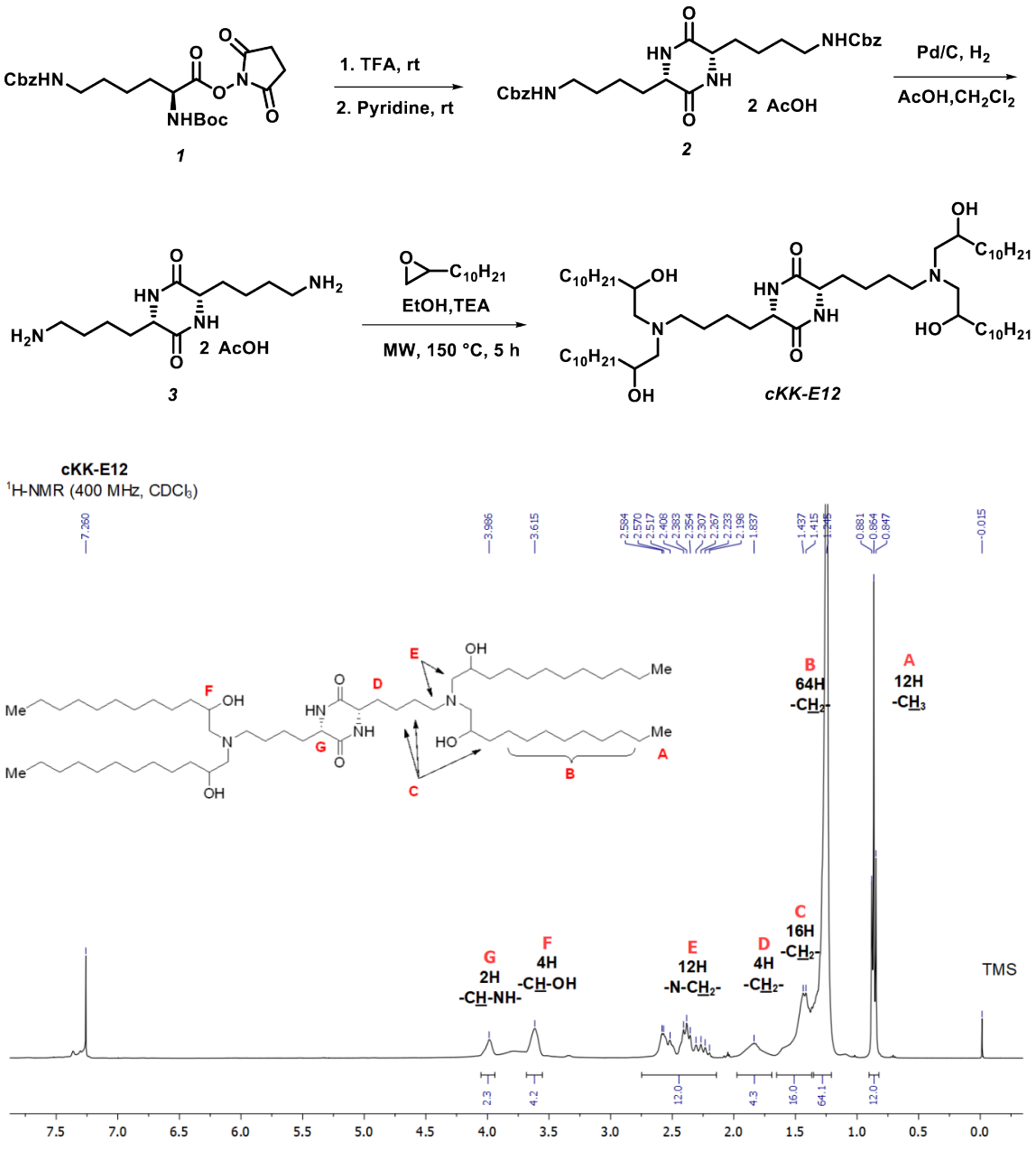
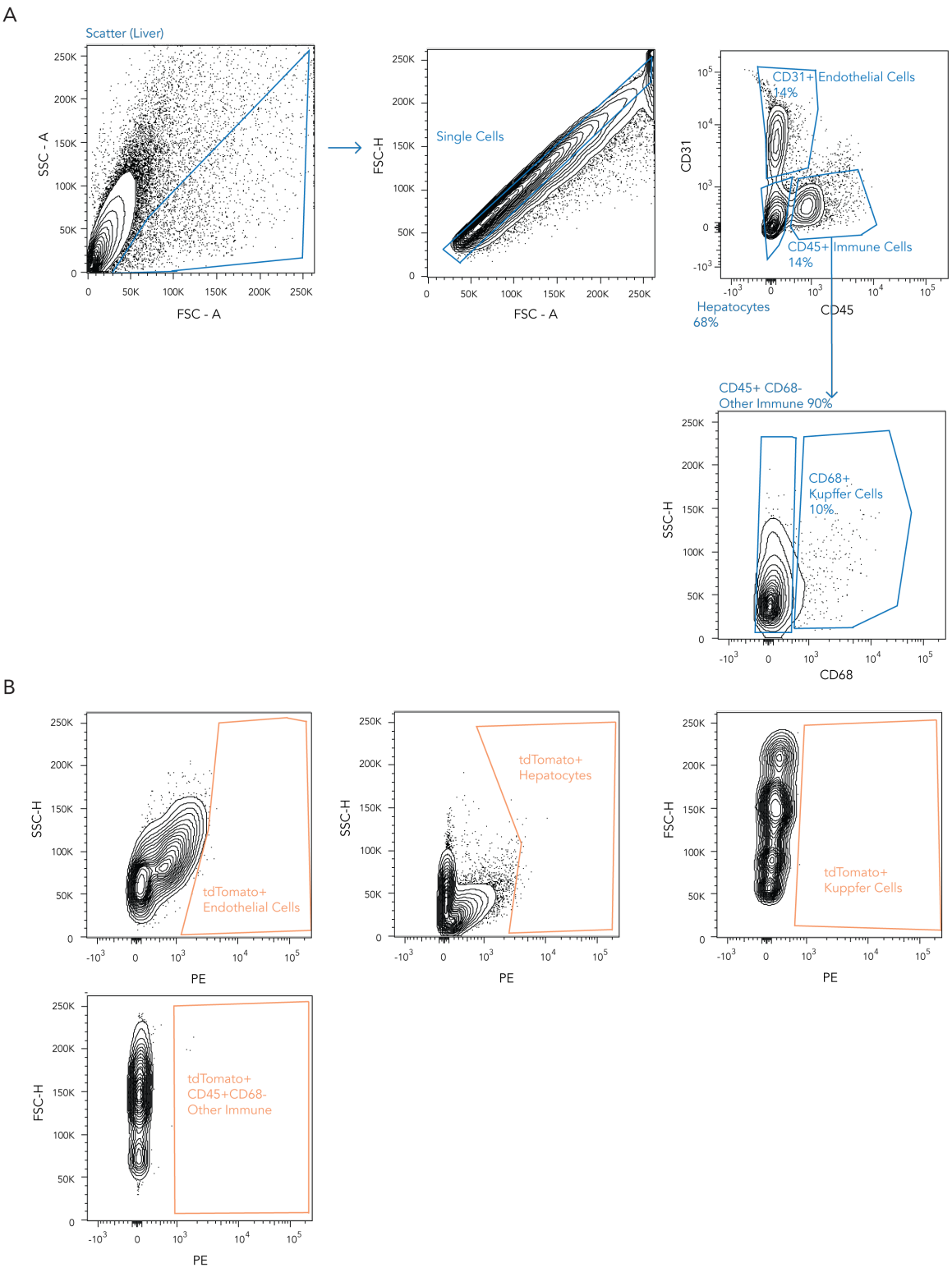
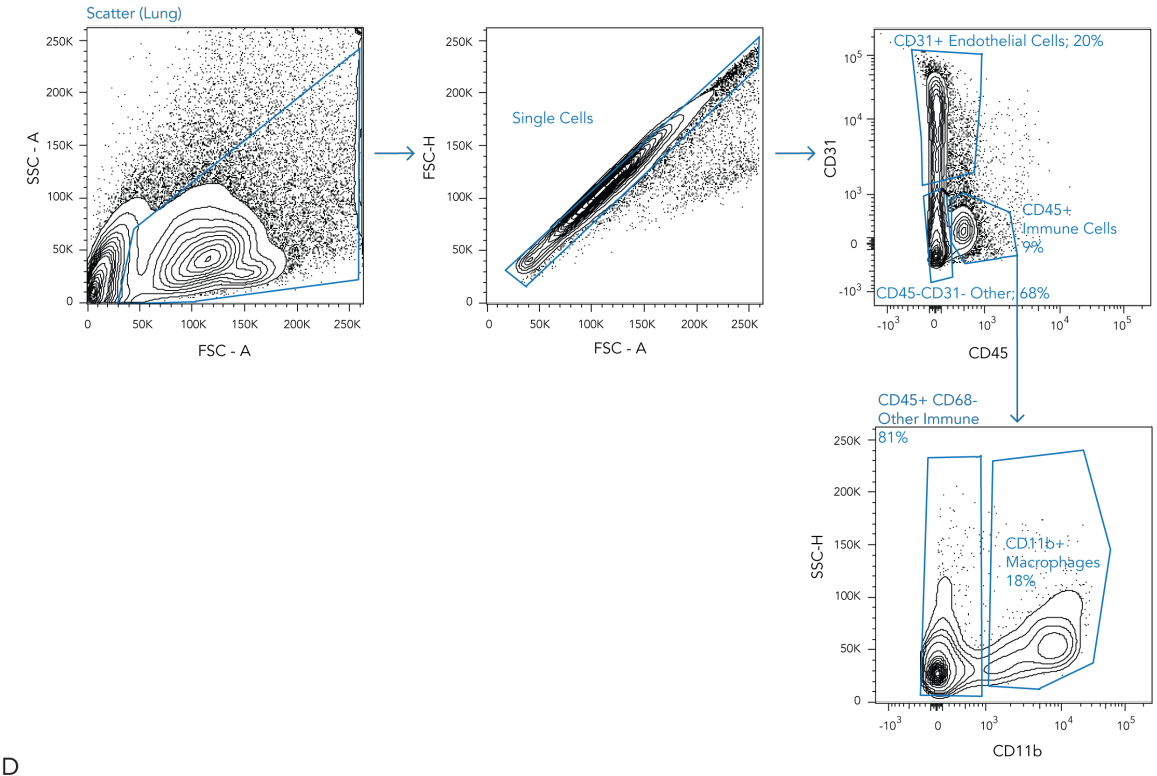


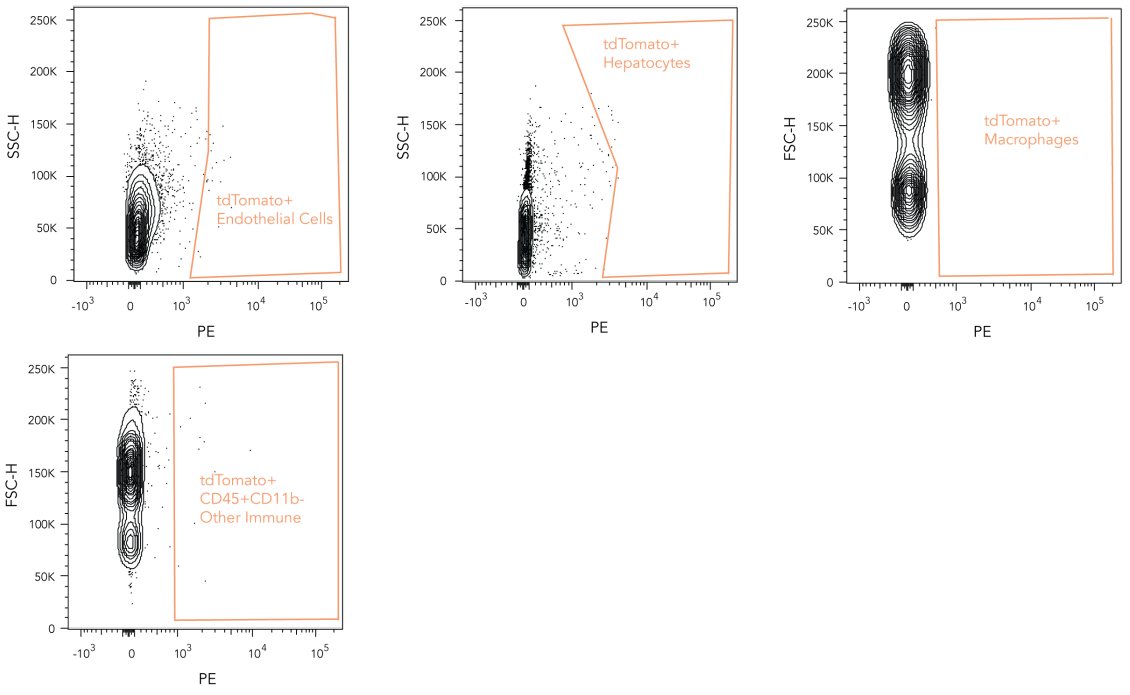
Figure 5.13 Chemical synthesis of CKK-E12. (A) CKK-E12 synthesis. (B) ^{1}H-NMR spectra confirming CKK-E12 structure.


A
Scatter (Liver)
SSC - A
FSC - A
Single Cells
FSC-H
FSC - A
CD31+ Endothelial Cells
14%
CD31
CD45+ Immune Cells
14%
Hepatocytes
68%
CD45
CD45+ CD68-
Other Immune 90%
CD68+
Kupffer Cells
10%
SSC-H
CD68
B
tdTomato+
Endothelial Cells
SSC-H
PE
tdTomato+
Hepatocytes
SSC-H
PE
tdTomato+
Kuppfer Cells
FSC-H
PE
tdTomato+
CD45+CD68-
Other Immune
FSC-H
PE

C


Scatter (Lung)
SSC - A
FSC - A
Single Cells
FSC-H
CD31+ Endothelial Cells; 20%
CD31
CD45+
Immune Cells
9%
CD45-CD31- Other; 68%
CD45
CD45+ CD68-
Other Immune
81%
CD11b+
Macrophages
18%
SSC-H
CD11b

D


tdTomato+
Endothelial Cells
SSC-H
PE
tdTomato+
Hepatocytes
tdTomato+
Macrophages
FSC-H
tdTomato+
CD45+CD11b-
Other Immune

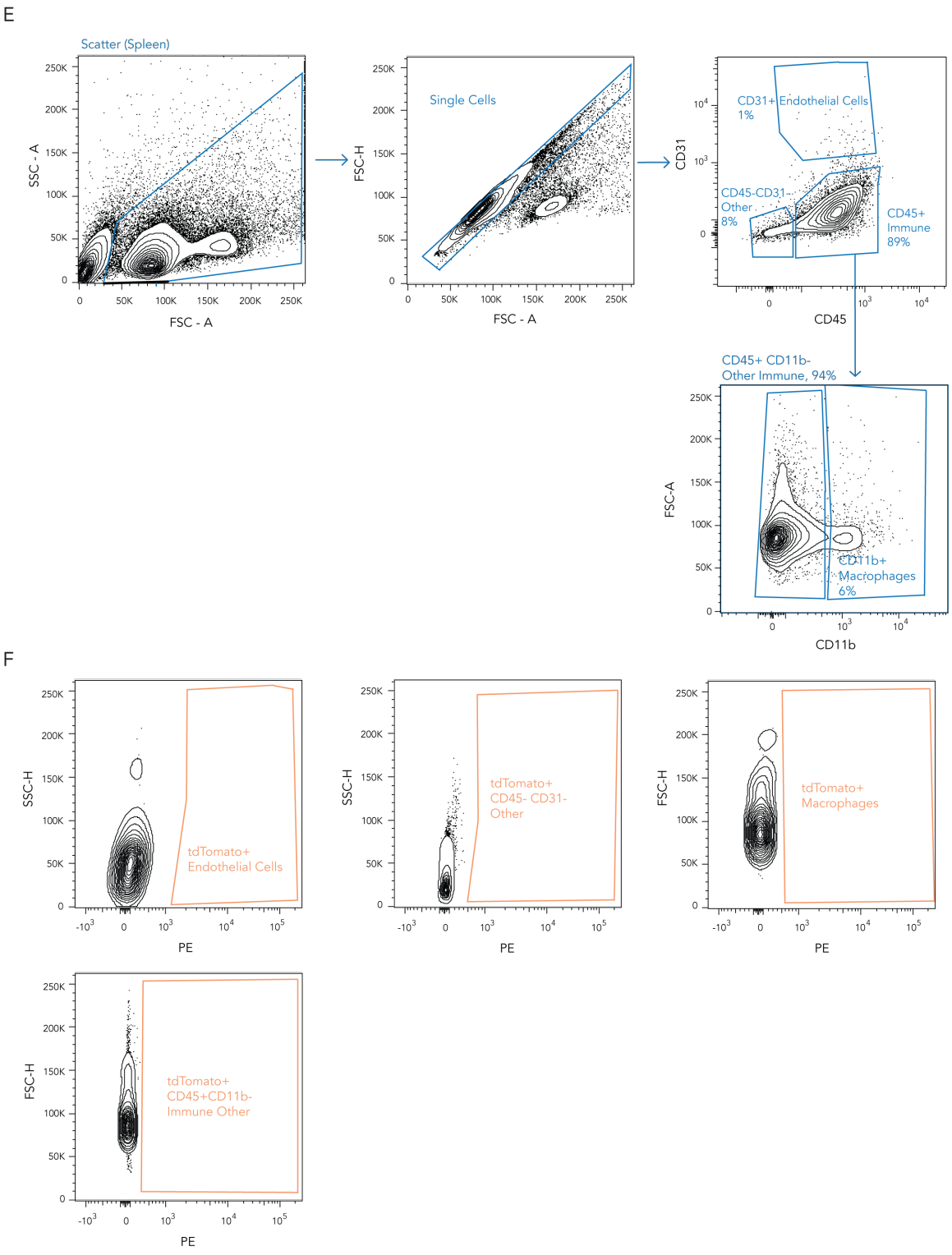


Figure 5.14 Representative *in vivo* gating strategies. Strategies for cell types in the (A, B) liver, (C, D) lung, and (E, F) spleen

CHAPTER 6. PERSPECTIVES AND FUTURE WORK

6.1 Need for a universal high-throughput screening system

Current LNP barcoding strategies are limited in the extent to which biological questions can be asked. Despite the fact that we have developed *in vivo* screening systems that can identify LNPs that efficiently deliver siRNA[79] and mRNA[9] *in vivo*, both of these screening systems either (i) rely on specialized mice to screen in or (ii) only allow for screening in specific cell types. Current siRNA screening systems are optimized to look for knockdown in ICAM-2 or GFP expressing cells as opposed to all cell types. There is a need for a high-throughput functional delivery screening system that can be used with genetic knockout mice. Currently, if the field wants to understand the mechanism of action of a particular LNP, each LNP needs to be assessed for its functional activity in knockout mice. Creating a screening system that works in knockout mice will allow for screening of a pool of LNPs in both WT and knockout mice, allowing researchers to get a broader sense of an LNPs mechanism. It has previously been shown that some LNPs rely on lipoprotein-dependent mechanisms to access cells[84], increasing interest in a screening platform that will allow researchers to identify ApoE-dependent and ApoE-independent LNPs. While we previously tried to answer pertinent question about the impact of specific genes (e.g. VLDLR, LDLR) on nanoparticle delivery using high-throughput LNP barcoding, we were constrained by the system that being used to ask these questions.

To this end, we have begun search for a gene that is ubiquitously expressed in all cell types across all tissues **(Figure 6.1)**, for which we will optimize the siRNA sequence that targets that gene, and begin work to recapitulate our siRNA screening system *in vivo*

using this new gene. In order to validate siRNAs targeting candidate genes, we designed chemically modified siRNAs targeting these genes. There are a number of potential gene candidates for such a system such as CD47 (marker of self) or beta-2-microglobulin (B2M). Both of these proteins are ubiquitously expressed on the surface of all cells. However, finding a gene that, when knocked down, will not lead to the destruction of the cell is necessary for looking at knockdown long-term. It is unclear whether knocking down CD47 or B2M using siCD47 or siB2M will lead to cell destruction via an immune response[250].

6.2 The utility of new sequencing technologies in studying nanomedicine

The perspectives and future work depicted here is an excerpt from Paunovska K, Loughrey DA, Sago CD, Langer R, Dahlman JE (2019). "Using Large Datasets to Understand Nanotechnology" *Advanced Materials.*

6.2.1 NGS approaches enable single cell and multiomics analyses

A suite of technologies based on high-throughput NGS have been created and validated. All of these are driven by advances in sequencing-by-synthesis, which allows scientists to characterize millions of molecules at the same time. These omics techniques, referred to as "sequence census" methods, can examine the genome (DNA), transcriptome (RNA), and epigenome (DNA modifications). All exploit the fact that DNA sequences can function as a digital substrate that is easily counted[251].

These technologies have evolved rapidly. Soon after NGS was reported, scientists designed ways to sequence DNA[252], and later, RNA[253], from single cells. Advances in

single cell required specific advances in acquiring and analyzing data. In particular, when acquiring scRNA-seq data, it is important to understand 'drop-out', an effect wherein datasets contain many genes with no expression. By developing standardized methodologies to overcome drop-out, single cell techniques have enabled targeted RNA[254, 255] and whole transcriptome[256] analysis. By sequencing RNA from single cells, scientists improved their fundamental understanding in many fields of biology, examining everything from the diversity of microbial ecosystems to the intratumor heterogeneity and clonal evolution of human cancer[257, 258]. As an example, scRNA-seq studies have been used to differentiate subclasses of a given cell type (e.g. neurons[259-261], or immune cells[262-264]), or study heterogeneous cell responses to a given biological stimulus[265]. In one representative example, Villani *et al*. performed unbiased scRNA-seq on 2400 peripheral blood mononuclear cells. By analyzing the subsequent gene expression data, they identified new subtypes of dendritic cells and monocytes in human blood, enabling more accurate immune monitoring in disease[263].

In order to generate these single cell data, authors combined an experimental and computational strategy to identify discriminative surface markers in clusters of cells that were similar to each other. They isolated the cells using these markers and validated the identity of these inferred subtypes using scRNA-seq. In order to ensure the data were robust, the authors corroborated their findings by analyzing peripheral blood mononuclear cells from ten independent healthy individuals. Although scRNA-seq approaches are not frequently used to study cellular response to nanomaterials, we are optimistic this approach will be important to the nanomedicine field for 2 reasons. First, scRNA-seq is now easy to use. In fact, there is an ongoing effort called the 'Human Cell Atlas' that aims to perform

scRNA-sequencing on as many cell types as possible[266]. Second, in the papers cited above, authors found that a collection of cells thought to be homogenous exhibit a high degree of genetic and functional heterogeneity. These data suggest that gene expression and subsequent cell function, even within a given cell type, exists on a spectrum. These approaches could similarly reveal subtypes of immune cells that readily interact with nanomaterials. By studying the different gene expression profiles in immune cells that do (or do not) respond to nanotechnology, master regulatory genes that trigger immune responses to nanomaterials or promote effective endosomal release may be identified.

More recently, the integration of diverse platforms (multiomics) has begun. In these examples, large scale analysis of multiple biomolecules is performed[267-269]. One key aspect of multiomic data generation is the fact that scientists must (i) process cells and (ii) design sequencing pipelines that allows several datasets to be acquired. In one example, scientists measured the genomic copy-number variations, transcriptome and DNA methylome of 25 single cancer cells. The authors were able to acquire these multiomic data using a gentle lysis procedure that dismantled the cellular membrane of an individual cell while keeping the nucleus intact. This preserved nucleus was used as a substrate for single cell DNA methylomic analysis, while the cytoplasmic lysate was used to acquire transcriptomic information from the same cell. They identified two distinct subpopulations within these cells and showed the transcriptomic heterogeneity within each subpopulation[270] affected cell function. In another example, scientists used NGS to concurrently measure transcriptomic and epigenomic data, in order to evaluate the mechanisms of neurodegeneration in Alzheimer's disease, and how the environment and the genome act through different cell types[271]. Once again, the authors used a novel experimental approach

to acquire the data; more specifically, the authors performed in parallel chromatin immunoprecipitation and RNA sequencing on harvested mouse hippocampus. This allowed seven different epigenetic modifications that mark distinct functional chromatin states and the corresponding changes in gene expression to be analyzed simultaneously. By profiling transcriptional and chromatin state dynamics, they found that immune-cell-specific enhancer regions and response genes were more accessible to transcription factors, suggesting the pathogenic capacity of the immune system in Alzheimer's disease. A coordinated decrease in synaptic plasticity genes was also found, linking these multiomic readouts to a potential mechanism of disease progression.

The coupling of protein mass spectrometry to genomics, known as proteogenomics[272, 273], is another new class of technologies to generate multiomic datasets. Although mass spectrometry has analytical limitations[274], these are being addressed. To date, proteogenomics has been applied to traditional biological problems. For example, scientists characterized human colon and rectal cancer[272]; using proteogenomics, the authors identified 4 subtypes of diffuse gastric cancers, associated with proliferation, immune response, metabolism, and invasion, respectively[275]. However, through these studies, best practices have been established that provide a framework to characterize protein-nanomaterial interactions. Thus, proteogenomics has the potential to be applied to the protein corona and other interactions between nanomaterials and proteins. Although multiomics approaches have not – to date – been applied to nanomaterials, these techniques permit scientists to characterize complex cellular responses[276, 277]. It is therefore very likely that multiomics can help elucidate how cells respond to nanomaterials.

6.2.2 Transcriptomics can uncover how cells respond to nanoparticles

In contrast with multiomics, transcriptomics has already been used to interpret the complex effects that nanomaterials and biomaterials have on gene expression. There are a number of recent examples of the nanotechnology field taking advantage of transcriptomics, both *in vitro* and *in vivo*. Carrow *et al*. recently used RNA-seq to identify more than 4000 genes whose RNA expression changed when human mesenchymal stem cells (hMSCs) interacted with nanosilicates[278]. Notably, they found that particular signaling pathways were upregulated, including the stress-responsive and surface receptor-mediated MAPK pathways. The authors also characterized a number of biophysical and biochemical cellular behavior and found that nanosilicates promote stem cell osteochondral differentiation. In particular, by analyzing changes in genes that are part of biological pathways related to osteogenesis, researchers saw that hMSCs exposed to nanosilicates tended to favor bone and cartilage lineages. They found that genes such as cartilage oligomeric matrix protein, aggrecan, and collagen type I α1 chain were upregulated; these genes are characteristic of osteochondral differentiation. Taken together, these data suggest that proliferation and differentiation pathways were influenced by nanomaterials.

As another example, Feliu *et al*. utilized primary human bronchial epithelial cells to show that cationic dendrimers caused significant changes in gene expression, even at doses that did not lead to acute or overt signs of cytotoxicity[279]. After administering a dose of 0.1 μM PAMAM dendrimers – which translates to a dose of roughly 1.4 μg/mL *in vitro* – to these cells, they found that the expression of 203 genes changed. Interestingly, by performing gene ontology enrichment analysis, the authors found that many of these genes were part of pathways related to cell division and cell cycle regulation. The authors created network diagrams to visualize predicted impacts on downstream pathways after

upregulation and downregulation of specific genes. These results are important, given that many studies rely on overt assays to screen for nanoparticle toxicity. The results may also have implications for tumor-targeted nanoparticles, since tumor growth can be driven by aberrant cell division and cell cycle regulation. In another example, Lucafò *et al*. reported the interaction of fullerenes with human MCF7 tumor cells showing that they cause a time-dependent alteration of gene expression, arresting cell cycle progression and promoting the entry in G0 phase[280]. By performing whole-transcriptome RNA-seq analysis on cells exposed to fullerenes, the authors found that mTOR signaling, which regulates cell growth and proliferation, was inhibited while genes upstream of TGF-β, important for cell remodeling and adhesion, were upregulated – suggesting that nanoparticles can alter cell cycle regulation. In addition, Gliga *et al*. showed that cerium oxide nanoparticles negatively affect neuronal differentiation and interfere with cytoskeletal organization in the murine cell line C17.2, which can be used as a model for developmental neural stem cells[281]. Cerium oxide nanoparticles were known to show cytoprotective effects. However, by analyzing gene expression using RNA sequencing this study found that the expression of at least 795 genes changed over a 7-day period after C17.2 cells were exposed to nanoceria. Changes in gene expression were compared to changes elicited with a common antioxidant, N-acetylcysteine, and samarium-doped nanoceria, which has previously been shown to have lower antioxidant activity than nanoceria alone. Notably, the authors found that nanoceria inhibited neuronal stem cell differentiation extensively, compared to N-acetylcysteine and samarium-doped nanoceria, when they analyzed the genes that were changed, illustrating that antioxidant properties were not necessarily beneficial in all cases. In *Chlamydomonas reinhardtii,* a model organism, authors found that exposure to four

different commonly used metal nanoparticles – nano-Ag, nano-TiO_2, nano-ZnO, and CdTe/CdS quantum dots (QD) – had both similar and relatively distinct effects on the transcriptome. More specifically, Zn, QD and Ti based nanoparticles had upregulation and/or downregulation of similar genes, whereas Ag elicited an opposite transcriptional response in *Chlamydomonas reinhardtii* when compared to the other three nanoparticles. Notably, some of the changes included potential proteasome inhibition which could suggest interest as a cancer chemotherapy agent[282]. Also, in *C. reinhardtii*, Beauvais-Flück *et al*. showed that up to 4784 transcripts were dysregulated when exposed to subnanomolar methyl mercury even after two hours. Genes involved in cell motility, nutrition, and photosynthesis were among the main regulated transcripts highlighting the tolerance mechanisms for microalgae at sublethal methyl-mercury concentrations[283].

Finally, additional evidence that nanoparticles alter genome-wide gene expression has been found *in vivo*; engineered iron sulfide nanoparticles were shown to cause substantial gene expression alterations in pathways related to immune and inflammatory responses, detoxification, oxidative stress and DNA repair and damage, in adult zebrafish[284]. These results illustrate that major transcriptional changes can be tracked *in vivo* when an organism is exposed to a nanoparticle. These examples are complemented by evidence suggesting the composition, size, or shape of a biomaterial potentiates the cellular response to that material[285]. Studies that record the cellular response to biomaterials have been collated in the Compendium for Biomaterial Transcriptomics (cBiT)[286], a collection of transcriptional profiles of cells after biomaterial exposure; this resource will likely continue to become even more valuable as more data become available.

As demonstrated by the studies above, best practices for RNA-seq data generation and gene expression analysis are established[287]. The first step is to clearly define a biological question. One simple test case would be 'What RNAs are affected by a given nanomaterial and can the RNAs identify a specific cellular signaling cascade that responds to that nanomaterial?'. Second, extract the cellular RNA and convert it to a countable pool of complementary DNA (cDNA) via reverse transcription using polydT or random hexamers using standard kits. Third, sequence this pool of DNA using NGS. Fourth, perform quality control analyses on the data in order to statistically correct biases that arise during sample preparation or sequencing. Fifth, analyze the 'clean' data using an appropriate bioinformatics pipeline, thereby identifying genes with up- or down-regulated expression in response to the nanomaterial[287]. Sixth, use network analysis or cell ontology-based approaches to understand whether alterations in gene expression can identify cellular pathways altered by the nanomaterial. Finally, once pathways are identified, it is feasible to make predictions about how the nanomaterial will affect the cellular phenotype (cell growth, death, toxicity, etc.).

6.2.3 Methods to analyze large datasets appropriately

As the output from sequencing platforms reaches the order of terabytes (and billions of sequencing reads), it will be increasingly important to visualize and interpret the data related to biomaterials using best practices. Here we describe common issues faced when interpreting data sets of this size, as well as ways to ensure the data interpretation is appropriate[189, 287]. One important consideration when analyzing large data sets is dimensionality. For example, some transcriptomic studies can have 20,000 dimensions; each dimension is the expression of a gene. Given that visualizing data on 20,000 axes is

not feasible, datasets are reduced to a smaller number of dimensions so they can be visualized (**Fig. 6.2A**). High-dimensional objects are replotted in a low-dimensional map; individual objects are represented by a point, and objects that behave similarly are 'clustered' nearby. In addition to making data easier to interpret visually, reducing dimensionality can be used to identify important variables in a complex, multivariable experiment.

Dimensionality reduction is often performed using principle component analysis (PCA)[288, 289]. Put succinctly, PCA provides a statistical framework whereby the maximum amount of variance is captured with the lowest possible number of dimensions. In biological experiments, where there are usually many more observations than variables, the number of principle components (PCs) is the same as the number of variables. The PCs are sorted by their statistical importance. For example, suppose factors contributing to the cost of a car were studied by generating a dataset with the cost, size, brand, color, and number of wheels of different vehicles. Since all cars have 4 wheels, this variable will not contribute to the variance in car costs. However, the cost might matter, as might the brand, and these two factors co-vary. In this case, principal component 1 (PC1) would be the linear combination of variables that contributed the most amount to variance (e.g. PC1 = 4*cost + 2.4*brand + 1.1*size + 0.3*color + 0.001*num. wheels). In this linear combination, the number of wheels negligibly contributes to the variance, and is therefore unimportant. Then, after factoring in PC1, a second set of relationships can be seen, where (for example) the size and color might co-vary: (e.g. PC2 = 2*size + 1.2*color + 0.7*brand + 0.1*cost + 0.001*num. wheels). Every factor contributes to each PC, but only the factors that explain a lot of the variance and are correlated have high weights for the same PC. In the case of

studying nanomaterial-biological interactions, the factors may be the sets of genes that are up or down-regulated in response to a specific nanomaterial. One important limitation to PCA however is that relationships between variables are often non-linear. In addition, PCA is usually specific for each dataset, making it difficult to compare PCs across studies. As a result, when considering whether a nanomaterial dataset can be analyzed with PCA, it is best to consult an expert in data analysis.

Even with these nuances, PCA can still be used effectively to reduce dimensionality. In biological applications, applying PCA to data with N variables will generate N PCs; if the first PC is responsible for a large percentage of the variance, the dimensionality of the dataset can be reduced by excluding PCs with much smaller contributions. PCA is commonly applied to biological datasets in order to identify experimental conditions that drive variance in gene expression[289]; in a typical gene expression profiling experiment, the first 5 PCs drive up to 50% of the variance, while the remainder explain just one or two percent of the variance and can be ignored. As a result, although nuances in the data can be lost during dimensionality reduction, the general structure of the dataset is preserved. As PCA is applied to nanomaterials, experiments will need to be designed in order to maximize the number of repetitions per observation. Another useful feature of PCA is that once the PC is identified, it can help identify what drives similarities among samples, and remove unimportant sources of variation. Supervised and algorithmic options for analyzing these factors are widely used in transcriptomics[290, 291], and therefore, should be applied to nanomaterial datasets.

Currently, PCA is used in biology to answer questions related to (i) genetic differences between cell populations or (ii) gene importance when it comes to

understanding a cellular response to specific stimuli. This can be closely related to nano-bio interactions, which would replace a normal biological stimulus (e.g., a cancer drug) with a nanomaterial, thus allowing scientists to probe mechanisms behind these interactions. However, since PCA is easy to perform, it can be applied to datasets inappropriately[292]. For instance, PCA is typically not useful when (i) the variance is somewhat evenly distributed among the principle components, and (ii) the dataset is small and the number of variables and variance within the dataset is large. What constitutes an appropriately large nanomaterial dataset? As larger datasets are generated using nanotechnology, this question will need to be addressed. Once again, consulting with scientists who specialize in PCA will be important for nanomaterial labs. However, lessons from biological studies may help answer the question. It is generally accepted that biological studies with a large number of replicates can be analyzed with PCA, whereas studies with a small number of biological replicates (e.g., N=3 or fewer), and therefore, relatively high experimental variability, cannot. As a control, biological replicates should cluster together. The larger the number of variables being analyzed; the more technical and biological replicates are required to make statistically powered statements about data. For biological, and nano-related applications, biological replicates should be strongly correlated. Minimizing biological variance within an experiment is also crucial to correct analysis of data. For example, when analyzing nanoparticle delivery data, it will be necessary to separate cells that had ‘low’, ‘medium’, and ‘high’ levels of delivery, in order to obtain interpretable data. Given that the absolute values of low, medium, and high can vary with the type of drug being delivered, nanotechnologists will need to provide the rationale for their selection clearly. The advantages and limitations of PCA, as well as best

practices, have been reviewed in other fields[288, 289]. These best practices will be a useful starting point for nanotechnologists.

PCA is a dimensionality reduction technique that is mathematically designed to identify axes with maximum variance. However, in some cases, preserving small differences between similar objects is preferred[293]. For example, single-cell sequencing experiments regularly reveal heterogeneity amongst cells that were previously thought to be homogenous[294, 295], and often identify important rare cell sub-populations. For example, Shalek *et al*. found that the core antiviral response in pathogenically stimulated primary mouse bone-marrow-derived dendritic cells was coordinated by only a small proportion of the population[296]. In particular, the group found that only 0.8% of the 1700 sequenced cells exhibited antiviral gene expression very early, thereby leading to a larger response from the entire population. Given that immune cell subpopulations have been found in many other biological contexts, these approaches may be useful in overcoming three key limitations to nanomaterials. First, nanomaterials are cleared by circulating immune cells as well as immune cells within tissues. We find it likely that subsets of immune cells – driven by particular signaling pathways - respond more 'aggressively' to nanomaterials. Understanding these pathways could lead to pre-emptive, transient interventions designed to reduce nanoparticle toxicity. Second, nanoparticles can interact with cells via surface receptors. It is feasible that cell subpopulations express higher levels of a given surface receptor, thereby making it easier to specifically target that cell subtype. Third, since many nanoparticles enter cells via endocytic pathways, escaping the endosome is critical. It may be possible to identify cell subsets that are particularly amenable to drug delivery, simply due to the expression of genes related to endosomal escape. In order to identify cell

subpopulations with these phenotypes, the best practice would be to analyze single cells, measuring immunostimulation, biodistribution, or cytoplasmic release, and, at the same time, measuring the transcriptomic profile of the cell. In these experiments, it would be important to group cells so small differences between cell types are preserved. For such situations, algorithms like t-distributed stochastic neighbor embedding (t-SNE) are appropriate. T-SNE, first described by Maaten and Hinton in 2008[293], has allowed researchers to analyze cell heterogeneity in new ways[294, 295]. Algorithms to visualize t-SNE plots have been adapted for use in multiple languages, including R, python, and MATLAB, making the technique easy to use. Biological predictions made by t-SNE have also been validated using traditional biochemical techniques. For example, DroNc-seq, a method that combines single cell and single nuclei RNA sequencing, was used to identify distinct cell populations with t-SNE. These populations were then confirmed using immunohistochemistry and other methods[294]. t-SNE is useful as an alternative cell clustering and visualization tool when trying to understand cell response to nanomaterials.

Although t-SNE has generated validated predictions when used correctly, it can also be used to draw incorrect conclusions. t-SNE plots are generated using several input parameters, most notably perplexity and the number of iterations run[297]. Authors have shown that selecting incorrect input variables can lead to images that contain clusters when in fact no clusters exist[297] (**Fig. 6.2B**); these are analogous to false positives. Moreover, every time a t-SNE plot is generated, the plot changes slightly, since all t-SNE plots are stochastic[293]. As a result, although the general structure is preserved and has meaning, interpreting relationships between individual points on the plot is inappropriate since the position of each individual point varies each time the analysis is performed (**Fig. 6.2C**).

A second approach used to analyze large datasets is unbiased clustering. Unbiased clustering helps visualize experimental groups that performed more similarly to one another than they did to other groups, without losing any information. Since clustering algorithms rely on different mathematical assumptions, it is important that clustering is performed with the appropriate algorithm, and that altering the algorithm does not dramatically alter the clustering pattern[189]. The most common algorithms are hierarchical, centroid/partition (e.g. k-means), density-based (e.g. DBSCAN)[298], and self-organizing maps (SOMs)[299]. In k-means clustering, the user selects a k value based on the number of clusters that the data will be partitioned into. If the user expects there to be many clusters, a high k number is selected; if the user expects few clusters, a low k number is selected. The algorithm associates nearby values based on their means; as more values are associated, the mean of all the values becomes the new mean until k clusters are formed[300]. Conversely, DBSCAN clusters are based on how closely points pack together and outliers are determined based on their presence in low density regions[298]. When measuring how cellular mRNA expression changes with response to a drug (or a biomaterial), hierarchical clustering or SOMs are often used. The appropriateness of a given clustering algorithm depends on the size and complexity of the dataset, as well as the research question being asked[301, 302], and guides to select the correct clustering algorithm have been published[303, 304]. Using appropriate clustering algorithms when analyzing biomaterial data will be important. For example, if k-mean clustering is employed, how is the number of clusters selected? Scientists studying biomaterials can learn from examples in other fields[305]. Unbiased clustering has been utilized in order to analyze how cells cluster based on nanoparticle functional delivery as well as how nanoparticles cluster based on material

properties[8, 78]. Given enough of this type of data, these analyses could be instrumental for intuitively designing future generations of nanoparticles.

To help evaluate whether the data are suitable for a given clustering algorithm, validation algorithms have been developed. Validation algorithms are based on metrics that evaluate how tight data within a given cluster are, and what the distance between clusters is[306, 307]. Validation algorithms are often subdivided by the type of clustering they employ; these include compactness, separation, and connectedness[307]. For example, to validate k-nearest neighbor clustering, a validation algorithm was developed based on the following idea: if we take a data point from a cluster, its k-nearest neighbors should be in the same cluster[306]. Put simply, the k-nearest neighbor is determined by assigning a value to each object; the value is proportional to its distance from the object. Then, depending on the k constant, the objects are group based on closeness; when k = 1, the nearest neighbor is clustered with the object of interest.

6.2.4 Visualizing complex biological interactions

Additional techniques are then required to visualize large datasets. Two common methods of data visualization are network diagrams and heatmaps. Network diagrams integrate data from many sources to model interactions within a biological system. As an example, scientists generate networks combining gene expression and other omics data[308]. Since looking at raw network diagrams can be challenging, they are simplified using algorithms that cluster the raw network[308]. This clustering utilizes gene expression data to quantify correlation values between genes. If the expression of A and B always change in the same direction, the algorithm tends to cluster them together. Given that even these

clustered networks can be difficult to interpret, manual editing of the network diagram can be employed to emphasize a specific component of the biological pathway. Alternatively, the gene expression may be overlaid on validated pathways using the Kyoto encyclopedia of genes and genomes (KEGG)[309-311] or the gene ontology consortium[312, 313]. These network diagrams – which are visual and qualitative – are also often augmented by including quantitative metrics derived from the dataset. As an example, information from gene or protein expression profiles can be included in network diagrams by making over or under-expressed genes/proteins stand out on the network. A common tool for creating integrated network diagrams is Cytoscape[314].

One related question that will need to be addressed as network analyses are used to understand biomaterial/cell interactions is the extent to which subtle biological interactions matter. In some cases, studying single genes will suffice. For example, the gene ApoE was shown to dramatically impact the delivery of a lipid nanoparticle *in vivo;* with ApoE, the nanoparticle was effective, and without it, the nanoparticle stopped functioning entirely[84]. However, it is likely that most nanoparticle-biological interactions will be driven by many genes interacting with one another. In the cases where many genes influence delivery, network analysis could focus on interactions between genes involved in endocytosis, metabolism, or intra/intercellular transport. To understand how many genes work in concert, network diagrams can be used to show interactions between hundreds or thousands of genes in a more unbiased way. Once these interactions are identified, scientists can evaluate whether the individual interactions are synergistic, additive, or antagonistic. If two genes interact synergistically, their effect on a phenotype is greater than the sum of each gene's individual impact. If they interact antagonistically, their effect is less than if they

were additive. Importantly, it is possible to evaluate how single genes and collections of genes can synergize or antagonize one another in a biological pathway[315].

Like network diagrams, heatmaps can be used to qualitatively highlight regions of interest in multivariate data. For example, gene expression heatmaps can identify genes that have high and low expression profiles if they cluster. If a clear and broad pattern exists within a dataset, heatmaps can highlight that pattern. Heatmaps are regularly used to supplement biological analyses. As an example, Subramanian *et al*. used hierarchical clustering to compare how 6 human cancer cell types clustered when analyzed using their profiling method, L1000, compared to Affymetrix and Illumina microarrays, and NGS-based RNA-seq, showing that each cell type clustered with itself independent of the sequencing/profiling system used[316]. They also analyzed 3333 drugs and 2418 additional compounds and showed that many of the drugs had potential off-target effects and potentially acted on multiple pathways. Honing in on the histone deacetylase (HDAC) superfamily of proteins, they were able to cluster inhibitors based on their selectivity for 13 different HDAC proteins[316]. Similarly, Hughes *et al*. assessed the effects of 300 different mutations and chemical treatments on *S. cerevisiae* and used hierarchical clustering to show that subtle changes in expression profiles can be tolerated and studied[317]. This is especially useful when looking at the effects of knocking out uncharacterized genes on a variety of cell processes. Heatmap analysis of sequencing data can be useful for identifying how a gene's expression changes over time in response to a biomaterial, and has been used to identify nanoparticles that efficiently deliver drugs[7, 78], identify cell types that are targeted by similar nanoparticles[78], and to identify nanoparticle chemical properties that tend to promote *in vivo* delivery.

Like other big data tools, it is important to ensure heatmaps are interpreted correctly. As an example, heatmaps use color to denote differences between samples; but the same color looks different when placed next to different colors[318] (**Fig. 6.3A**). In addition, data can be scaled by row or by column – this decision is dictated by what differences are being emphasized within a dataset. For example, a test dataset may have 'cell types' as column labels and 'genes' as row labels. The scaling method will dictate whether differences in the expression of one gene throughout multiple cell types (scaling by row), or differences in multiple genes' expression throughout one cell type (scaling by column) is emphasized. Attempting to qualitatively interpret data between rows if scaling colors by row or between columns when scaling colors by column would be incorrect – the colors may appear similar, but the absolute values would differ (**Fig. 6.3B**). Similarly, if the dataset has many more dimensions in one variable (e.g., genes) than another variable (e.g., cell types), it is best to cluster by the variable with fewer dimensions[189]. For example, if the expression of 20,000 genes is analyzed in 80 cell types, it is better to cluster by cell type first. Finally, data normalization (e.g. centering/scaling data around the mean, median, standard deviation (STD)) as well as the method used for clustering (e.g. Ward's, average, single, or complete) can change how the data cluster (**Fig 6.3C**). Finally, it is important to avoid dropping samples from the dataset, since this can have a large effect on how the rest of the samples cluster, as well as how the data is normalized. By understanding the limitations of over interpreting the color of a single box, running the data through more than one clustering algorithm (to ensure the clustering pattern does not dramatically change), analyzing the colors within the right 'direction' (i.e., column or row), and avoiding dropping data from the dataset, heatmaps can be generated that provide

compelling evidence of trends within complex biological systems; in many cases, these trends would be difficult to identify using other methods.

It is similarly important to understand the variance associated with your large data set; variance can be biological or technical. Biological variance is understood and can largely be mitigated by using a large number of replicates. Technical variance is still less well understood and can change with the experiment. As an example, reverse transcribing RNA can lead to bias that alters RNA-seq datasets[319]. Scientists also found that specific sequencing machines perform differently[320] and can generate bias[321]. There are simple ways to minimize variance. For example, including a sufficient number of biological replicates, and including appropriate positive and negative controls. One additional control that is important to consider when analyzing many biomolecules at once is the 'input'. For example, if you administer a pool of DNA-barcoded cells to an animal, it is important to sequence that 'input' pool, so you can normalize your output appropriately. Finally, any hits identified with any initial high-throughput screen should be independently validated using a tool like quantitative PCR, although previous studies have shown high correlation between the two techniques[322-324].

6.2.5 Future Perspectives

High-throughput data generation and analysis is not without difficulties, but this does not downplay its potential impact on nanomedicine. Recent clinical results using nanomedicine are cause for great excitement; these advances can be furthered using sequencing technologies. For example, nanoparticles carrying small molecules have been safely administered to patients[100], and siRNA delivered to hepatocytes by GalNAc

conjugates[101] or lipid nanoparticles[1] have treated genetic disease. At the same time, the need for systemically administered nanomedicines that target non-hepatocytes is significant, since most systemically administered drug delivery systems are still sequestered in the liver. The need for drug delivery is also growing. Traditional small molecule therapies have been joined by drugs based on proteins, siRNA, miRNA, mRNA, lncRNA, ASOs, ZFNs, TALENs, and CRISPR-Cas proteins. Each class of drugs will present numerous opportunities for nanotechnologists; as an example, the nanoparticle formulation that delivers a Cas9 mRNA is unlikely to be the best nanoparticle formulation for a Cas9 ribonucleoprotein. One additional example is whether the design rules for nanomedicines delivering one drug class (e.g., small molecules or proteins) will pertain to nanomedicines delivering another drug class (e.g., siRNA or mRNA). On one hand, it is possible to foresee a gene acting as a semi-master regulator of drug delivery. On the other, the biological response to nanoparticles containing proteins may be entirely different than the biological response to nanoparticles containing nucleic acids.

Using NGS, scientists can now quantify how thousands of nanoparticles target cells directly *in vivo* by formulating nanoparticle to carry rationally designed 'DNA barcodes'[8, 9, 72, 78, 79, 206, 207, 325] (**Fig. 6.4A**). In a separate example, scientists have used non-NGS forms of DNA analysis to perform high throughput *in vivo* assays of chemotherapy delivery[162] (**Fig. 6.4B**). These high throughput *in vivo* studies may eventually relate nanomaterial structure to *in vivo* delivery. However, future advances still need to be made, particularly in the ability to perform multivariate analysis on these large datasets. For example, when one of the components making up the nanoparticle is varied (e.g. PEG), interpreting causality in the dataset is difficult. If two nanoparticles with varying PEG molar ratios are

tested, and the nanoparticle with high molar percentages of PEG performs well, is it due to increased PEG, or decreased non-PEG components? PCA, t-SNE and other dimensionality reduction techniques are equipped for complex analyses like this. If this high throughput *in vivo* approach is coupled to an improved mathematical framework that permits scientists to understand how multivariate changes in nanoparticle structure alter delivery, nanoparticles with improved traits can be designed. For example, one key limitation in nanoparticle delivery is the unwanted clearance by immune cells, particularly in the liver and spleen. By quantifying how thousands of chemically distinct nanoparticles deliver drugs to on-target cells as well as these off-target cells, scientists may be able to 'evolve' nanoparticles that interact with clearance organs less frequently.

One way in which sequencing technologies may improve nanomedicine is by making the pre-clinical 'pipeline' used to discover nanoparticles more efficient. For example, the standard in the field is to synthesize chemically distinct nanoparticles, screen them *in vitro,* and select a small number of compounds for *in vivo* studies. However, *in vitro* nanoparticle delivery can be a poor predictor of systemic *in vivo* nanoparticle delivery[8]. At the same time, certain *in vitro* systems that recapitulate organ physiology may predict *in vivo* delivery. We envision high throughput studies comparing *in vivo* delivery to organ-on-chip systems[158] using thousands of nanoparticles[206]. By statistically comparing how thousands of different nanoparticles behave, these studies could elucidate the engineering (or biological) variables that make organ-on-chip systems predictive of *in vivo* behavior. A second inefficiency in the nanoparticle discovery pipeline is the unknown relationship between nanoparticle delivery in a mouse, and nanoparticle delivery in a rat, pig, NHP, or human. A systematic study of small animal models designed to identify a

'gold standard' animal to predict delivery in large animals has not been reported; this would constitute a significant advance for the field. We anticipate these studies may reveal that a given nanomedicine behaves differently in different mouse strains. Mouse strain-specific delivery has been observed with a promising virus[326, 327] selected using a novel *in vivo* viral evolution based approach[327]. The correct pre-clinical animal model may also change with the desired tissue; as an example, compared to mice, ferrets are better models for human airborne viral transduction[328]. By testing thousands of nanomaterials *in vivo* and understanding how strain- and species-dependent biological factors influence delivery, these large datasets may help improve how well pre-clinical models predict delivery in humans.

Big datasets may also be useful for understanding how to design nanotechnologies. For example, a method for de novo protein design[329] was recently reported; using machine learning, Butterfield *et al*. created a large library of protein-based nanocages (**Fig. 6.4C**). By applying selection pressures, nanocages were evolved using a 'bottom-up' approach to carry their own mRNA genome. Specifically, the authors performed multiple rounds of selection to identify the important nucleocapsid features for enhanced genome packaging, nuclease protection, and circulation time *in vivo*, without compromising the architecture of the structure. This was the first reported case of a non-viral container that can encapsulate its own genome and evolve in a complex extracellular environment, with the synthetic systems serving to rival the best recombinant AAVs. Using a similar approach, scientists used computational modeling to design and evolve proteins with different functions, including dimerization[330] and decreased side effects in a pre-clinical tumor model[331]. In particular, the authors \ designed a variant of interleukin-2 (IL-2) that would bind its

receptor on the target cell (T cells) without binding off-target receptors. The authors found that by redesigning one of the four helices on native IL-2 protein, they could increase on-target binding to the IL-2 receptor βγc heterodimer, while decreasing off-target binding to IL-2Rα (CD25), thereby driving toxicity. By redesigning these motifs, the authors improved IL-2 efficacy in mouse models of melanoma and colon cancer. Using a different approach, Guerette *et al*. coupled transcriptomics and proteomics data to design and predict the behavior of biomimetic materials[332]. The authors were able to rapidly process structural and functional novel high-performance eco-friendly materials pertaining to embryo protection, predation and adhesion. For example, they engineered silk-like materials from squid sucker ring teeth proteins that exceed the mechanical properties of many natural and synthetic polymers. Of particular note, the authors found a structural protein, suckerin-39, that surrounds squid sucker ring teeth and has high homology to silk, which would not have been discovered without the use of a combinatorial approach.

More recently, a series of papers have generated large biomaterial datasets without using NGS. In one example, quantitative structure-property relationship was retrospectively performed on a dataset describing nanoparticle formation; using this analysis, the authors found specific molecular variables associated with the drugs encapsulated in the nanoparticles were predictive of nanoparticle formation. Interestingly, the variables were related directly with the electronic configuration of the atoms making up the drug. Using only the molecular structure information of drug compounds, the authors rationally designed nanoparticles that delivered chemotherapeutics to tumors in mice[333], exploiting caveolin-dependent nanoparticle endocytosis. Specifically, the authors explored a number of different nanomaterial groups (e.g. detergents, azo dyes, and

polyelectrolytes) and used their quantitative structure-nanoparticle assembly prediction model to predict, and then validate, whether 400 different hydrophobic drugs would formulate into nanoparticles. Taken together, these examples constitute an innovative approach to coupling computational techniques, experiments, and unbiased screens, in order to improve nanomaterial design. In a third example, Yamankurt et al. developed a high throughput method based on mass spectrometry to monitor how immune cells responded to spherical nucleic acid nanomedicines[334]. The authors designed a library of 960 nanomedicines, varying the nanoparticle core (e.g. cholesterol, phospholipid), oligonucleotide shell (e.g. phosphodiester or phosphorothioate backbone, and sequence), and peptide antigen (e.g. OVA or E7). Their high throughput cell toxicity assay led to several structure-function relationships. First, spherical nucleic acid nanomedicines elicit more immune activation than linear oligonucleotides, and linear oligonucleotide immune activation is dependent on what the oligonucleotide is conjugated to (e.g. cholesterol, DOPE) as well as its backbone. Notably, the authors used the data to 'train' a machine learning algorithm, in order to identify non-linear property interactions (e.g. if there are 5 different properties, what is the interdependent effect of each property on the other). This is important because it can be difficult to decouple the effect of one property on another in a high-throughput screen where lots of variables are being changed, thus making it challenging to predict the biological response to a nanomedicine. Most recently, Rath et al. released a pre-print describing VSEPRnet, a method by which the physical and chemical traits of biomolecules are encoded in a way that enables neural network algorithms to make predictions[335]. The authors used this approach to predict binding between small peptides and allele-specific MHC-Class-1 molecules.

One need in the emerging field of large datasets and nanomedicine is the development of selection pressures that can be used to isolate nanoparticles that have performed a desired function *in vivo*. In biological studies, selection pressures are often based on cell death/proliferation, or alternatively, on fluorescence of a reporter gene[336-338]. High throughput nanotechnology screens will require assays with their own robust selection pressures, including biodistribution, functional cytoplasmic delivery, nuclear delivery, immunogenicity, and others. These will all generate different readouts. For example, nanoparticle delivery can be classified as (i) non-functional biodistribution, and (ii) functional, cytoplasmic delivery. In (i) a nanomaterial adhered to a cell is not distinguished from one that gets endocytosed, degraded in a lysosome, or delivered to the cytoplasm. However, in (ii) a nanoparticle must reach the cytoplasm of a cell, which ensures that only cells functionally delivered to are analyzed. These nanomaterial selection pressures can then be sub-divided into (i) up, or (ii) down-screens. Cells functionally delivered to in an up-screen change from no signal to a strong 'on' signal, whereas cells functionally delivered to in a down-screen change from high signal to 'low' signal.

Finally, well-designed studies could help answer key questions pertaining to the biology of delivery. First, which molecules play a predominant role in delivery? Proteins and lipids affect delivery, but carbohydrates require further exploration. Second, is a nanoparticle's delivery more likely to be due to a small number of master regulatory genes, or many genes acting in concert? Third, do lncRNAs and epigenetic modifications alter the cellular response to nanoparticles? Given that these molecules regulate many biological phenotypes[339], we find it likely. Fourth, are there *in vitro* systems that efficiently recapitulate and predict *in vivo* delivery? Organ-on-chip systems may be poised to answer

important biological questions. Finally, is there a 'gold standard' animal that can be used to predict delivery in large animals? The translation from delivery in small animal models (e.g. mice, rats) to efficient delivery in large animals (e.g. pigs, NHPs, humans) is still largely unknown. The network analyses needed to answer these questions will be aided by multiomics. For example, sequencing technologies that concurrently measure mRNA expression and protein expression have been developed[340]. Multiomics analyses may also aid nanomedicines by improving the drugs nanomedicines are meant to deliver. For example, the efficacy of RNA therapies is strongly affected by chemical modifications to the RNA[341]. Transcriptomics can identify splicing patterns, as well as the frequency with which RNAs are affected by modifications. These modifications are known to affect maturation, folding, and metabolism[342-344] of mRNAs; understanding the relationship between modifications and RNA transport could lead to nucleic acid therapeutics with improved safety profiles.

The interface between materials, medicine, and high-throughput sequencing marks a significant opportunity for researchers. To take full advantage of novel technologies, nanotechnologists will need to understand molecular biology, data analysis, and data visualization. Currently, scientists who design nanoparticles do not typically work alongside scientists who study omics-sized data sets. One way to accelerate the marriage of omics and nanotechnology is to teach concepts like PCR, primer design, sequencing preparation, PCA, and biostatistics in standard engineering and chemistry curricula. Until that time, if a chemist, materials scientists, or nanomedicine scientist would like to initiate an omics-based experiment, it will be important to consider the following steps. First, identify the types of data that are necessary. Is it important to understand the transcriptomic

response, epigenetic response, proteomic response, or some combination thereof? Is it sufficient to collect these data from many cells, or is it important to measure single cells individually? Second, seek out statisticians and bioinformaticians, in order to design your experiment correctly. How many groups or experimental conditions should be analyzed? What type of data analysis and visualization will be required? What types of experimental and technical controls are needed in order to believe the results? Answering these five questions will not guarantee the experiment is a success, but it will improve the odds that the data can be interpreted. Scientists who embrace NGS and analytics will be positioning themselves at the forefront of innovative new approaches that could accelerate the development of new materials and broadly benefit precision medicine and human health.

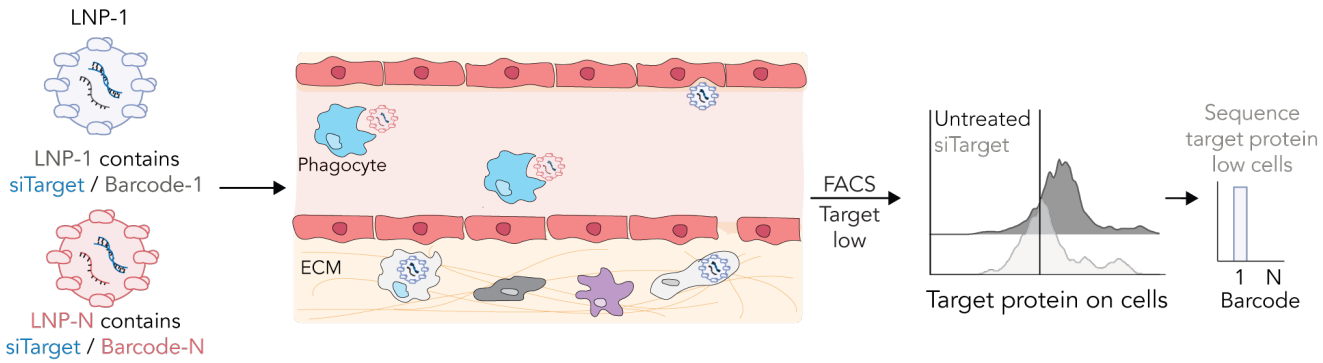


Figure 6.1 Development of a high-throughput screening system is necessary to be able to screen LNPs for functional delivery in knockout mice.

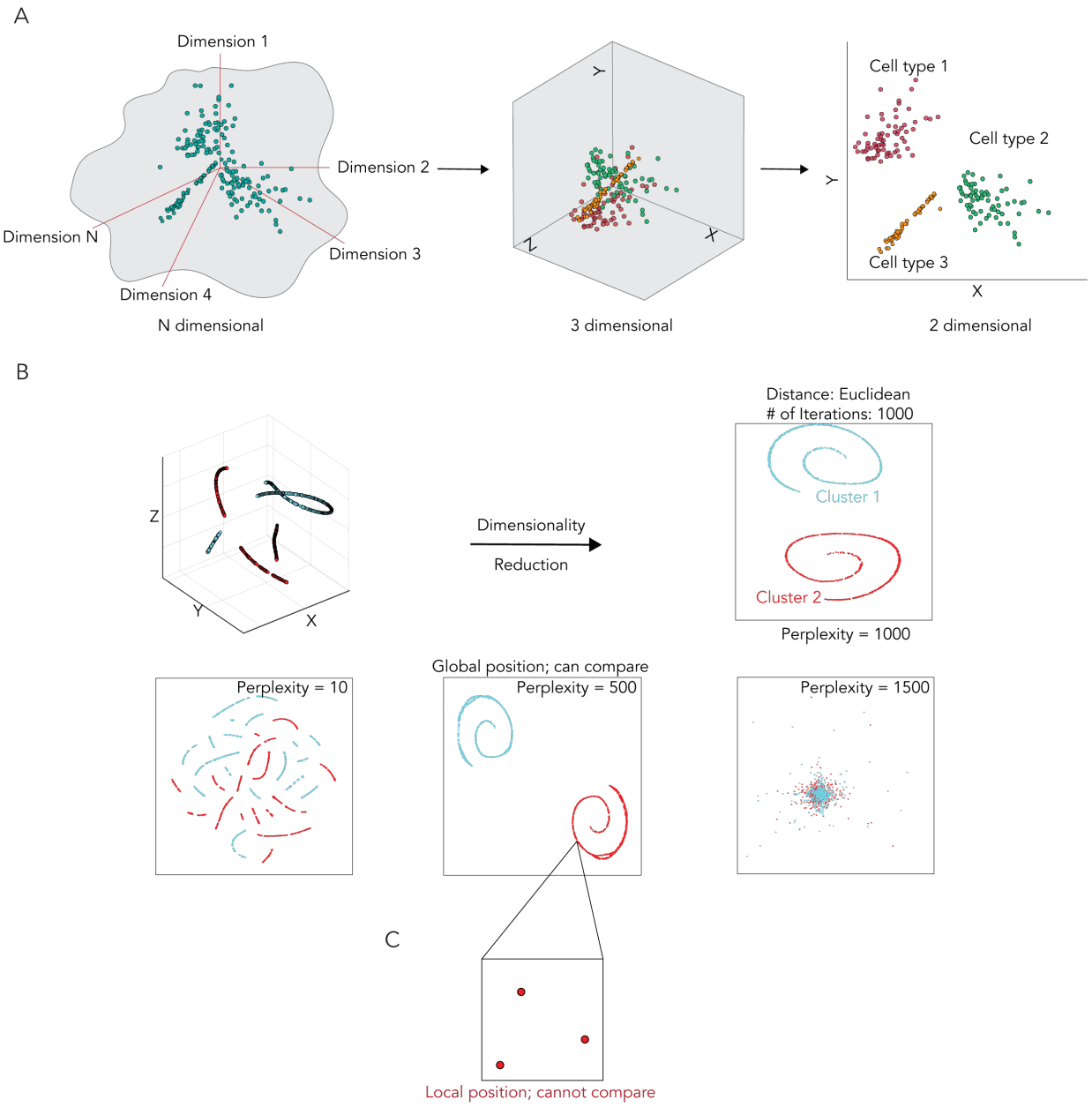


Figure 6.2 After generating large datasets, (A) data can be reduced to a smaller number of dimensions. This is done so data can be clearly visualized after identifying the most important variables in the experiment. (B) When reducing data dimensionality, selecting incorrect input variables can lead to images that contain clustered data when no clusters actually exist. In this example, varying the perplexity variable alters clustering. (C) Interpreting relationships between individual points in a t-SNE plot is not appropriate since the position of individual dots varies with each run of the analysis. Interpreting broad relationships from the data is appropriate.

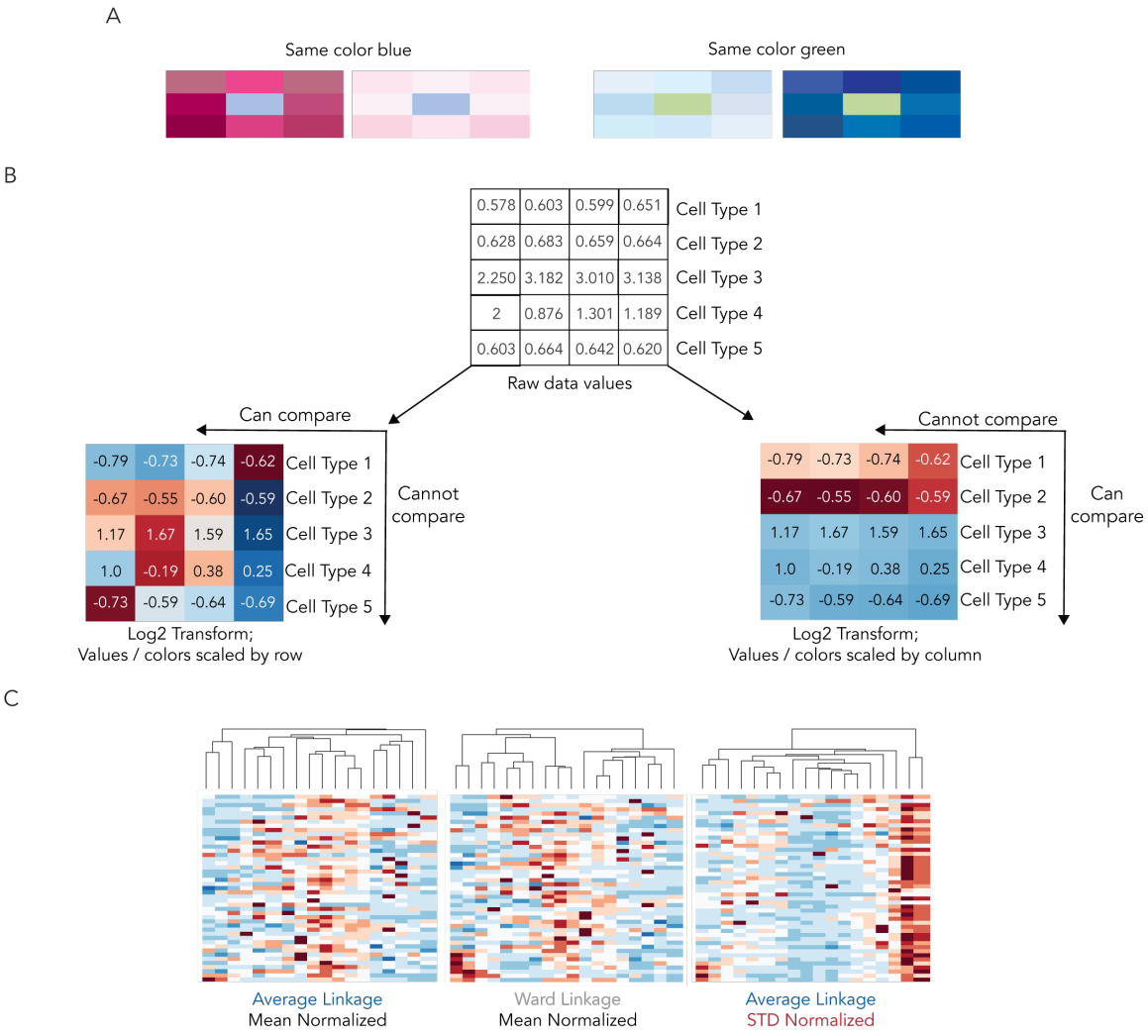


Figure 6.3 Heatmap generation and interpretation depends on the algorithms, conditions, and colors used. (A) The same color can look different when surrounded by different colors. (B) Heatmaps can be scaled by row or column. If scaling by row, colors can be compared within a row. If scaling by column, colors can be compared within the column. (D) Dendrogram clusters vary as a function of the normalization method and clustering algorithms.

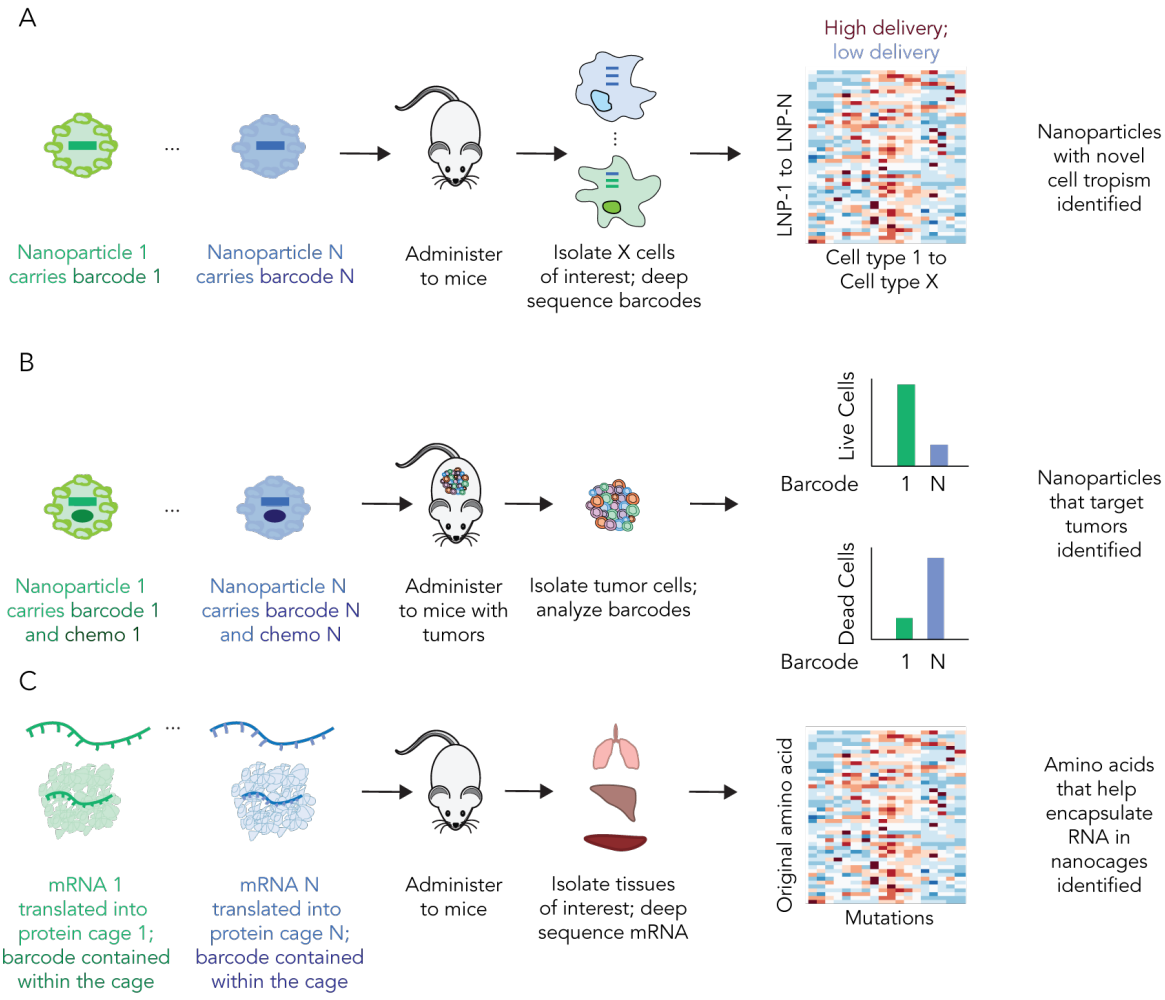


Figure 6.4 High throughput in vivo assays have been used to study nanomedicines. A) Nanoparticles were formulated to carry DNA barcodes. Nanoparticle 1, with chemical structure 1, was made to carry DNA barcode 1; nanoparticle N, with chemical structure N, was made to carry DNA barcode N. All N nanoparticles were administered to mice, cells of interest were isolated, and next-generation sequencing was using to quantify delivery of all N nanoparticles simultaneously. B) Liposome 1 was formulated to carry DNA barcode 1 and a chemotherapy; liposome N was formulated to carry DNA barcode N and a chemotherapy. Tumor delivery was quantified by measuring live/dead cells isolated from the tumor. C) Nanocages consisting of a different protein shell were encoded with mRNAs. The protein nanocages were administered to mice, and the effective nanocages were isolated from tissues. Sequencing was used to determine the mRNAs, and thus, by extension, the protein nanocages that survived in vivo.

REFERENCES

1. Adams, D. et al. Patisiran, an RNAi Therapeutic, for Hereditary Transthyretin Amyloidosis. *N Engl J Med* **379**, 11-21 (2018).

2. Ray, K.K. et al. Inclisiran in Patients at High Cardiovascular Risk with Elevated LDL Cholesterol. *N Engl J Med* **376**, 1430-1440 (2017).

3. Lorenzer, C., Dirin, M., Winkler, A.M., Baumann, V. & Winkler, J. Going beyond the liver: progress and challenges of targeted delivery of siRNA therapeutics. *J Control Release* **203**, 1-15 (2015).

4. Raesch, S.S. et al. Proteomic and lipidomic analysis of nanoparticle corona upon contact with lung surfactant reveals differences in protein, but not lipid composition. *ACS nano* **9**, 11872-11885 (2015).

5. Cedervall, T. et al. Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proc Natl Acad Sci U S A* **104**, 2050-2055 (2007).

6. Dowdy, S.F. Overcoming cellular barriers for RNA therapeutics. *Nature Biotechnology* **35**, 222-229 (2017).

7. Dahlman, J.E. et al. Barcoded nanoparticles for high throughput *in vivo* discovery of targeted therapeutics. *Proc Natl Acad Sci U S A* **114**, 2060-2065 (2017).

8. Paunovska, K. et al. A Direct Comparison of *in Vitro* and *in Vivo* Nucleic Acid Delivery Mediated by Hundreds of Nanoparticles Reveals a Weak Correlation. *Nano Letters* **18**, 2148-2157 (2018).

9. Sago, C.D. et al. High-throughput in vivo screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing. *Proceedings of the National Academy of Sciences* **115**, E9944-E9952 (2018).

10. Avery, O.T., MacLeod, C.M. & McCarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *The Journal of experimental medicine* **79**, 137-158 (1944).

11. Watson, J.D. & Crick, F.H. Molecular structure of nucleic acids. *Nature* **171**, 737-738 (1953).

12. Crick, F. Central dogma of molecular biology. *Nature* **227**, 561-563 (1970).

13. Ha, M. & Kim, V.N. Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology* **15**, 509-524 (2014).

14. Doench, J.G. & Sharp, P.A. Specificity of microRNA target selection in translational repression. *Genes & development* **18**, 504-511 (2004).

15. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233 (2009).

16. Grimson, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell* **27**, 91-105 (2007).

17. Peter, M. Targeting of mRNAs by multiple miRNAs: the next step. *Oncogene* **29**, 2161-2164 (2010).

18. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* **447**, 1130-1134 (2007).

19. Fitzgerald, K. et al. A Highly Durable RNAi Therapeutic Inhibitor of PCSK9. *N Engl J Med* **376**, 41-51 (2017).

20. Lam, J.K., Chow, M.Y., Zhang, Y. & Leung, S.W. siRNA versus miRNA as therapeutics for gene silencing. *Molecular Therapy-Nucleic Acids* **4**, e252 (2015).

21. Setten, R.L., Rossi, J.J. & Han, S.-p. The current state and future directions of RNAi-based therapeutics. *Nature Reviews Drug Discovery* **18**, 421-446 (2019).

22. Burnett, J.C. & Rossi, J.J. RNA-based therapeutics: current progress and future prospects. *Chemistry & biology* **19**, 60-71 (2012).

23. Sridharan, K. & Gogtay, N.J. Therapeutic nucleic acids: current clinical status. *British journal of clinical pharmacology* **82**, 659-672 (2016).

24. Levin, A.A. Treating disease at the RNA level with oligonucleotides. *New England Journal of Medicine* **380**, 57-70 (2019).

25. Finkel, R.S. et al. Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *The Lancet* **388**, 3017-3026 (2016).

26. Doudna, J.A. & Lorsch, J.R. Ribozyme catalysis: not different, just worse. *Nature structural & molecular biology* **12**, 395-402 (2005).

27. Mulhbacher, J., St-Pierre, P. & Lafontaine, D.A. Therapeutic applications of ribozymes and riboswitches. *Current opinion in pharmacology* **10**, 551-556 (2010).

28. Khan, A.U. Ribozyme: a clinical tool. *Clinica chimica acta* **367**, 20-27 (2006).

29. Bunka, D.H. & Stockley, P.G. Aptamers come of age–at last. *Nature Reviews Microbiology* **4**, 588-596 (2006).

30. Doudna, J.A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).

31. Platt, R.J. et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440-455 (2014).

32. Sander, J.D. & Joung, J.K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature biotechnology* **32**, 347 (2014).

33. Konermann, S. et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583-588 (2015).

34. Dahlman, J.E. et al. Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nature biotechnology* **33**, 1159 (2015).

35. Abudayyeh, O.O. et al. RNA targeting with CRISPR–Cas13. *Nature* **550**, 280-284 (2017).

36. Cox, D.B. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019-1027 (2017).

37. Liu, J.-J. et al. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* **566**, 218-223 (2019).

38. Chen, J.S. et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* **360**, 436-439 (2018).

39. Finn, J.D. et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent *In Vivo* Genome Editing. *Cell Reports* **22**, 2227-2235 (2018).

40. Nelson, C.E. et al. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403-407 (2016).

41. Nelson, C.E. et al. Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. *Nature medicine* **25**, 427-432 (2019).

42. Schwank, G. et al. Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell stem cell* **13**, 653-658 (2013).

43. Park, C.-Y. et al. Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. *Cell stem cell* **17**, 213-220 (2015).

44. Kleinstiver, B.P. et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495 (2016).

45. Charlesworth, C.T. et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nature medicine* **25**, 249-254 (2019).

46. Wagner, D.L. et al. High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nature medicine* **25**, 242-248 (2019).

47. Wolff, J.A. et al. Direct gene transfer into mouse muscle in vivo. *Science* **247**, 1465-1468 (1990).

48. Petsch, B. et al. Protective efficacy of in vitro synthesized, specific mRNA vaccines against influenza A virus infection. *Nature biotechnology* **30**, 1210 (2012).

49. Bogers, W.M. et al. Potent immune responses in rhesus macaques induced by nonviral delivery of a self-amplifying RNA vaccine expressing HIV type 1 envelope with a cationic nanoemulsion. *The Journal of infectious diseases* **211**, 947-955 (2015).

50. Pardi, N. et al. Administration of nucleoside-modified mRNA encoding broadly neutralizing antibody protects humanized mice from HIV-1 challenge. *Nature communications* **8**, 1-8 (2017).

51. Richner, J.M. et al. Modified mRNA vaccines protect against Zika virus infection. *Cell* **168**, 1114-1125. e1110 (2017).

52. Kowalski, P.S., Rudra, A., Miao, L. & Anderson, D.G. Delivering the messenger: advances in technologies for therapeutic mRNA delivery. *Molecular Therapy* **27**, 710 (2019).

53. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology* **31**, 822-826 (2013).

54. Presnyak, V. et al. Codon optimality is a major determinant of mRNA stability. *Cell* **160**, 1111-1124 (2015).

55. Kanasty, R., Dorkin, J.R., Vegas, A. & Anderson, D. Delivery materials for siRNA therapeutics. *Nature materials* **12**, 967-977 (2013).

56. Karikó, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23**, 165-175 (2005).

57. Khvorova, A. & Watts, J.K. The chemical evolution of oligonucleotide therapies of clinical utility. *Nature biotechnology* **35**, 238 (2017).

58. Eckstein, F. Phosphorothioates, essential components of therapeutic oligonucleotides. *Nucleic acid therapeutics* **24**, 374-387 (2014).

59. Putney, S.D., Benkovic, S.J. & Schimmel, P.R. A DNA fragment with an alpha-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. *Proceedings of the National Academy of Sciences* **78**, 7350-7354 (1981).

60. Bartlett, D.W. & Davis, M.E. Effect of siRNA nuclease stability on the in vitro and in vivo kinetics of siRNA-mediated gene silencing. *Biotechnology and bioengineering* **97**, 909-921 (2007).

61. Layzer, J.M. et al. In vivo activity of nuclease-resistant siRNAs. *Rna* **10**, 766-771 (2004).

62. Prakash, T.P. et al. Synergistic effect of phosphorothioate, 5′-vinylphosphonate and GalNAc modifications for enhancing activity of synthetic siRNA. *Bioorganic & medicinal chemistry letters* **26**, 2817-2820 (2016).

63. Mäkilä, J. et al. Synthesis of multi-galactose-conjugated 2′-O-methyl oligoribonucleotides and their in vivo imaging with positron emission tomography. *Bioorganic & medicinal chemistry* **22**, 6806-6813 (2014).

64. Akinc, A. et al. A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. *Nat Biotechnol* **26**, 561-569 (2008).

65. Semple, S.C. et al. Rational design of cationic lipids for siRNA delivery. *Nat Biotechnol* **28**, 172-176 (2010).

66. Zimmermann, T.S. et al. RNAi-mediated gene silencing in non-human primates. *Nature* **441**, 111-114 (2006).

67. Alabi, C.A. et al. Multiparametric approach for the evaluation of lipid nanoparticles for siRNA delivery. *Proc Natl Acad Sci U S A* **110**, 12881-12886 (2013).

68. Love, K.T. et al. Lipid-like materials for low-dose, *in vivo* gene silencing. *Proc Natl Acad Sci U S A* **107**, 1864-1869 (2010).

69. Mui, B.L. et al. Influence of Polyethylene Glycol Lipid Desorption Rates on Pharmacokinetics and Pharmacodynamics of siRNA Lipid Nanoparticles. *Mol. Ther.—Nucleic Acids* **2**, e139 (2013).

70. Akinc, A. et al. Development of lipidoid–siRNA formulations for systemic delivery to the liver. *Molecular Therapy* **17**, 872-879 (2009).

71. Dong, Y. et al. Lipopeptide nanoparticles for potent and selective siRNA delivery in rodents and nonhuman primates. *Proc Natl Acad Sci U S A* **111**, 3955-3960 (2014).

72. Paunovska, K. et al. Nanoparticles Containing Oxidized Cholesterol Deliver mRNA to the Liver Microenvironment at Clinically Relevant Doses. *Advanced Materials*, 1807748 (2019).

73. Lokugamage, M.P. et al. Mild Innate Immune Activation Overrides Efficient Nanoparticle-Mediated RNA Delivery. *Advanced Materials* **32**, 1904905 (2020).

74. Fenton, O.S. et al. Bioinspired alkenyl amino alcohol ionizable lipid materials for highly potent in vivo mRNA delivery. *Advanced materials* **28**, 2939-2943 (2016).

75. Semple, S.C. et al. Rational design of cationic lipids for siRNA delivery. *Nature biotechnology* **28**, 172-176 (2010).

76. Dahlman, J.E. et al. *In vivo* endothelial siRNA delivery using polymeric nanoparticles with low molecular weight. *Nature Nanotechnology* **9**, 648-655 (2014).

77. Lokugamage, M.P., Sago, C.D., Gan, Z., Krupczak, B.R. & Dahlman, J.E. Constrained nanoparticles deliver siRNA and sgRNA to T cells in vivo without targeting ligands. *Advanced Materials* **31**, 1902251 (2019).

78. Paunovska, K. et al. Analyzing 2000 in Vivo Drug Delivery Data Points Reveals Cholesterol Structure Impacts Nanoparticle Delivery. *ACS Nano* **12**, 8341-8349 (2018).

79. Sago, C.D. et al. Nanoparticles that deliver RNA to bone marrow identified by in vivo directed evolution. *Journal of the American Chemical Society* **140**, 17095-17105 (2018).

80. Fenton, O.S. et al. Synthesis and biological evaluation of ionizable lipid materials for the in vivo delivery of messenger RNA to B lymphocytes. *Advanced Materials* **29**, 1606944 (2017).

81. Kauffman, K.J. et al. Rapid, Single-Cell Analysis and Discovery of Vectored mRNA Transfection In Vivo with a loxP-Flanked tdTomato Reporter Mouse. *Mol Ther Nucleic Acids* **10**, 55-63 (2018).

82. Cheng, Q. et al. Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR-Cas gene editing. *Nat Nanotechnol* **15**, 313-320 (2020).

83. Kulkarni, J.A., Cullis, P.R. & van der Meel, R. Lipid Nanoparticles Enabling Gene Therapies: From Concepts to Clinical Utility. *Nucleic Acid Therapeutics* **28**, 146-157 (2018).

84. Akinc, A. et al. Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms. *Mol Ther* **18**, 1357-1364 (2010).

85. Cheng, X. & Lee, R.J. The role of helper lipids in lipid nanoparticles (LNPs) designed for oligonucleotide delivery. *Advanced Drug Delivery Reviews* **99**, 129-137 (2016).

86. Kumar, V. et al. Shielding of Lipid Nanoparticles for siRNA Delivery: Impact on Physicochemical Properties, Cytokine Induction, and Efficacy. *Mol Ther Nucleic Acids* **3**, e210 (2014).

87. Chen, S. et al. Development of lipid nanoparticle formulations of siRNA for hepatocyte gene silencing following subcutaneous administration. *Journal of Controlled Release* **196**, 106-112 (2014).

88. Gilleron, J. et al. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat Biotechnol* **31**, 638-646 (2013).

89. Pillay, S. et al. An essential receptor for adeno-associated virus infection. *Nature* **530**, 108-112 (2016).

90. Summerford, C. & Samulski, R.J. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *Journal of virology* **72**, 1438-1445 (1998).

91. Nonnenmacher, M. & Weber, T. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell host & microbe* **10**, 563-576 (2011).

92. Nonnenmacher, M. & Weber, T. Intracellular transport of recombinant adeno-associated virus vectors. *Gene therapy* **19**, 649-658 (2012).

93. Sun, J., Anand-Jawa, V., Chatterjee, S. & Wong, K. Immune responses to adeno-associated virus and its recombinant vectors. *Gene therapy* **10**, 964-976 (2003).

94. Mendell, J.R. et al. Single-dose gene-replacement therapy for spinal muscular atrophy. *New England Journal of Medicine* **377**, 1713-1722 (2017).

95. Russell, S. et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *The Lancet* **390**, 849-860 (2017).

96. Duan, D. Systemic AAV micro-dystrophin gene therapy for Duchenne muscular dystrophy. *Molecular Therapy* **26**, 2337-2356 (2018).

97. Bowles, D.E. et al. Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Molecular Therapy* **20**, 443-455 (2012).

98. Cheng, C.J., Tietjen, G.T., Saucier-Sawyer, J.K. & Saltzman, W.M. A holistic approach to targeting disease with polymeric nanoparticles. *Nature Reviews: Drug Discovery* **14**, 239-247 (2015).

99. Tsoi, K.M. et al. Mechanism of hard-nanomaterial clearance by the liver. *Nat Mater* **15**, 1212-1221 (2016).

100. Caster, J.M., Patel, A.N., Zhang, T. & Wang, A. Investigational nanomedicines in 2016: a review of nanotherapeutics currently undergoing clinical trials. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **9** (2017).

101. Pasi, K.J. et al. Targeting of Antithrombin in Hemophilia A or B with RNAi Therapy. *N Engl J Med* **377**, 819-828 (2017).

102. Bahl, K. et al. Preclinical and Clinical Demonstration of Immunogenicity by mRNA Vaccines against H10N8 and H7N9 Influenza Viruses. *Molecular Therapy* (2017).

103. Ashton, S. et al. Aurora kinase inhibitor nanoparticles target tumors with favorable therapeutic index in vivo. *Sci Transl Med* **8**, 325ra317 (2016).

104. Nel, A.E. et al. Understanding biophysicochemical interactions at the nano–bio interface. *Nature materials* **8**, 543-557 (2009).

105. El-Sayed, A. & Harashima, H. Endocytosis of gene delivery vectors: from clathrin-dependent to lipid raft-mediated endocytosis. *Molecular Therapy* **21**, 1118-1130 (2013).

106. Wang, F. et al. The biomolecular corona is retained during nanoparticle uptake and protects the cells from the damage induced by cationic nanoparticles until degraded in the lysosomes. *Nanomedicine* **9**, 1159-1168 (2013).

107. Vilanova, O. et al. Understanding the kinetics of protein–nanoparticle corona formation. *ACS nano* **10**, 10842-10850 (2016).

108. Lundqvist, M. et al. The evolution of the protein corona around nanoparticles: a test study. *ACS nano* **5**, 7503-7509 (2011).

109. Albanese, A. et al. Secreted biomolecules alter the biological identity and cellular interactions of nanoparticles. (2014).

110. Schöttler, S. et al. Protein adsorption is required for stealth effect of poly (ethylene glycol)-and poly (phosphoester)-coated nanocarriers. *Nature nanotechnology* **11**, 372-377 (2016).

111. Salvati, A. et al. Transferrin-functionalized nanoparticles lose their targeting capabilities when a biomolecule corona adsorbs on the surface. *Nature nanotechnology* **8**, 137-143 (2013).

112. Verano-Braga, T. et al. Insights into the cellular response triggered by silver nanoparticles using quantitative proteomics. *ACS nano* **8**, 2161-2175 (2014).

113. Tenzer, S. et al. Rapid formation of plasma protein corona critically affects nanoparticle pathophysiology. *Nature nanotechnology* **8**, 772-781 (2013).

114. Blanco, E., Shen, H. & Ferrari, M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nature biotechnology* **33**, 941-951 (2015).

115. Wan, C., Allen, T. & Cullis, P. Lipid nanoparticle delivery systems for siRNA-based therapeutics. *Drug delivery and translational research* **4**, 74-83 (2014).

116. Zuckerman, J.E., Choi, C.H., Han, H. & Davis, M.E. Polycation-siRNA nanoparticles can disassemble at the kidney glomerular basement membrane. *Proc Natl Acad Sci U S A* **109**, 3137-3142 (2012).

117. Aird, W.C. Endothelial cell heterogeneity. *Critical care medicine* **31**, S221-S230 (2003).

118. Aird, W.C. Endothelial cell heterogeneity. *Cold Spring Harbor perspectives in medicine* **2**, a006429 (2012).

119. Aird, W.C. Phenotypic heterogeneity of the endothelium. *Circulation research* **100**, 174-190 (2007).

120. Aird, W.C. Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circulation research* **100**, 174-190 (2007).

121. Augustin, H.G. & Koh, G.Y. Organotypic vasculature: From descriptive heterogeneity to functional pathophysiology. *Science* **357**, eaal2379 (2017).

122. Kheirolomoom, A. et al. Multifunctional Nanoparticles Facilitate Molecular Targeting and miRNA Delivery to Inhibit Atherosclerosis in ApoE(-/-) Mice. *ACS Nano* **9**, 8885-8897 (2015).

123. Sahay, G. et al. Efficiency of siRNA delivery by lipid nanoparticles is limited by endocytic recycling. *Nat Biotechnol* **31**, 653-658 (2013).

124. Wittrup, A. et al. Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown. *Nature biotechnology* **33**, 870-876 (2015).

125. Patel, S. et al. Boosting Intracellular Delivery of Lipid Nanoparticle-Encapsulated mRNA. *Nano Lett* **17**, 5711-5718 (2017).

126. Rohner, N.A. & Thomas, S.N. Melanoma growth effects on molecular clearance from tumors and biodistribution into systemic tissues versus draining lymph nodes. *J Control Release* **223**, 99-108 (2016).

127. Bertrand, N. et al. Mechanistic understanding of *in vivo* protein corona formation on polymeric nanoparticles and impact on pharmacokinetics. *Nature Communications* **8**, 777 (2017).

128. Wolfram, J. et al. A chloroquine-induced macrophage-preconditioning strategy for improved nanodelivery. *Sci Rep* **7**, 13738 (2017).

129. Kishimoto, T.K. et al. Improving the efficacy and safety of biologic drugs with tolerogenic nanoparticles. *Nature nanotechnology* **11**, 890-899 (2016).

130. Albanese, A., Tang, P.S. & Chan, W.C. The effect of nanoparticle size, shape, and surface chemistry on biological systems. *Annual review of biomedical engineering* **14**, 1-16 (2012).

131. Chithrani, B.D., Ghazani, A.A. & Chan, W.C. Determining the size and shape dependence of gold nanoparticle uptake into mammalian cells. *Nano lett* **6**, 662-668 (2006).

132. Chithrani, B.D. & Chan, W.C. Elucidating the mechanism of cellular uptake and removal of protein-coated gold nanoparticles of different sizes and shapes. *Nano lett* **7**, 1542-1550 (2007).

133. Vácha, R., Martinez-Veracoechea, F.J. & Frenkel, D. Receptor-mediated endocytosis of nanoparticles of various shapes. *Nano letters* **11**, 5391-5395 (2011).

134. Sykes, E.A., Chen, J., Zheng, G. & Chan, W.C. Investigating the impact of nanoparticle size on active and passive tumor targeting efficiency. (2014).

135. Lloyd-Evans, E. et al. Niemann-Pick disease type C1 is a sphingosine storage disease that causes deregulation of lysosomal calcium. *Nature medicine* **14**, 1247-1255 (2008).

136. Lifland, A.W. et al. Human respiratory syncytial virus nucleoprotein and inclusion bodies antagonize the innate immune response mediated by MDA5 and MAVS. *Journal of virology* **86**, 8245-8258 (2012).

137. Tibbitt, M.W., Dahlman, J.E. & Langer, R. Emerging frontiers in drug delivery. *Journal of the American Chemical Society* **138**, 704-717 (2016).

138. Sahin, U., Kariko, K. & Tureci, O. mRNA-based therapeutics--developing a new class of drugs. *Nat Rev Drug Discov* **13**, 759-780 (2014).

139. Hao, J. et al. Rapid Synthesis of a Lipocationic Polyester Library via Ring-Opening Polymerization of Functional Valerolactones for Efficacious siRNA Delivery. *J Am Chem Soc* **29**, 9206-9209 (2015).

140. Siegwart, D.J. et al. Combinatorial synthesis of chemically diverse core-shell nanoparticles for intracellular delivery. *Proc Natl Acad Sci U S A* **108**, 12996-13001 (2011).

141. Zhang, Y. et al. Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37-53 (2016).

142. Chen, D. et al. Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation. *J Am Chem Soc* **134**, 6948-6951 (2012).

143. Carmeliet, P. & Jain, R.K. Molecular mechanisms and clinical applications of angiogenesis. *Nature* **473**, 298-307 (2011).

144. Swirski, F.K. & Nahrendorf, M. Leukocyte behavior in atherosclerosis, myocardial infarction, and heart failure. *Science* **339**, 161-166 (2013).

145. Sager, H.B. et al. RNAi targeting multiple cell adhesion molecules reduces immune cell recruitment and vascular inflammation after myocardial infarction. *Science Translational Medicine* **8**, 342ra380-342ra380 (2016).

146. Yun, S. et al. Interaction between integrin alpha5 and PDE4D regulates endothelial inflammatory signalling. *Nat Cell Biol*, 1043–1053 (2016).

147. Xue, W. et al. Small RNA combination therapy for lung cancer. *Proc Natl Acad Sci U S A* **111**, E3553-E3561 (2014).

148. Sager, H.B. et al. Proliferation and Recruitment Contribute to Myocardial Macrophage Expansion in Chronic Heart Failure. *Circ Res* **119**, 853-864 (2016).

149. White, K. et al. Genetic and hypoxic alterations of the microRNA-210-ISCU1/2 axis promote iron-sulfur deficiency and pulmonary hypertension. *EMBO Mol Med* **7**, 695-713 (2015).

150. Khan, O.F. et al. Dendrimer-Inspired Nanomaterials for the in Vivo Delivery of siRNA to Lung Vasculature. *Nano Lett* **15**, 3008-3016 (2015).

151. Khan, O.F. et al. Ionizable amphiphilic dendrimer-based nanomaterials with alkyl-chain-substituted amines for tunable siRNA delivery to the liver endothelium *in vivo*. *Angew Chem Int Ed Engl* **53**, 14397-14401 (2014).

152. Geary, R.S., Norris, D., Yu, R. & Bennett, C.F. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv Drug Deliv Rev* **87**, 46-51 (2015).

153. Crooke, S.T., Wang, S., Vickers, T.A., Shen, W. & Liang, X.H. Cellular uptake and trafficking of antisense oligonucleotides. *Nat Biotechnol* **35**, 230-237 (2017).

154. Ni, C.W., Kumar, S., Ankeny, C.J. & Jo, H. Development of immortalized mouse aortic endothelial cell lines. *Vasc Cell* **6**, 7 (2014).

155. Tao, W. et al. Noninvasive imaging of lipid nanoparticle-mediated systemic delivery of small-interfering RNA to the liver. *Mol Ther* **18**, 1657-1666 (2010).

156. Kranz, L.M. et al. Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. *Nature* **534**, 396-401 (2016).

157. Whitehead, K.A. et al. In vitro-in vivo translation of lipid nanoparticles for hepatocellular siRNA delivery. *ACS Nano* **6**, 6922-6929 (2012).

158. Bhatia, S.N. & Ingber, D.E. Microfluidic organs-on-chips. *Nat Biotechnol* **32**, 760-772 (2014).

159. Fatehullah, A., Tan, S.H. & Barker, N. Organoids as an in vitro model of human development and disease. *Nat Cell Biol* **18**, 246-254 (2016).

160. Koga, J. et al. Macrophage Notch Ligand Delta-Like 4 Promotes Vein Graft Lesion Development: Implications for the Treatment of Vein Graft Failure. *Arterioscler Thromb Vasc Biol* **35**, 2343-2353 (2015).

161. Araki, K. et al. Translation is actively regulated during the differentiation of CD8+ effector T cells. *Nat Immunol* **18**, 1046-1057 (2017).

162. Yaari, Z. et al. Theranostic barcoded nanoparticles for personalized cancer medicine. *Nature Communications* **7**, 13325 (2016).

163. Zhou, K. et al. Modular degradable dendrimers enable small RNAs to extend survival in an aggressive liver cancer model. *Proc Natl Acad Sci U S A* **113**, 520-525 (2016).

164. Yoshida, H. & Kisugi, R. Mechanisms of LDL oxidation. *Clinica Chimica Acta* **411**, 1875-1882 (2010).

165. Ikonen, E. Cellular cholesterol trafficking and compartmentalization. *Nature Reviews: Molecular Cell Biology* **9**, 125-138 (2008).

166. Cases, S. et al. ACAT-2, a second mammalian acyl-CoA: cholesterol acyltransferase its cloning, expression, and characterization. *Journal of Biological Chemistry* **273**, 26755-26764 (1998).

167. Cases, S. et al. Identification of a gene encoding an acyl CoA: diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proc Natl Acad Sci U S A* **95**, 13018-13023 (1998).

168. Wasan, K.M., Brocks, D.R., Lee, S.D., Sachs-Barrable, K. & Thornton, S.J. Impact of lipoproteins on the biological activity and disposition of hydrophobic

drugs: implications for drug discovery. *Nature Reviews: Drug Discovery* **7**, 84-99 (2008).

169. Steinbrecher, U.P., Zhang, H. & Lougheed, M. Role of oxidatively modified LDL in atherosclerosis. *Free Radical Biology and Medicine* **9**, 155-168 (1990).

170. Hulthe, J. & Fagerberg, B. Circulating oxidized LDL is associated with subclinical atherosclerosis development and inflammatory cytokines (AIR Study). *Arteriosclerosis, thrombosis, and vascular biology* **22**, 1162-1167 (2002).

171. Tall, A.R. & Yvan-Charvet, L. Cholesterol, inflammation and innate immunity. *Nature Reviews: Immunology* **15**, 104-116 (2015).

172. Miller, Y.I. et al. Oxidation-specific epitopes are danger-associated molecular patterns recognized by pattern recognition receptors of innate immunity. *Circulation research* **108**, 235-248 (2011).

173. Li, D. & Mehta, J.L. (Elsevier Science, 2005).

174. Manninen, V. et al. Joint effects of serum triglyceride and LDL cholesterol and HDL cholesterol concentrations on coronary heart disease risk in the Helsinki Heart Study. Implications for treatment. *Circulation* **85**, 37-45 (1992).

175. Armstrong, V. et al. The association between serum Lp (a) concentrations and angiographically assessed coronary atherosclerosis: dependence on serum LDL levels. *Atherosclerosis* **62**, 249-257 (1986).

176. Kontush, A. & Chapman, M.J. Antiatherogenic small, dense HDL—guardian angel of the arterial wall? *Nature Clinical Practice Cardiovascular Medicine* **3**, 144–153 (2006).

177. Holvoet, P., Keyzer, D.D. & Jacobs Jr, D. Oxidized LDL and the metabolic syndrome. *Future lipidology* **3**, 637-649 (2008).

178. Eckel, R.H., Grundy, S.M. & Zimmet, P.Z. The metabolic syndrome. *Lancet* **365**, 1415-1428 (2005).

179. Coelho, T. et al. Safety and efficacy of RNAi therapy for transthyretin amyloidosis. *N Engl J Med* **369**, 819-829 (2013).

180. Ishibashi, S. et al. Hypercholesterolemia in low density lipoprotein receptor knockout mice and its reversal by adenovirus-mediated gene delivery. *J Clin Invest* **92**, 883-893 (1993).

181. Frykman, P.K., Brown, M.S., Yamamoto, T., Goldstein, J.L. & Herz, J. Normal plasma lipoproteins and fertility in gene-targeted mice homozygous for a disruption in the gene encoding very low density lipoprotein receptor. *Proc Natl Acad Sci U S A* **92**, 8453-8457 (1995).

182. Moore, K.J., Sheedy, F.J. & Fisher, E.A. Macrophages in atherosclerosis: a dynamic balance. *Nature Reviews: Immunology* **13**, 709-721 (2013).

183. Choudhury, R.P., Lee, J.M. & Greaves, D.R. Mechanisms of Disease: macrophage-derived foam cells emerging as therapeutic targets in atherosclerosis. *Nature Clinical Practice Cardiovascular Medicine* **2**, 309–315 (2005).

184. Moore, K.J. & Tabas, I. Macrophages in the pathogenesis of atherosclerosis. *Cell* **145**, 341-355 (2011).

185. Brown, M.S. & Goldstein, J.L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* **89**, 331-340 (1997).

186. Chawla, A., Saez, E. & Evans, R.M. Don't know much bile-ology. *Cell* **103**, 1-4 (2000).

187. Hassan, H.H., Denis, M., Krimbou, L., Marcil, M. & Genest, J. Cellular cholesterol homeostasis in vascular endothelial cells. *Canadian Journal of Cardiology* **22**, 35B-40B (2006).

188. O'Connell, B.J., Denis, M. & Genest, J. Cellular physiology of cholesterol efflux in vascular endothelial cells. *Circulation* **110**, 2881-2888 (2004).

189. Ronan, T., Qi, Z. & Naegle, K.M. Avoiding common pitfalls when clustering biological data. *Science Signaling* **9**, re6-re6 (2016).

190. Yin, H. et al. Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing. *Nature Biotechnology*, nbt. 4005 (2017).

191. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research* **42**, e168 (2014).

192. Miller, J.B. et al. Non-Viral CRISPR/Cas Gene Editing *In Vitro* and *In Vivo* Enabled by Synthetic Nanoparticle Co-Delivery of Cas9 mRNA and sgRNA. *Angew Chem Int Ed Engl* **56**, 1059-1063 (2017).

193. Jiang, C. et al. A non-viral CRISPR/Cas9 delivery system for therapeutically targeting HBV DNA and pcsk9 *in vivo*. *Cell Res* **27**, 440-443 (2017).

194. Hsu, P.D., Lander, E.S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278 (2014).

195. Ulitsky, I. & Bartel, D.P. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**, 26-46 (2013).

196. Poisson, J. et al. Liver sinusoidal endothelial cells: Physiology and role in liver diseases. *J Hepatol* **66**, 212-227 (2017).

197. Kuai, R., Li, D., Chen, Y.E., Moon, J.J. & Schwendeman, A. High-Density Lipoproteins (HDL) – Nature's Multi-Functional Nanoparticles. *ACS Nano* **10**, 3015-3041 (2016).

198. Monopoli, M.P., Aberg, C., Salvati, A. & Dawson, K.A. Biomolecular coronas provide the biological identity of nanosized materials. *Nature Nanotechnology* **7**, 779-786 (2012).

199. Hajj, K.A. & Whitehead, K.A. Tools for translation: non-viral materials for therapeutic mRNA delivery. *Nature Reviews Materials* **2**, 17056 (2017).

200. Jones, S.W. et al. Nanoparticle clearance is governed by Th1/Th2 immunity and strain background. *J Clin Invest* **123**, 3061-3073 (2013).

201. Mutemberezi, V., Guillemot-Legris, O. & Muccioli, G.G. Oxysterols: From cholesterol metabolites to key mediators. *Prog Lipid Res* **64**, 152-169 (2016).

202. Brown, A.J. & Jessup, W. Oxysterols: Sources, cellular storage and metabolism, and new insights into their roles in cholesterol homeostasis. *Mol Aspects Med* **30**, 111-122 (2009).

203. Kulig, W., Cwiklik, L., Jurkiewicz, P., Rog, T. & Vattulainen, I. Cholesterol oxidation products and their biological importance. *Chem Phys Lipids* **199**, 144-160 (2016).

204. Mahley, R.W. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science* **240**, 622-630 (1988).

205. Palm, W. & Thompson, C.B. Nutrient acquisition strategies of mammalian cells. *Nature* **546**, 234-242 (2017).

206. Lokugamage, M.P., Sago, C.D. & Dahlman, J.E. Testing thousands of nanoparticles in vivo using DNA barcodes. *Current Opinion in Biomedical Engineering* **7**, 1-8 (2018).

207. Sago, C.D. et al. Modifying a Commonly Expressed Endocytic Receptor Retargets Nanoparticles in Vivo. *Nano Lett* (2018).

208. Rizk, M. & Tuzmen, S. Update on the clinical utility of an RNA interference-based treatment: focus on Patisiran. *Pharmgenomics Pers Med* **10**, 267-278 (2017).

209. Kolios, G., Valatas, V. & Kouroumalis, E. Role of Kupffer cells in the pathogenesis of liver disease. *World J Gastroenterol* **12**, 7413-7420 (2006).

210. Akinc, A. & Battaglia, G. Exploiting endocytosis for nanomedicines. *Cold Spring Harb Perspect Biol* **5**, a016980 (2013).

211. Voigt, J., Christensen, J. & Shastri, V.P. Differential uptake of nanoparticles by endothelial cells through polyelectrolytes with affinity for caveolae. *Proc Natl Acad Sci U S A* **111**, 2942-2947 (2014).

212. Maugeri, M. et al. Linkage between endosomal escape of LNP-mRNA and loading into EVs for transport to other cells. *Nat Commun* **10**, 4333 (2019).

213. Anselmo, A.C. & Mitragotri, S. Nanoparticles in the clinic: An update. *Bioengineering & Translational Medicine* **4**, e10143 (2019).

214. Lewis, G.F. & Rader, D.J. New insights into the regulation of HDL metabolism and reverse cholesterol transport. *Circ Res* **96**, 1221-1232 (2005).

215. Antonescu, C.N., McGraw, T.E. & Klip, A. Reciprocal regulation of endocytosis and metabolism. *Cold Spring Harb Perspect Biol* **6**, a016964 (2014).

216. Ochaba, J. et al. A novel and translational role for autophagy in antisense oligonucleotide trafficking and activity. *Nucleic acids research* **47**, 11284-11303 (2019).

217. Park, W.S. et al. Comprehensive identification of PIP3-regulated PH domains from C. elegans to H. sapiens by model prediction and live imaging. *Mol Cell* **30**, 381-392 (2008).

218. Dibble, C.C. & Cantley, L.C. Regulation of mTORC1 by PI3K signaling. *Trends Cell Biol* **25**, 545-555 (2015).

219. Vivanco, I. & Sawyers, C.L. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat Rev Cancer* **2**, 489-501 (2002).

220. Laketa, V. et al. PIP(3) induces the recycling of receptor tyrosine kinases. *Sci Signal* **7**, ra5 (2014).

221. Fields, I.C., King, S.M., Shteyn, E., Kang, R.S. & Folsch, H. Phosphatidylinositol 3,4,5-trisphosphate localization in recycling endosomes is necessary for AP-1B-dependent sorting in polarized epithelial cells. *Mol Biol Cell* **21**, 95-105 (2010).

222. Hay, N. The Akt-mTOR tango and its relevance to cancer. *Cancer Cell* **8**, 179-183 (2005).

223. Khan, O.F. et al. Endothelial siRNA delivery in nonhuman primates using ionizable low-molecular weight polymeric nanoparticles. *Sci Adv* **4**, eaar8409 (2018).

224. Wang, C., de Jong, E., Sjollema, K.A. & Zuhorn, I.S. Entry of PIP3-containing polyplexes into MDCK epithelial cells by local apical-basal polarity reversal. *Sci Rep* **6**, 21436 (2016).

225. McMahon, H.T. & Boucrot, E. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* **12**, 517-533 (2011).

226. Orlandi, P.A. & Fishman, P.H. Filipin-dependent inhibition of cholera toxin: evidence for toxin internalization and activation through caveolae-like domains. *J Cell Biol* **141**, 905-915 (1998).

227. Kirschman, J.L. et al. Characterizing exogenous mRNA delivery, trafficking, cytoplasmic release and RNA-protein correlations at the level of single cells. *Nucleic Acids Res* **45**, e113 (2017).

228. Sommerfeld, S.D. et al. Interleukin-36gamma-producing macrophages drive IL-17-mediated fibrosis. *Sci Immunol* **4** (2019).

229. Chen, F. et al. mTOR mediates IL-23 induction of neutrophil IL-17 and IL-22 production. *The Journal of Immunology* **196**, 4390-4399 (2016).

230. Moreno-Marín, N. et al. Aryl hydrocarbon receptor promotes liver polyploidization and inhibits PI3K, ERK, and Wnt/β-catenin signaling. *iScience* **4**, 44-63 (2018).

231. Yan, J., Wang, R. & Horng, T. mTOR is key to T cell transdifferentiation. *Cell metabolism* **29**, 241-242 (2019).

232. Karmaus, P.W. et al. Metabolic heterogeneity underlies reciprocal fates of TH 17 cell stemness and plasticity. *Nature* **565**, 101-105 (2019).

233. Meng, H. et al. Low density lipoprotein receptor-related protein-1 (LRP1) regulates thrombospondin-2 (TSP2) enhancement of Notch3 signaling. *J Biol Chem* **285**, 23047-23055 (2010).

234. Harayama, T. & Riezman, H. Understanding the diversity of membrane lipid composition. *Nature reviews. Molecular cell biology* (2018).

235. Malapaka, R.R. et al. Identification and mechanism of 10-carbon fatty acid as modulating ligand of peroxisome proliferator-activated receptors. *J Biol Chem* **287**, 183-195 (2012).

236. Song, G., Ouyang, G. & Bao, S. The activation of Akt/PKB signaling pathway and cell survival. *J Cell Mol Med* **9**, 59-71 (2005).

237. Lo, Y.Y. & Cruz, T.F. Involvement of reactive oxygen species in cytokine and growth factor induction of c-fos expression in chondrocytes. *J Biol Chem* **270**, 11727-11730 (1995).

238. Gatica, D., Lahiri, V. & Klionsky, D.J. Cargo recognition and degradation by selective autophagy. *Nat Cell Biol* **20**, 233-242 (2018).

239. Willoughby, J.L.S. et al. Evaluation of GalNAc-siRNA Conjugate Activity in Pre-clinical Animal Models with Reduced Asialoglycoprotein Receptor Expression. *Mol Ther* **26**, 105-114 (2018).

240. Ostergaard, M.E. et al. Efficient Synthesis and Biological Evaluation of 5'-GalNAc Conjugated Antisense Oligonucleotides. *Bioconjug Chem* **26**, 1451-1455 (2015).

241. Nair, J.K. et al. Multivalent N-acetylgalactosamine-conjugated siRNA localizes in hepatocytes and elicits robust RNAi-mediated gene silencing. *J Am Chem Soc* **136**, 16958-16961 (2014).

242. Sago, C.D., Krupczak, B.R., Lokugamage, M.P., Gan, Z. & Dahlman, J.E. Cell Subtypes Within the Liver Microenvironment Differentially Interact with Lipid Nanoparticles. *Cellular and Molecular Bioengineering* (2019).

243. Stocker, H. et al. Living with lethal PIP3 levels: viability of flies lacking PTEN restored by a PH domain mutation in Akt/PKB. *Science* **295**, 2088-2091 (2002).

244. Fine, B. et al. Activation of the PI3K pathway in cancer through inhibition of PTEN by exchange factor P-REX2a. *Science* **325**, 1261-1265 (2009).

245. Kanani, H., Chrysanthopoulos, P.K. & Klapa, M.I. Standardizing GC-MS metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* **871**, 191-201 (2008).

246. Kanani, H.H. & Klapa, M.I. Data correction strategy for metabolomics analysis using gas chromatography-mass spectrometry. *Metab Eng* **9**, 39-51 (2007).

247. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121 (2001).

248. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

249. Kuleshov, M.V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97 (2016).

250. Oldenborg, P.-A. et al. Role of CD47 as a marker of self on red blood cells. *Science* **288**, 2051-2054 (2000).

251. Kahvejian, A., Quackenbush, J. & Thompson, J.F. What would you do if you could sequence everything? *Nature Biotechnology* **26**, 1125 (2008).

252. Navin, N. et al. Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90-94 (2011).

253. Tang, F. et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods* **6**, 377 (2009).

254. Patel, A.P. et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401 (2014).

255. Ting, D.T. et al. Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell reports* **8**, 1905-1918 (2014).

256. Streets, A.M. et al. Microfluidic single-cell whole-transcriptome sequencing. *Proc Natl Acad Sci U S A* **111**, 7048-7053 (2014).

257. Pérez-Torrado, R. et al. Ecological interactions among Saccharomyces cerevisiae strains: insight into the dominance phenomenon. *Scientific reports* **7**, 43603 (2017).

258. Meyer, M. et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proceedings of the National Academy of Sciences* **112**, 851-856 (2015).

259. Treutlein, B. et al. Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534**, 391-395 (2016).

260. Zeisel, A. et al. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138-1142 (2015).

261. Usoskin, D. et al. Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. *Nature neuroscience* **18**, 145-153 (2015).

262. Villani, A.C. & Shekhar, K. Single-Cell RNA Sequencing of Human T Cells. *Methods Mol Biol* **1514**, 203-239 (2017).

263. Villani, A.-C. et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, eaah4573 (2017).

264. Papalexi, E. & Satija, R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* (2017).

265. Dixit, A. et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **167**, 1853-1866.e1817 (2016).

266. Rozenblatt-Rosen, O., Stubbington, M.J.T., Regev, A. & Teichmann, S.A. The Human Cell Atlas: from vision to reality. *Nature* **550**, 451-453 (2017).

267. Dey, S.S., Kester, L., Spanjaard, B., Bienko, M. & van Oudenaarden, A. Integrated genome and transcriptome sequencing of the same cell. *Nature Biotechnology* **33**, 285 (2015).

268. Angermueller, C. et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nature Methods* **13**, 229 (2016).

269. Cheow, L.F. et al. Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nature Methods* **13**, 833 (2016).

270. Hou, Y. et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res* **26**, 304-319 (2016).

271. Gjoneska, E. et al. Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. *Nature* **518**, 365-369 (2015).

272. Zhang, B. et al. Proteogenomic characterization of human colon and rectal cancer. *Nature* **513**, 382 (2014).

273. Alfaro, J.A., Sinha, A., Kislinger, T. & Boutros, P.C. Onco-proteogenomics: cancer proteomics joins forces with genomics. *Nature methods* **11**, 1107 (2014).

274. Geyer, P.E. et al. Plasma proteome profiling to assess human health and disease. *Cell systems* **2**, 185-195 (2016).

275. Mun, D.G. et al. Proteogenomic Characterization of Human Early-Onset Gastric Cancer. *Cancer Cell* **35**, 111-124.e110 (2019).

276. Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature biotechnology* **32**, 381 (2014).

277. Yan, J., Risacher, S.L., Shen, L. & Saykin, A.J. Network approaches to systems biology analysis of complex disease: integrative methods for multi-omics data. *Briefings in bioinformatics* (2017).

278. Carrow, J.K. et al. Widespread changes in transcriptome profile of human mesenchymal stem cells induced by two-dimensional nanosilicates. *Proc Natl Acad Sci U S A* **115**, E3905-e3913 (2018).

279. Feliu, N. et al. Next-generation sequencing reveals low-dose effects of cationic dendrimers in primary human bronchial epithelial cells. *ACS nano* **9**, 146-163 (2014).

280. Lucafò, M. et al. Profiling the molecular mechanism of fullerene cytotoxicity on tumor cells by RNA-seq. *Toxicology* **314**, 183-192 (2013).

281. Gliga, A.R. et al. Cerium oxide nanoparticles inhibit differentiation of neural stem cells. *Scientific reports* **7**, 9284 (2017).

282. Simon, D.F. et al. RNA-Seq analysis of the effects of metal nanoparticle exposure on the transcriptome of Chlamydomonas reinhardtii. *Applied and environmental microbiology*, AEM. 00998-00913 (2013).

283. Beauvais-Flück, R., Slaveykova, V.I. & Cosio, C. Transcriptomic and physiological responses of the green microalga Chlamydomonas reinhardtii during short-term exposure to subnanomolar methylmercury concentrations. *Environmental science & technology* **50**, 7126-7134 (2016).

284. Zheng, M., Lu, J. & Zhao, D. Toxicity and Transcriptome Sequencing (RNA-seq) Analyses of Adult Zebrafish in Response to Exposure Carboxymethyl Cellulose Stabilized Iron Sulfide Nanoparticles. *Scientific reports* **8**, 8083 (2018).

285. Veiseh, O. et al. Size- and shape-dependent foreign body immune response to materials implanted in rodents and non-human primates. *Nat Mater* **14**, 643-651 (2015).

286. Hebels, D., Carlier, A., Coonen, M.L.J., Theunissen, D.H. & de Boer, J. cBiT: A transcriptomics database for innovative biomaterial engineering. *Biomaterials* **149**, 88-97 (2017).

287. Conesa, A. et al. A survey of best practices for RNA-seq data analysis. *Genome Biology* **17**, 13 (2016).

288. Novembre, J. & Stephens, M. Interpreting principal component analyses of spatial population genetic variation. *Nat Genet* **40**, 646-649 (2008).

289. Ringnér, M. What is principal component analysis? *Nature biotechnology* **26**, 303-304 (2008).

290. Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat Protoc* **7**, 500-507 (2012).

291. Mecham, B.H., Nelson, P.S. & Storey, J.D. Supervised normalization of microarrays. *Bioinformatics* **26**, 1308-1315 (2010).

292. Lever, J., Krzywinski, M. & Altman, N. Points of Significance: Principal component analysis. *Nature Methods* **14**, 641-642 (2017).

293. Maaten, L.v.d. & Hinton, G. Visualizing data using t-SNE. *Journal of machine learning research* **9**, 2579-2605 (2008).

294. Habib, N. et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nature methods* **14**, 955 (2017).

295. Macosko, E.Z. et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202-1214 (2015).

296. Shalek, A.K. et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363-369 (2014).

297. Wattenberg, M.a.V., Fernanda and Johnson, Ian How to Use t-SNE Effectively. *Distill* (2016).

298. Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. in Kdd, Vol. 96 226-231 (1996).

299. Kohonen, T. The self-organizing map. *Neurocomputing* **21**, 1-6 (1998).

300. Hartigan, J.A. & Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **28**, 100-108 (1979).

301. Wiwie, C., Baumbach, J. & Röttger, R. Comparing the performance of biomedical clustering methods. *Nature methods* **12**, 1033-1038 (2015).

302. Xu, C. & Su, Z. Identification of cell types from single-cell transcriptomes using a novel clustering method. *Bioinformatics* **31**, 1974-1980 (2015).

303. Jiang, D., Tang, C. & Zhang, A. Cluster analysis for gene expression data: A survey. *IEEE Transactions on knowledge and data engineering* **16**, 1370-1386 (2004).

304. Quackenbush, J. Computational genetics: computational analysis of microarray data. *Nature reviews genetics* **2**, 418 (2001).

305. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recogn. Lett.* **31**, 651-666 (2010).

306. Ding, C. & He, X. in Proceedings of the 2004 ACM symposium on Applied computing 584-589 (ACM, 2004).

307. Handl, J., Knowles, J. & Kell, D.B. Computational cluster validation in post-genomic data analysis. *Bioinformatics* **21**, 3201-3212 (2005).

308. Gehlenborg, N. et al. Visualization of omics data for systems biology. *Nature methods* **7**, S56 (2010).

309. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**, D353-d361 (2017).

310. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).

311. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457-462 (2016).

312. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res* **45**, D331-d338 (2017).

313. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29 (2000).

314. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498-2504 (2003).

315. Tong, A.H.Y. et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* **294**, 2364-2368 (2001).

316. Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437-1452. e1417 (2017).

317. Hughes, T.R. et al. Functional discovery via a compendium of expression profiles. *Cell* **102**, 109-126 (2000).

318. Wong, B. Points of view: Color coding. *Nature Methods* **7**, 573-573 (2010).

319. Hansen, K.D., Brenner, S.E. & Dudoit, S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic acids research* **38**, e131-e131 (2010).

320. Loman, N.J. et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nature biotechnology* **30**, 434 (2012).

321. Ross, M.G. et al. Characterizing and measuring bias in sequence data. *Genome Biol* **14**, R51 (2013).

322. Asmann, Y.W. et al. 3'tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. *BMC genomics* **10**, 531 (2009).

323. Wu, A.R. et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nature methods* **11**, 41-46 (2014).

324. Griffith, M. et al. Alternative expression analysis by RNA sequencing. *Nature methods* **7**, 843 (2010).

325. Sago, C.D. et al. Barcoding chemical modifications into nucleic acids improves drug stability in vivo. *Journal of Materials Chemistry B* **6**, 7197-7203 (2018).

326. Chan, K.Y. et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat Neurosci* **20**, 1172-1179 (2017).

327. Deverman, B.E. et al. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nat Biotechnol* **34**, 204-209 (2016).

328. Belser, J.A., Katz, J.M. & Tumpey, T.M. The ferret as a model organism to study influenza A virus infection. *Dis Model Mech* **4**, 575-579 (2011).

329. Butterfield, G.L. et al. Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature* **552**, 415 (2017).

330. Chen, Z. et al. Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106-111 (2019).

331. Silva, D.A. et al. De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186-191 (2019).

332. Guerette, P.A. et al. Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science. *Nat Biotechnol* **31**, 908-915 (2013).

333. Shamay, Y. et al. Quantitative self-assembly prediction yields targeted nanomedicines. *Nat Mater* **17**, 361-368 (2018).

334. Yamankurt, G. et al. Exploration of the nanomedicine-design space with high-throughput screening and machine learning. *Nat Biomed Eng* **3**, 318-327 (2019).

335. Rath, S. et al. VSEPRnet: Physical structure encoding of sequence-based biomolecules for functionality prediction: Case study with peptides. *bioRxiv*, 656033 (2019).

336. Shalem, O., Sanjana, N.E. & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* **16**, 299-311 (2015).

337. Chen, S. et al. Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. *Cell* **160**, 1246-1260 (2015).

338. Shalem, O. et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87 (2014).

339. Batista, P.J. & Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298-1307 (2013).

340. Darmanis, S. et al. Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell reports* **14**, 380-389 (2016).

341. Foster, D.J. et al. Advanced siRNA Designs Further Improve In Vivo Performance of GalNAc-siRNA Conjugates. *Mol Ther* **26**, 708-717 (2018).

342. Vu, L.P. et al. The N6-methyladenosine (m6A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat Med* **23**, 1369-1376 (2017).

343. Roundtree, I.A., Evans, M.E., Pan, T. & He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **169**, 1187-1200 (2017).

344. Zhao, B.S., Roundtree, I.A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**, 31-42 (2017).