nature nanotechnology

ARTICLES

https://doi.org/10.1038/s41565-021-01030-y

Species-dependent in vivo mRNA delivery and cellular responses to nanoparticles

Marine Z. C. Hatit[1,3], Melissa P. Lokugamage[1,3], Curtis N. Dobrowolski[1,3], Kalina Paunovska[1], Huanzhen Ni[1], Kun Zhao[1], Daryll Vanover[2], Jared Beyersdorf[2], Hannah E. Peck[2], David Loughrey[1], Manaka Sato[1], Ana Cristian[1], Philip J. Santangelo[1,2] and James E. Dahlman[1]✉

Nanoparticles are tested in mice and non-human primates before being selected for clinical trials. Yet the extent to which mRNA delivery, as well as the cellular response to mRNA drug delivery vehicles, is conserved across species in vivo is unknown. Using a species-independent DNA barcoding system, we have compared how 89 lipid nanoparticles deliver mRNA in mice with humanized livers, primatized livers and four controls: mice with 'murinized' livers as well as wild-type BL/6, Balb/C and NZB/BINJ mice. We assessed whether functional delivery results in murine, non-human primate and human hepatocytes can be used to predict delivery in the other species in vivo. By analysing in vivo hepatocytes by RNA sequencing, we identified species-dependent responses to lipid nanoparticles, including mRNA translation and endocytosis. These data support an evidence-based approach to making small-animal preclinical nanoparticle studies more predictive, thereby accelerating the development of RNA therapies.

mRNA, short interfering RNA (siRNA) and antisense oligonucleotides have been used to treat disease in humans[1–8]. The subset of these drugs administered intravenously requires nucleic acid delivery to diseased cells. To date, US Food and Drug Administration (FDA)-approved siRNA drugs use carbohydrate conjugates[9] or lipid nanoparticles (LNPs)[10], both of which exploit endogenous trafficking pathways to deliver siRNA into hepatocytes[11,12]. These insights into evolutionary conserved pathways that promote drug delivery have proven valuable as drugs have progressed from mice to non-human primates (NHPs) and humans. Furthermore, by comparing delivery in murine hepatocytes and human hepatocytes from 'humanized' mice[13–16], which are mice engineered to include human cells, often hepatocytes, the adeno-associated virus (AAV) field has demonstrated that humanized mice are more predictive of AAV delivery in human cells than wild-type mice.

These examples underscore the value of understanding how genes that alter nanoparticle delivery vary between preclinical species and humans. Yet no study has systematically (that is, using dozens of chemically distinct nanoparticles) compared LNP delivery across species in vivo. We therefore tested three hypotheses. First, we would observe species-dependent differences in hepatocyte LNP delivery. Second, the differences between NHP delivery and human delivery would be smaller than the differences between murine and human delivery. Third, transcriptomic analyses of hepatocytes from different species would identify cell signalling pathways associated with differences in delivery. However, testing these hypotheses is challenging because humanized mice cost several thousand dollars per animal.

To overcome these issues, we developed species-agnostic nanoparticle delivery screening (SANDS). SANDS has three traits required for these studies. First, it is high throughput and in vivo. Second, it reports functional LNP delivery (in this case, mRNA translated into functional protein) at the cell-type level. Third, it is species independent, as it does not require Cre reporters or other genomic DNA constructs. These studies also required mice with humanized and primatized livers, as well as murinized mice, which have the same immune system as primatized and humanized mice, and undergo the same cell engraftment process. Using SANDS, we quantified how 89 chemically distinct LNPs functionally deliver mRNA in six murine models: BL/6, Balb/C, NZB/BlNJ (NZB), Fah/Rag2/interleukin gamma chain (FRG) knockout (KO) mice[17,18] on a BL/6 background with humanized livers (humanized), NHP livers (primatized) and murine livers (murinized). In addition to quantifying how well mouse, NHP and human hepatocytes predict delivery in one another, we identified species-dependent responses to LNPs.

Species-agnostic nanoparticle delivery screening (SANDS)

We designed SANDS to quantify functional hepatocyte mRNA delivery mediated by dozens of LNPs across species. Each LNP was formulated by microfluidic mixing[19] and carried a functional mRNA encoding a reporter and a DNA barcode: LNP-1 carried mRNA and DNA barcode 1, whereas LNP-N carried mRNA and DNA barcode N. We performed quality control on individual LNPs by quantifying their hydrodynamic diameter and polydispersity by dynamic light scattering; stable, monodisperse LNPs with a hydrodynamic diameter less than 200 nm were pooled together. Twenty-four hours after LNPs were administered, reporter$^+$ cells were isolated by fluorescence-activated cell sorting (FACS). By sequencing the DNA barcodes in these cells, we identified barcodes that colocalized with cells in which functional mRNA delivery occurred (Fig. 1a,b). Using DNA barcodes optimized to facilitate sensitive in vivo readouts[20], we formulated LNPs carrying mRNA and DNA barcodes in a ratio of 10:1 (ref. [21]).

We considered several reporter mRNAs for SANDS. We excluded Cre (ref. [22]), which requires transgenes unavailable in

[1]Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, USA. [2]Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA, USA. [3]These authors contributed equally: Marine Z.C. Hatit, Melissa P. Lokugamage, Curtis N. Dobrowolski. ✉e-mail: james.dahlman@bme.gatech.edu

humanized mice, and luciferase, which requires permeabilization and intracellular staining for flow cytometry. Instead, we formulated LNPs to carry mRNA encoding a glycosylphosphatidylinositol (GPI)-anchored camelid VHH antibody (anchored-VHH, aVHH). Linking the VHH domain with a GPI anchor induces cell-surface aVHH expression[23], allowing aVHH$^+$ cells to be detected with an anti-camelid VHH antibody (Fig. 1c). We formulated aVHH mRNA with Lipofectamine MessengerMax and transfected A549 cells with 1 μg mRNA per well. After 24 h, we observed aVHH expression by immunofluorescent staining (Fig. 1d). We then confirmed that the expression of the mRNA reporter was similar in human and mouse cells in vitro. We formulated aVHH mRNA with polyethyleneimine (PEI) and transfected Fa2N-4 and AML-12 cells with 0.5 μg mRNA per well. After 24 h, we found 48% of aVHH$^+$ cells in human and 46% in mouse using flow cytometry (Fig. 1e). Finally, we quantified aVHH expression in vivo after formulating a previously reported[24] LNP with aVHH mRNA and intravenously injecting BL/6 mice with 1 mg kg^{-1} mRNA. The percentage of functionally transfected hepatocytes (CD31$^-$CD45$^-$aVHH$^+$ cells) was highest at 24 h (Fig. 1f). We therefore selected a timepoint of 24 h for subsequent studies.

In vivo LNP delivery across cells from different species

We used SANDS to quantify how 89 LNPs delivered mRNA in vivo to cells derived from different species. We chose a combinatorial library design using four components (Fig. 1g,h): lipomers (7C1, cKK-E12 and cKK-E15)[25–27], cholesterol (cholesterol and 20α-hydroxycholesterol)[24,28,29], polyethyleneglycol (PEG)-lipids ($C_{14}PEG_{2000}$ and $C_{18}PEG_{2000}$)[30] and helper lipids (DOPE, DOTAP and DOTMA)[31–33]. Of the 144 barcoded LNPs that we formulated, 89 met quality control criteria and were pooled together and sterile-filtered. As a control, we plotted the hydrodynamic diameter of the LNP pool, which was within the range of the 89 individual LNPs, indicating that pooled LNPs did not come out of solution (Fig. 1i). We then analysed the hydrodynamic diameter as a function of ionizable lipid (Fig. 1j), cholesterol (Fig. 1k), PEG-lipid (Fig. 1l) and helper lipid (Fig. 1m), and found that LNP stability was not significantly impacted by any individual component.

We intravenously injected the LNP pool, containing aVHH mRNA and DNA barcodes, into six mouse strains at a total nucleic acid dose of 1.5 mg kg^{-1} (each particle was injected with an average of 0.017 mg total nucleic acid per mouse kg; Fig. 2a). We also injected control mice for each strain with phosphate-buffered saline (PBS). After 24 h, we isolated human- and NHP-repopulated (anti-human CD47$^+$) as well as murine resident (anti-mouse CD31$^-$CD45$^-$) hepatocytes from FRG mice using FACS (Supplementary Fig. 1). We denote human- and NHP-repopulated hepatocytes humanized/H and primatized/P, respectively, and murine hepatocytes found in human and primatized FRG mice are denoted humanized/M and primatized/M, respectively. Finally, mouse hepatocytes from murinized mice are denoted murinized/M (Fig. 2b,c). We observed a significant increase in the percentage of aVHH$^+$ humanized/H and primatized/P hepatocytes compared with humanized/M and primatized/M hepatocytes (Fig. 2d). These data compare hepatocytes isolated from the same animal, excluding the possibility that differences are due to tissue handling or missed injections. We then performed an in situ hybridization analysis of delivered aVHH mRNA and species-specific cell markers (Supplementary Fig. 2). Consistent with our aVHH flow cytometry data, we observed increased aVHH colocalization with human or NHP hepatocytes. We did not observe overt changes to liver tissue structure that would promote delivery to human and NHP cells, relative to mouse cells. As an additional control, we compared the percentage of aVHH$^+$ murinized/M and BL/6 hepatocytes and observed similar delivery profiles to humanized/M and primatized/M hepatocytes (Fig. 2d).

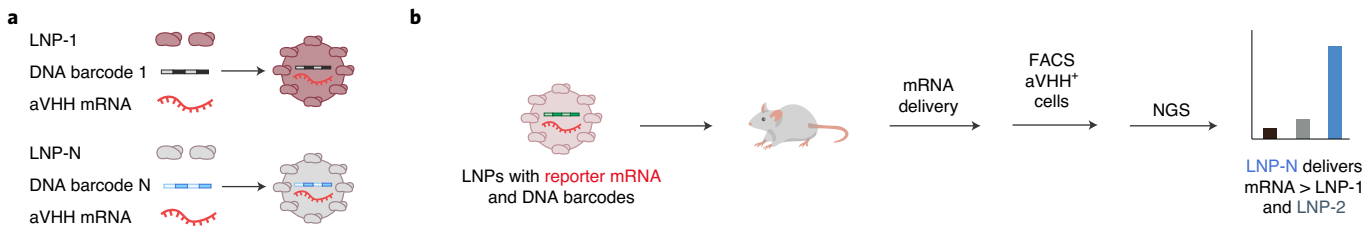
We then quantified how delivery varied across species by calculating the correlation between normalized delivery in humanized/H, primatized/P and murinized/M hepatocytes. Normalized delivery is a validated way to quantify LNP delivery using DNA barcodes[34]. As expected, unencapsulated barcodes entered cells less efficiently than barcodes formulated into LNPs (Fig. 2e). We then calculated the correlation of normalized delivery for primatized/P hepatocytes and humanized/H hepatocytes and found it to be high ($R^2 = 0.83$, Fig. 2f). The correlation between humanized/H and humanized/M hepatocytes was lower ($R^2 = 0.53$, Fig. 2g), even though these cells were taken from the same mice, excluding the possibility that species differences were caused by tissue handling. A similar correlation ($R^2 = 0.58$, Fig. 2h) was observed between primatized/P and primatized/M hepatocytes isolated from the same animals. We calculated correlations for humanized/H hepatocytes and four additional mouse strains, including murinized/M control hepatocytes; these correlations ranged between 0.31 and 0.64 (Supplementary Fig. 3a–h). These data suggest that delivery to humanized/H cells is best predicted by primatized/P cells, relative to murine cells.

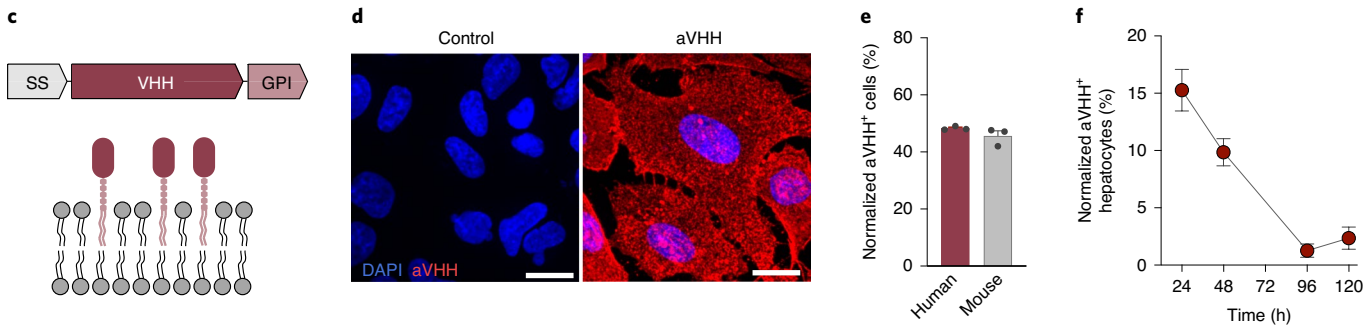
These correlations include all 89 LNPs. However, the best LNPs in mice were selected for NHP or human studies. We therefore compared delivery across species for top-performing LNPs. We generated a heatmap of normalized delivery for all 89 LNPs in humanized/M, primatized/P and humanized/H hepatocytes (Fig. 2i). We noted that five LNPs exhibited high delivery in humanized/M hepatocytes but low delivery in primatized/P or humanized/H hepatocytes. Additionally, we found three LNPs had high delivery in humanized/H and low delivery in humanized/M hepatocytes. This suggests that screening in mice can generate false positive LNPs (that is, LNPs predicted to deliver mRNA in NHP or human hepatocytes that fail to do so) as well as false negative LNPs. We then rank-ordered the top ten LNPs in humanized/M hepatocytes and plotted their normalized delivery in humanized/H hepatocytes (Fig. 2j). Some top LNPs performed well in murine and human hepatocytes. However, several LNPs performed well in murine hepatocytes but not in human hepatocytes. We similarly

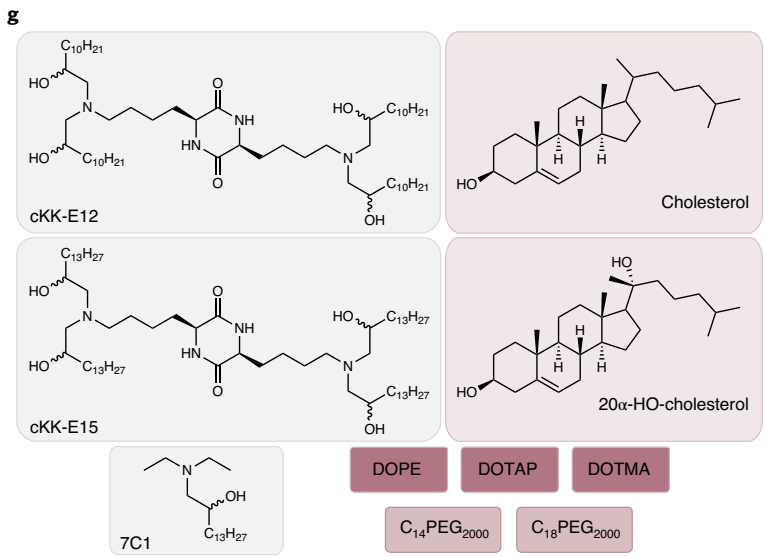
Fig. 1 | Characterizing SANDS. a, LNP-1 is formulated to carry mRNA encoding aVHH as well as DNA barcode 1, and LNP-N is formulated to carry mRNA encoding aVHH as well as DNA barcode N. **b**, Stable LNPs are pooled and injected into mice. aVHH$^+$ cells are later isolated by FACS, and the barcodes within the cells are identified using next-generation sequencing (NGS). **c**, aVHH was selected as reporter because it is embedded in the cell bilayer and can be quantified using antibodies against VHH. SS, signal sequence. **d**, aVHH protein expression, measured by immunofluorescent staining, 24 h after A549 cells were transfected with aVHH mRNA. DAPI, 4′,6-diamidino-2-phenylindole. Scale bars, 100 μm. **e**, aVHH protein expression, measured by flow cytometry, in Fa2N-4 human and AML-12 mouse hepatocytes 24 h after transfection with aVHH mRNA formulated with 0.5 μg PEI. $N = 3$ per group, average ± s.e.m. Paired *t*-test, not significant (n.s.) = 0.2474. **f**, The percentage of aVHH$^+$ hepatocytes (that is, CD31$^-$CD45$^-$aVHH$^+$) at different timepoints following treatment with LNPs carrying 1 mg kg^{-1} aVHH mRNA. $N = 3$ per group, average ± s.e.m. **g**, The combinatorial design of 144 chemically distinct LNPs formulated carrying aVHH and DNA barcodes. DOPE, 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine; DOTAP, 1,2-dioleoyl-3-trimethylammonium-propane; DOTMA, 1,2-di-*O*-octadecenyl-3-trimethylammonium propane; $C_{14}PEG_{2000}$, 1,2-dimyristoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[methoxy (polyethyleneglycol)-2,000]; $C_{18}PEG_{2000}$, 1,2-distearoyl-*sn*-glycero-3-phosphoethanolamine-*N*-[methoxy(polyethyleneglycol)-2,000]. **h**, Four molar ratios were chosen to vary LNP formulations. **i**, Hydrodynamic diameters of 89 LNPs that were deemed stable enough to pool together, as well as the diameter of the pooled LNP control (red sphere). **j**–**m**, Hydrodynamic diameters of the 89 LNPs plotted as a function of LNP chemical property: ionizable lipid (**j**), cholesterol (**k**), PEG-lipid (**l**) and helper lipid (**m**). $N = 3$ per group, average ± standard deviation.

compared primatized/P and humanized/H hepatocytes and found that NHP delivery was similar to human delivery (Fig. 2j). These analyses led us to conclude that murine hepatocytes do not predict delivery in human cells as well as NHP hepatocytes.

To confirm that SANDS can predict NHP or human results more accurately, we performed two control experiments. First, we analysed mRNA delivery mediated by the top LNP from our screen in primary human and mouse hepatocytes in vitro (Supplementary Fig. 3i). We found that this LNP preferentially delivered mRNA to primary human hepatocytes, relative to mouse hepatocytes. By contrast, a control transfection reagent did not show species-dependent delivery. Second, we found a Moderna patent comparing hepatocyte mRNA delivery in NHP livers using 'Lipid H' and the FDA-approved Dlin-MC3-DMA lipid formulation[10]. We therefore formulated a

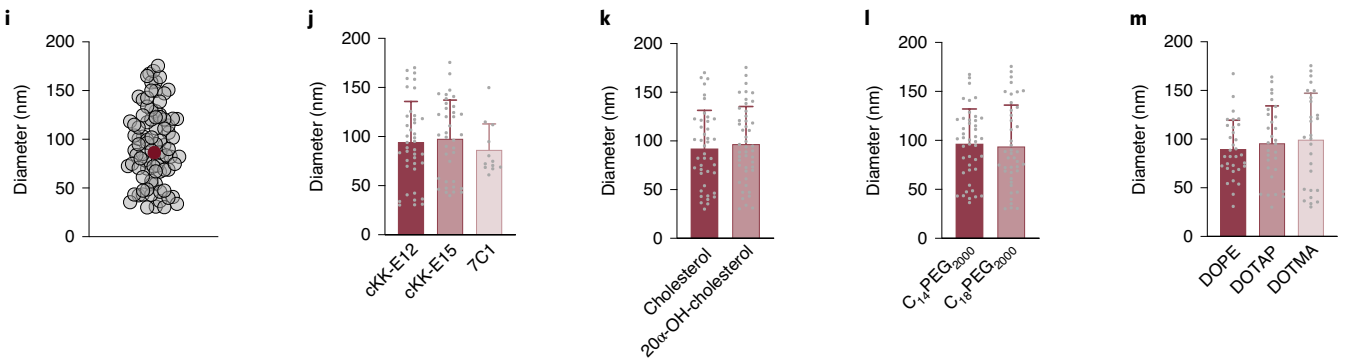






h

Molar ratio				
Ionizable lipid	60	50	45	35
Cholesterol	10	35	39.5	46.5
PEG-lipid	25	2.5	2.5	2.5
Phospholipid	5	12.5	13	16



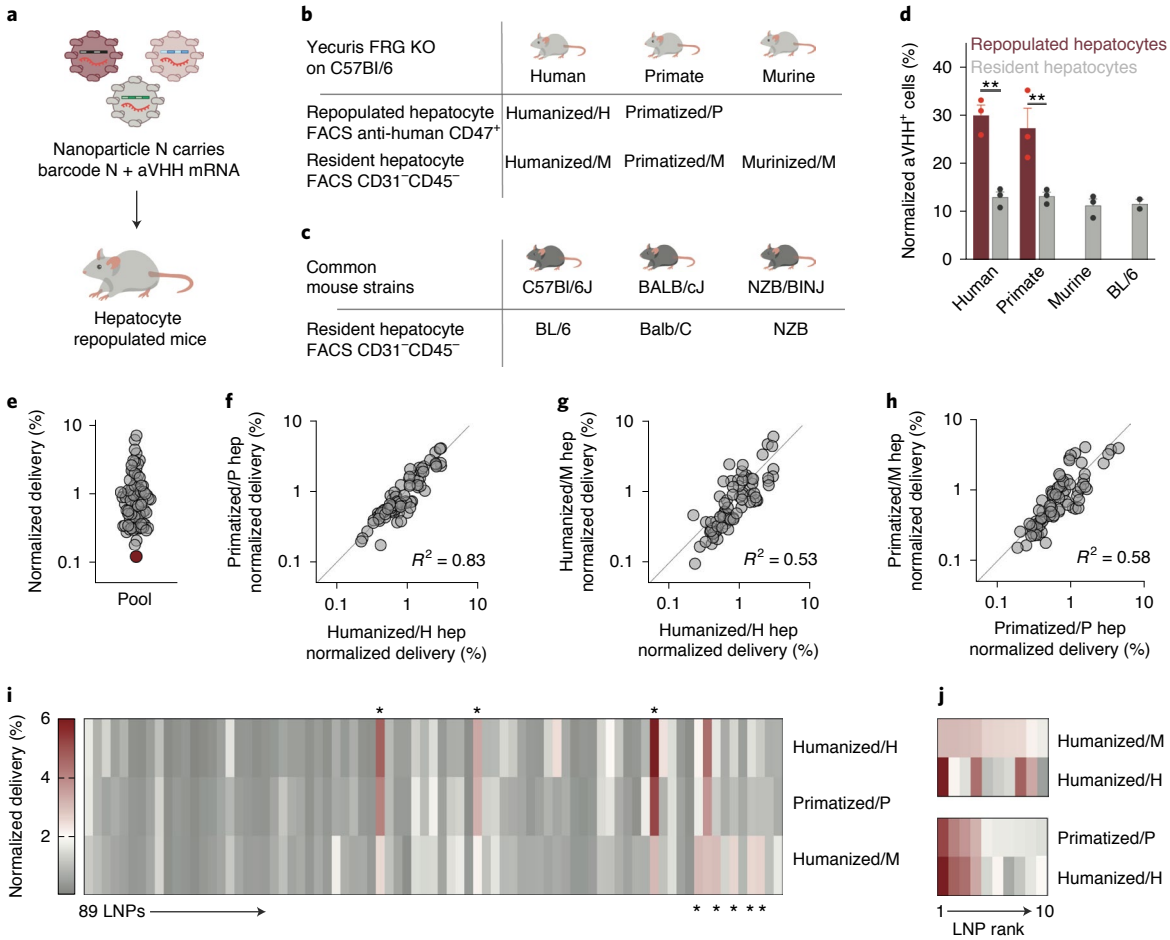


Fig. 2 | Nanoparticle delivery across murine, NHP and human hepatocytes in vivo. a–c, Eighty-nine LNPs labelled with SANDS were administered to six different murine models (**a**), including mice with humanized and primatized livers, as well as murinized control mice that underwent the same procedures as humanized mice (**b**) and three wild-type strains (**c**). **d**, Percentage of aVHH$^+$ hepatocytes across the three species and wild-type BL/6. Two-way analysis of variance: $^{**}P = 0.007$, n.s. = 0.9755. $N = 3$ per group, average ± s.e.m. **e**, Normalized delivery for all 89 LNPs, averaged across all samples. The unencapsulated DNA barcode, acting as a negative control (red sphere), was delivered into cells less efficiently than barcodes encapsulated by LNPs. **f–h**, R^2 analyses comparing delivery in human, NHP and murine hepatocytes: correlation of normalized delivery for primatized/P and humanized/H hepatocytes (**f**), humanized/H and humanized/M hepatocytes (**g**) and primatized/P and primatized/M hepatocytes (**h**). hep, hepatocyte. **i**, Heatmap of normalized delivery for all 89 LNPs in human, NHP and control mouse cells. The asterisks denote LNPs that delivered mRNA differently between the three species. **j**, Normalized delivery for the top ten-performing LNPs, ranked by delivery in murine or NHP cells. Delivery into human cells does not consistently align with murine delivery (top), but does align more consistently with NHP cells (bottom).

Moderna library and an MC3 library, using 32 LNPs per lipid, and compared delivery in humanized hepatocytes. The Moderna library resulted in twice as many aVHH$^+$ cells in humanized hepatocytes (Supplementary Fig. 3j), suggesting that SANDS can recapitulate reported datasets.

In addition to selecting individual LNPs for studies in larger animals, a second common practice is to use LNP delivery in mice to extract relationships between LNP chemical structures and in vivo delivery. We plotted normalized delivery for all 89 LNPs as a function of lipomer, cholesterol, PEG, helper lipid and molar ratio, and looked for relationships between LNP structure and normalized delivery across species. We did not observe species-dependent differences in cholesterol (Supplementary Fig. 3k). The relationship between helper lipid and delivery was also consistent across all three species (Supplementary Fig. 3l). We observed species differences in PEG-lipids: PEG-lipids with a C_{14} lipid tail tended to deliver barcodes to humanized/H hepatocytes significantly more than PEG-lipids with a C_{18} lipid tail. This difference was consistent but not significant in primatized/P hepatocytes, and was not observed in murinized/M hepatocytes (Supplementary Fig. 3m). Similarly, the type of lipomer that promoted delivery in murinized/M hepatocytes was distinct from the lipomer that promoted delivery in humanized/H and primatized/P hepatocytes (Supplementary Fig. 3n). We also found a statistical difference in molar ratios but were unable to clearly interpret these differences (Supplementary Fig. 3o–q). Additional studies are required to understand whether the PEG- and lipomer-based differences we observed are consistent across LNPs.

Species-dependent cellular responses to LNPs

We then returned to our observation that LNP delivery led to a higher percentage of aVHH$^+$ humanized/H and primatized/P hepatocytes, relative to humanized/M, primatized/M and murinized/M hepatocytes (Fig. 2d). We hypothesized that the difference in aVHH$^+$ hepatocytes was driven by overt differences in how much nucleic acid reached the cells. We quantified LNP biodistribution in aVHH$^+$ hepatocytes using droplet digital PCR (ddPCR); the barcodes included a ddPCR probe site for sensitive biodistribution

readouts in FACS-sorted cells[20]. The increase in aVHH$^+$ cells (Fig. 2d) was not observed when we measured biodistribution (Supplementary Fig. 3r), which led us to discard this hypothesis. We then posited that the species-dependent difference in the percentage of aVHH$^+$ cells was caused by species-dependent internal cell signalling, which can influence LNP mRNA delivery[27,35].

We compared the transcriptomic profiles of humanized/H and humanized/M cells. Using a modified Smart-seq2 protocol[36], we sequenced 1,000 cells from three pooled biological replicates (Fig. 3a). Counts were analysed using integrated Differential Expression and Pathway (iDEP) analysis, which allows for differential expression and pathway analysis of RNA sequencing (RNA-Seq) data[37] (Fig. 3b and Supplementary Fig. 4). Due to the high number of samples, biological replicates were pooled before sequencing, therefore increasing the total number of reads per condition. We also analysed the data using CORNAS, a tool specifically developed for analysing gene expression data without biological replicates[38]. We obtained similar results for both CORNAS and DESeq2 analyses. Using iDEP, we first performed a differential expression analysis (Fig. 3c) to determine the number of differentially expressed genes in humanized/H and humanized/M samples treated with either PBS or LNPs (Fig. 3c). We found 2.6-fold more differentially expressed genes (13,721) when comparing humanized/H hepatocytes treated with LNP, relative to their PBS control, than when we compared LNP-treated humanized/M hepatocytes, relative to their PBS control (5,338, Fig. 3c). These results provide one line of evidence that human and murine cells underwent different transcriptional responses to LNPs.

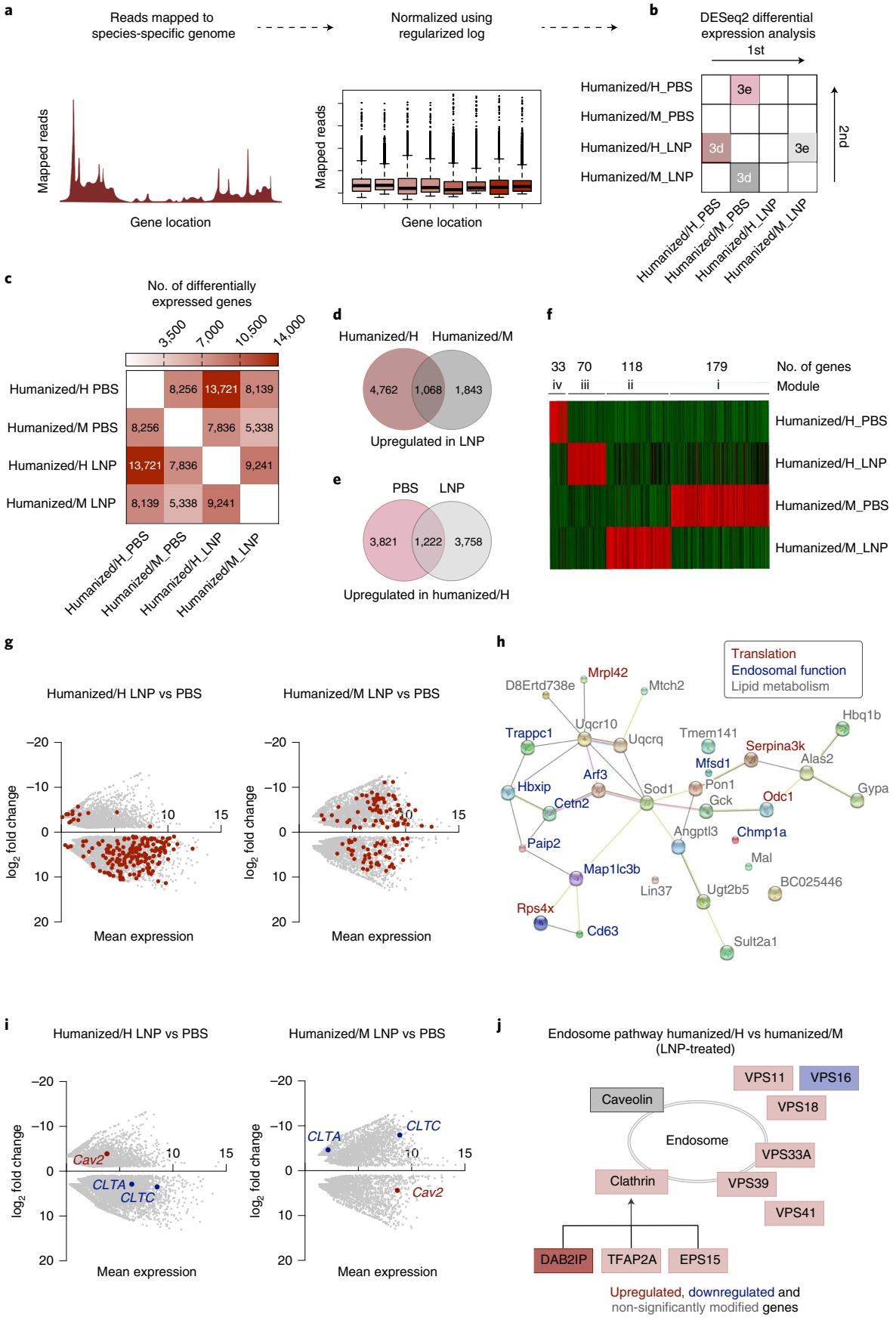
We noted that most of the differentially regulated genes, between the LNP- and PBS-treated samples, were upregulated (Supplementary Fig 4e). We therefore focused our analysis on those genes. We quantified the number of upregulated genes in LNP-treated humanized/M cells, relative to PBS-treated humanized/M cells (Fig. 3d), and found 2,911 upregulated genes. Notably, the number of upregulated genes in LNP-treated humanized/H cells, relative to PBS-treated humanized/H cells, was twofold higher (5,830). Of these upregulated genes, only 1,068 were shared, suggesting that upregulated genes differed between the two species. To control for basal species differences (that is, differences not driven by response to LNPs), we compared humanized/H PBS-treated hepatocytes with humanized/M PBS-treated hepatocytes. We found that 5,043 genes were upregulated in PBS-treated humanized/H hepatocytes. We then compared humanized/H LNP-treated hepatocytes with humanized/M LNP-treated hepatocytes and found that 4,980 genes were upregulated. We reasoned that if the differences in gene expression were driven by basal species differences, then most of the upregulated genes would be shared. However, of these upregulated genes, only 1,222 were shared (Fig. 3e). Finally, we confirmed that the expression of 80 housekeeping genes was comparable across samples (Supplementary Fig. 5). To validate that cells exhibited different transcriptional profiles, we performed a weighted correlation network analysis (WGCNA) using the 400 most upregulated genes for both humanized/H and humanized/M LNP-treated samples. This algorithm separated the 400 genes into four distinct modules, providing another line of evidence that murine and human hepatocytes responded differently to LNPs (Fig. 3f).

To understand the cellular functions of these genes, we analysed the top ten upregulated genes within the four WGCNA clusters. Five of the top ten upregulated genes in the LNP-treated humanized/H cluster were related to translation or post-translational processes, whereas none of the top ten humanized/M cluster genes were (Supplementary Fig. 4). To confirm this result, we entered read counts for all the genes into iDEP, for both humanized/H and humanized/M samples, with and without LNP treatment, and extracted a list of the most upregulated genes. We considered genes significant if they changed by at least twofold with a false discovery rate of less than 0.1. The most differentially regulated pathway between humanized/H and humanized/M cells, both treated with LNPs relative to their own PBS controls, was again translation (Fig. 3g and Supplementary Fig. 6a,b). Specifically, we found that 202 genes related to translation were upregulated in the LNP-treated humanized/H samples; 99 of these genes were upregulated in the LNP-treated humanized/M samples (Supplementary Fig. 6c).

Given that mRNA translation is a key cellular function, we reasoned that some aspects of the cellular response to LNPs would be conserved. Therefore, we entered the list of 99 genes upregulated in both humanized/H and humanized/M into the DAVID Bioinformatics database[39]. The genes upregulated in both conditions were related to eukaryotic translation initiation factor eIF2/eIF5 and the eukaryotic elongation factor 1 (eEF1), suggesting that exogenously delivered mRNA is impacted by similar initiation and elongation mechanisms in both species. We then entered the 103 genes that were only upregulated in human cells but were unable to identify specific aspects of translation that were clearly differentiated. We therefore explored interactions of these genes with other cellular functions. We used an unbiased Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) analysis to better understand how upregulated genes interact with each other. We took the top 30 upregulated genes in LNP-treated humanized/H and humanized/M hepatocytes[40] and input them in the database (Fig. 3h). As a control, we first investigated whether STRING analysis identified genes associated with translation, which it did. Further, the STRING analysis highlighted that those genes were interacting with other genes related to lipid metabolism and endosomal function, which may affect LNP delivery. Based on this, we evaluated the expression of *Cav2*, *CLTA* and *CLTC*, canonical genes associated with caveolin- and clathrin-mediated endocytosis, which LNPs use to enter cells. We found that *CLTA* and *CLTC* were upregulated in humanized/H hepatocytes treated with LNPs, relative to PBS, whereas *Cav2* was downregulated (Fig. 3i). However, we observed the opposite expression profile in humanized/M hepatocytes (Fig. 3i). These data highlight that canonical clathrin- and caveolin-mediated endocytosis genes were not regulated similarly in murine and human

Fig. 3 | Transcriptomic studies reveal species-dependent response to LNPs. a, Schematic of the Smart-seq2 protocol design. **b**, DESeq2 differential expression analysis uses a pairwise comparison to determine the number of differentially expressed genes between different conditions. Smart-seq2 on humanized mice was repeated twice, once with biological replicates pooled and once separated. **c**, DESeq2 differential expression analysis reveals that LNP-treated humanized/H contains a higher number of differentially expressed genes. **d,** The number of upregulated genes was determined between humanized/H and humanized/M species after LNP treatment relative to PBS-treated cells. **e**, The number of upregulated genes was determined between LNP- and PBS-treated humanized/H samples relative to the corresponding humanized/M cells. **f**, WGCNA analysis shows the 400 most upregulated genes in humanized/H and humanized/M hepatocytes: four unique modules were identified. **g**, Differential gene expression in the LNP-treated humanized/H and humanized/M hepatocytes was identified relative to PBS-treated hepatocytes. Genes associated with mRNA translation are highlighted in red. **h**, STRING analysis, based on the top 30 upregulated genes, across the LNP-treated humanized/H and humanized/M samples identified interactions between genes associated with translation, lipid metabolism and endosomal function. *P* value on upregulated genes = 3.07×10^{-4}. **i**, Differential expression of genes associated with caveolin- and clathrin-mediated endocytosis in both humanized/H and humanized/M hepatocytes indicates receptor-mediated endocytosis-related difference in delivery across species. **j**, As an additional line of evidence, the Reactome database was used to illustrate the differences in endocytosis-related gene expression between LNP-treated humanized/H and humanized/M hepatocytes.

hepatocytes after LNP exposure. To confirm these data, we used the Reactome database to analyse receptor-mediated endocytosis profiles in LNP-treated humanized/H and humanized/M samples. The Reactome data corroborated that humanized/H hepatocytes use clathrin-mediated endocytosis more predominantly than caveolin, relative to humanized/M hepatocytes after LNP exposure (Fig. 3j).



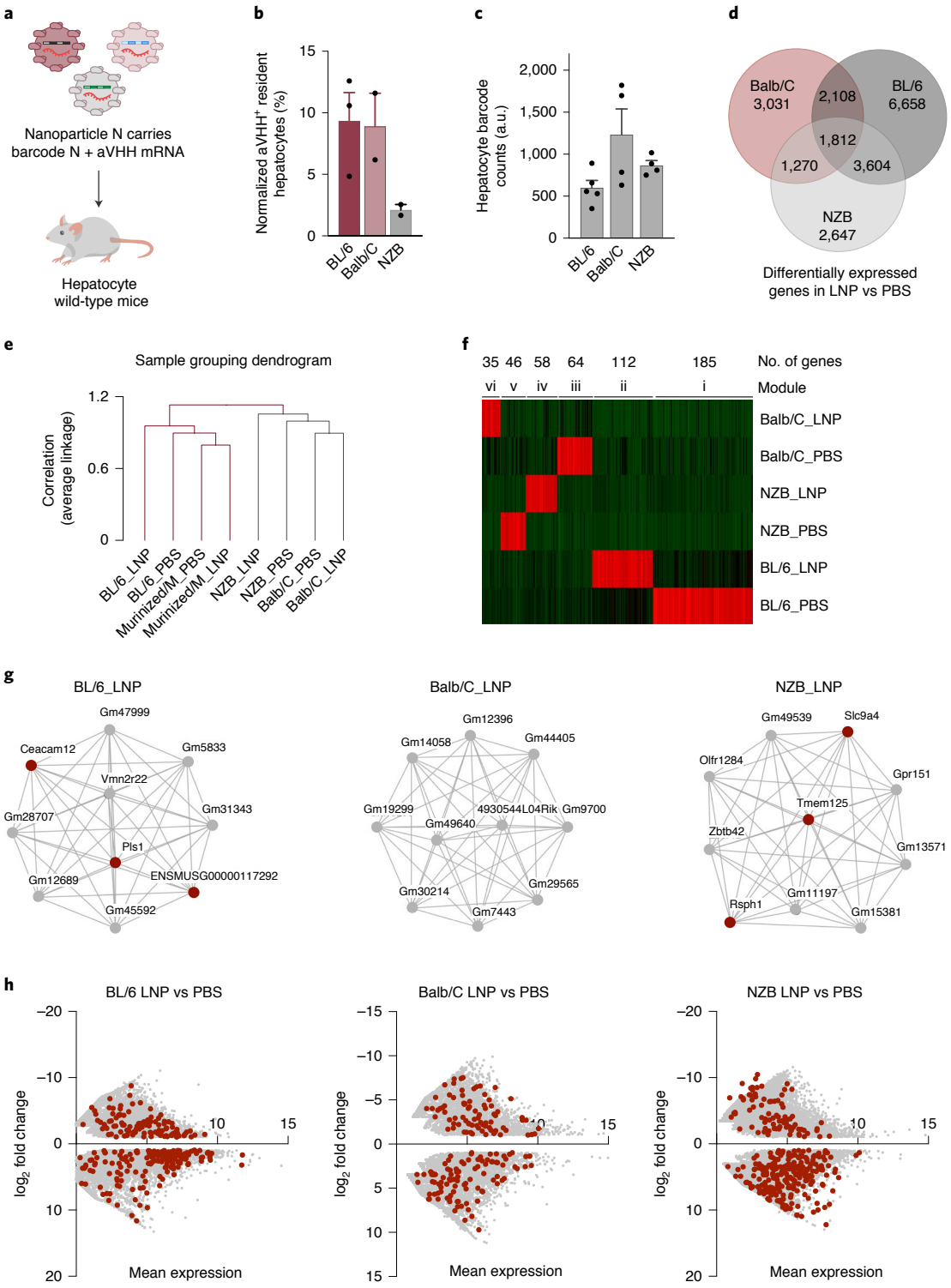


Fig. 4 | Inflammatory genes impact mRNA delivery across multiple mouse strains. a, The 89 LNPs previously described, carrying a DNA barcode and aVHH mRNA, were intravenously administered to wild-type BL/6, Balb/C and NZB mice. The Smart-seq2 protocol on wild-type mice was not repeated. **b**, The percentage of aVHH⁺ hepatocytes was analysed after 24 h. $N = 3$ per group, average ± s.e.m. **c**, The biodistribution of DNA barcodes was quantified using ddPCR. $N = 3$ per group, average ± s.e.m. **d**, The number of upregulated genes was determined between the different mouse strains after LNP treatment. **e**, Hierarchical clustering analysis was performed, indicating different gene expression profiles between Balb/C, NZB and BL/6 mice. **f**, WGCNA analysis shows the 500 most upregulated genes in wild-type murine hepatocytes: six unique modules were identified. **g**, Network analysis of the top ten upregulated genes found in the WGCNA analysis for each cluster of LNP-treated BL/6 (module ii in **f**), Balb/C (module vi) and NZB (module iv) mice. Genes associated with immune and disease responses are highlighted in red. **h**, Differential gene expression between the LNP- and PBS-treated samples was analysed for all three strains. Genes associated with the immune response are highlighted in red.

To confirm our conclusions, we repeated the experiment in humanized mice, keeping biological replicates separated during sequencing (Supplementary Figs. 7–9). We compared the transcriptomic profiles of humanized/H and humanized/M cells using the same modified Smart-seq2 protocol. We first performed a genome-wide analysis and compared overall gene expression levels between each set of conditions (pooled and separated) and found them to be similar ($R^2 = 0.89 \pm 0.047$, Supplementary Fig. 7). We then compared the expression level of genes related to caveolin- and clathrin-mediated endocytosis in both LNP-treated humanized/H and humanized/M with their respective PBS-treated samples (Supplementary Fig. 8b,c). We again observed that *CLTC* and *CLTA* genes were upregulated in LNP-treated humanized/H samples, relative to PBS, whereas *Cav2* was downregulated. The opposite trend was observed in LNP-treated humanized/M samples, relative to PBS. Lastly, we performed a STRING analysis to study the top ten upregulated genes in LNP-treated humanized/H and humanized/M hepatocytes, and found that four of those ten genes were also observed in the previous experiment (Supplementary Fig. 8e). These genes were found to be associated with translation, lipid metabolism and endosomal pathways, consistent with our previous experiment.

Based on these data, we concluded that the species-dependent increase in aVHH$^+$ humanized/H hepatocytes is potentially driven by a combination of two non-mutually exclusive mechanisms. The first mechanism is a species-dependent reliance on either caveolin- or clathrin-mediated endocytosis. Although both pathways are present in both species, the in vivo activity of one pathway relative to the other may not be conserved. The second potential mechanism stems from differences in translation that are still to be determined. We cannot identify the specific translational mechanism; however, our data do not support the hypothesis that differences are driven by initiation or elongation.

Strain-dependent cellular responses to LNPs

One limitation to several of the unbiased RNA-Seq analyses above is that some cellular functions (for example, translation) may have more genes associated with them and thus may come up more often as gene lists are surveyed. To control for this, we performed the same transcriptomic analyses but removed the species variable by comparing the delivery with three wild-type immunocompetent mouse strains. Using the same pool of 89 LNPs, we injected wild-type BL/6, Balb/C and NZB mice (Fig. 4a) and found a lower percentage of aVHH$^+$ hepatocytes in NZB mice compared with in BL/6 and Balb/C (Fig. 4b). We quantified the biodistribution in hepatocytes using ddPCR and found that the biodistribution in aVHH$^+$ cells between mice strains was similar and, once again, did not overlay with the percentage of aVHH$^+$ hepatocytes (Fig. 4c).

To analyse whether this difference in delivery was driven by the same cellular mechanisms observed between human and murine, we isolated hepatocytes and examined their transcriptomic profiles. We first quantified the number of differentially expressed genes and found differences in gene expression between the three mouse strains treated with LNPs, relative to their PBS controls (Fig. 4d). To better understand these differences and how they relate to each sample, we used hierarchical clustering to determine how similar each strain was to the others. We found that NZB and Balb/C mice clustered together, whereas BL/6 mice formed a separate cluster (Fig. 4e). As a control, we confirmed that murinized/M, which is on a BL/6 background, and wild-type BL/6 clustered together (Fig. 4e). To confirm different expression profiles between the three wild-type mice strains, we performed a WGCNA analysis and found unique gene clusters for each strain (Fig. 4f). We then looked at the top ten genes found within each cluster: most of the genes were unannotated and several of the annotated genes were related to immune and disease responses (Fig. 4g). We did not identify pathways for translation or endocytosis. To further explore the differences between mouse strains, we performed pathway analysis of all the genes using iDEP and found far more upregulated genes associated with immune response in BL/6 (273 genes) and NZB (324 genes) mice than in Balb/C (178 genes) mice (Supplementary Fig. 10). We confirmed these results by comparing MA plots of the differentially expressed genes in LNP-treated samples, relative to PBS, which highlighted different gene expression profiles between the three strains (Fig. 4h). These results implicate differential immune response as a potential mechanism for strain-dependent delivery, and these data led us to conclude that the translation and endocytosis differences observed in our species-to-species studies were not identified solely because these processes are well annotated in RNA-Seq analytical software.

Conclusions

SANDS may ease the hard-to-predict progression of drug delivery candidates from mice to NHPs to humans. NHP hepatocytes predicted delivery to human hepatocytes better than murine hepatocytes; however, murine hepatocytes were somewhat predictive. As a result, our results do not justify going straight into NHP one-by-one LNP screening, which is unethical. Instead, our data support a two-step model for selecting LNPs in mice. In the first step, dozens of LNPs are screened at once in humanized mice. In the second step, the top LNPs are tested individually in humanized mice and immunocompetent mice. We propose that in the individual LNP studies, the efficacy data in humanized mice supersede the efficacy data from wild-type mice, and the safety data in wild-type mice supersede the safety data in humanized mice.

Given that SANDS is species-agnostic, hundreds of LNPs in combinations of animal models, including those with disease phenotypes, may be evaluated in future studies. This confers two advantages. The first is the ability to screen in animals known to better recapitulate human physiology. The second advantage is the ability to test whether disease-specific genes influence delivery. In this study, we identified translation and endocytosis as two cellular processes that may respond to LNPs in a species-dependent manner. Notably, although the data were consistent across several experiments and analyses, some data were generated with a pooled sample of cells rather than biological replicates. These data were analysed with software specifically designed for pooled samples but may still be affected by biological variability and dispersion. If validated by other laboratories, then mice engineered to overexpress endocytosis genes used by human hepatocytes may be useful models for drug delivery: we envision mice with (1) a fully intact murine immune system and (2) hepatocytes with 'human-like' endocytosis profiles. Finally, it is important to note that we would have generated a different biological hypothesis if we had only studied LNPs using murine hepatocytes.

It is important to acknowledge the limitations of the study. One limitation is the use of humanized and primatized mice as proxies for humans and NHPs. One alternative is to screen directly in wild-type mice and NHPs, then compare results; however, this alternative would not include human cells. A second limitation is the use of LNPs we found likely to target hepatocytes using endogenous trafficking pathways. Other LNP chemical spaces may exhibit different species-to-species variability. A third limitation is our focus on hepatocytes; the mouse–NHP–human predictivity may change as a function of cell type. Despite these limitations, we believe this work may help streamline the preclinical development of RNA delivery vehicles.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of

data and code availability are available at https://doi.org/10.1038/s41565-021-01030-y.

Received: 13 July 2020; Accepted: 11 October 2021;
Published online: 07 February 2022

References

1. Adams, D. et al. Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. *N. Engl. J. Med.* **379**, 11–21 (2018).
2. Garrelfs, S. F. et al. Lumasiran, an RNAi therapeutic for primary hyperoxaluria type 1. *N. Engl. J. Med.* **384**, 1216–1226 (2021).
3. Balwani, M. et al. Phase 3 trial of RNAi therapeutic givosiran for acute intermittent porphyria. *N. Engl. J. Med.* **382**, 2289–2301 (2020).
4. Ray, K. K. et al. Two phase 3 trials of inclisiran in patients with elevated LDL cholesterol. *N. Engl. J. Med.* **382**, 1507–1519 (2020).
5. Finkel, R. S. et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1723–1732 (2017).
6. Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2020).
7. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
8. Sahin, U. et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **547**, 222–226 (2017).
9. 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).
10. Akinc, A. et al. The Onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. *Nat. Nanotechnol.* **14**, 1084–1087 (2019).
11. Akinc, A. et al. Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms. *Mol. Ther.* **18**, 1357–1364 (2010).
12. 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).
13. Lisowski, L. et al. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* **506**, 382–386 (2014).
14. Paulk, N. K. et al. Bioengineered AAV capsids with combined high human liver transduction in vivo and unique humoral seroreactivity. *Mol. Ther.* **26**, 289–303 (2018).
15. Vercauteren, K. et al. Superior in vivo transduction of human hepatocytes using engineered AAV3 capsid. *Mol. Ther.* **24**, 1042–1049 (2016).
16. Pei, X. et al. Development of AAV variants with human hepatocyte tropism and neutralizing antibody escape capacity. *Mol. Ther. Methods Clin. Dev.* **18**, 259–268 (2020).
17. Wilson, E. M. et al. Extensive double humanization of both liver and hematopoiesis in FRGN mice. *Stem Cell Res.* **13**, 404–412 (2014).
18. Foquet, L. et al. Successful engraftment of human hepatocytes in uPA-SCID and FRG® KO mice. *Methods Mol. Biol.* **1506**, 117–130 (2017).
19. 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).
20. Sago, C .D. et al. Modifying a commonly expressed endocytic receptor retargets nanoparticles in vivo. *Nano Lett.* **18**, 7590–7600 (2018).
21. Sago, C. D. et al. Nanoparticles that deliver RNA to bone marrow identified by in vivo directed evolution. *J. Am. Chem. Soc.* **140**, 17095–17105 (2018).
22. Sago, C. D. et al. High-throughput in vivo screen of functional mRNA delivery identifies nanoparticles for endothelial cell gene editing. *Proc. Natl Acad. Sci. USA* **115**, E9944–E9952 (2018).
23. Tiwari, P. M. et al. Engineered mRNA-expressed antibodies prevent respiratory syncytial virus infection. *Nat. Commun.* **9**, 3999 (2018).
24. Paunovska, K. et al. Nanoparticles containing oxidized cholesterol deliver mRNA to the liver microenvironment at clinically relevant doses. *Adv. Mater.* **31**, e1807748 (2019).
25. Dong, Y. et al. Lipopeptide nanoparticles for potent and selective siRNA delivery in rodents and nonhuman primates. *Proc. Natl Acad. Sci. USA* **111**, 3955–3960 (2014).
26. Dahlman, J. E. et al. In vivo endothelial siRNA delivery using polymeric nanoparticles with low molecular weight. *Nat. Nanotechnol.* **9**, 648–655 (2014).
27. Lokugamage, M. P. et al. Mild innate immune activation overrides efficient nanoparticle-mediated RNA delivery. *Adv. Mater.* **32**, 1904905 (2020).
28. Paunovska, K. et al. Analyzing 2000 in vivo drug delivery data points reveals cholesterol structure impacts nanoparticle delivery. *ACS Nano* **12**, 8341–8349 (2018).
29. Patel, S. et al. Naturally-occurring cholesterol analogues in lipid nanoparticles induce polymorphic shape and enhance intracellular delivery of mRNA. *Nat. Commun.* **11**, 983 (2020).
30. 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).
31. Kaczmarek, J. C. et al. Optimization of a degradable polymer-lipid nanoparticle for potent systemic delivery of mRNA to the lung endothelium and immune cells. *Nano Lett.* **18**, 6449–6454 (2018).
32. Kranz, L. M. et al. Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. *Nature* **534**, 396–401 (2016).
33. 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).
34. Lokugamage, M. P., Sago, C. D. & Dahlman, J. E. Testing thousands of nanoparticles in vivo using DNA barcodes. *Curr. Opin. Biomed. Eng.* **7**, 1–8 (2018).
35. Patel, S. et al. Boosting intracellular delivery of lipid nanoparticle-encapsulated mRNA. *Nano Lett.* **17**, 5711–5718 (2017).
36. Picelli, S. et al. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
37. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **19**, 534 (2018).
38. Low, J. Z. B., Khang, T. F. & Tammi, M. T. CORNAS: coverage-dependent RNA-Seq analysis of gene expression data without biological replicates. *BMC Bioinformatics* **18**, 575 (2017).
39. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
40. Szklarczyk, D. et al. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

Methods

Synthesis of lipomers. Lipomers 7C1, cKK-E12 and cKK-E15 were prepared according to literature procedures[25–27].

Preparation of 7C1. C_{13} lipids and PEI_{600} were combined and heated at 90 °C in EtOH for 48–72 h.

Preparation of cKK-E12. 2,5-dioxopyrrolidin-1-yl N6-((benzyloxy)carbonyl)-N2-(tert-butoxycarbonyl)-L-lysinate (compound **1**) (20 g, 41.9 mmol) was added to a 100-ml flask and trifluoroacetic acid (42 ml) was added slowly at 0 °C, then stirred at room temperature for 30 min. The solvent was evaporated under reduced pressure and the crude product, dissolved in *N*,*N*-dimethylformamide (DMF; 5 ml), was added dropwise to pyridine (300 ml) at 0 °C. The reaction mixture was then stirred at room temperature overnight. The solvents were evaporated under reduced pressure and washed with ethyl acetate to give pure dibenzyl (((2S,5S)-3,6-dioxopiperazine-2,5-diyl)bis(butane-4,1-diyl))dicarbamate (compound **2**) (8.4 g, 31% yield). To a solution of compound **2** in acetic acid–CH_2Cl_2 (1:1, 300 ml) was added Pd/C (10 wt%, 3.0 g). The black suspension was degassed for 5 min with hydrogen and then stirred at room temperature overnight under a hydrogen atmosphere. The reaction mixture was filtered through Celite and washed with MeOH. The combined filtrates were then concentrated to obtain a crude yellow viscous oil. The oil was washed with ethyl acetate to yield (3S,6S)-3,6-bis(4-aminobutyl)piperazine-2,5-dione (compound **3**), which was used without further purification. Next, triethylamine (0.12 ml, 0.88 mmol) was added to a solution of compound **3** (84 mg, 0.22 mmol) and 1,2-epoxydodecane (247 mg, 1.34 mmol) in EtOH (2 ml). After stirring for 30 min at room temperature, the reaction mixture was then irradiated in a microwave reactor at 150 °C for 5 h. Purification of the crude residue by flash column chromatography (gradient eluent: MeOH–CH_2Cl_2 20:1 to 7:1 at 1 column volume (CV)) afforded cKK-E12 (219 mg, 40% yield) as a yellow oil. ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 4.02$ (br. s, 2H), 3.71 (br. s, 4H), 3.06–2.19 (m, 12H), 1.86 (br. s, 4H), 1.43–1.21 (m, 80H), 0.87 ppm (t, $J = 8$ Hz, 12H). HRMS (ESI): *m*/*z* calcd $[M+2H]^{2+}$ for $C_{60}H_{121}O_6N_4$ 497.4677, found 497.4662; *m*/*z* calcd $[M+H]^+$ for $C_{60}H_{121}O_6N_4$ 993.9272, found 993.9281.

Preparation of cKK-E15. Compound **1** (20 g, 41.9 mmol) was added to a 100-ml flask and trifluoroacetic acid (42 ml) was added slowly at 0 °C and 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 °C. The reaction mixture was stirred at room temperature overnight. The solvents were evaporated under reduced pressure and the crude product washed with ethyl acetate to give pure compound **2** (8.4 g, 31% yield). To a solution of compound **2** in acetic acid–CH_2Cl_2 (1:1, 300 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 overnight under a hydrogen atmosphere. The reaction mixture was filtered through Celite and washed with MeOH. The combined filtrates were concentrated and the crude product was washed with ethyl acetate to yield compound **3** (4.8 g, 98% yield). Next, triethylamine (0.12 ml, 0.88 mmol) was added to a solution of compound **3** (84 mg, 0.22 mmol) and tridecyloxirane (302 mg, 1.34 mmol) in EtOH (2 ml). The reaction mixture was then irradiated in a microwave reactor at 150 °C for 5 h. Purification of the crude residue by flash column chromatography (gradient eluent: 1.0–2.0% MeOH–CH_2Cl_2, then 2.0–4.0% MeOH–CH_2Cl_2 containing 0.5% NH_4OH at 1 CV) afforded cKK-E15 (200 mg, 78% yield) as a light-yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$): $\delta = 0.87$ (t, $J = 6.5$ Hz, 12H), 1.22–1.32 (m, 88H), 1.36–1.56 (m, 16H), 1.73–1.99 (m, 4H), 2.17–2.68 (br. s, 12H), 3.62 (m, 4H), 3.98 ppm (m, 2H). HRMS (ESI): *m*/*z* calcd $[M+H]^+$ for $C_{72}H_{145}N_4O_6$ 1162.1159, found 1162.1153.

aVHH mRNA synthesis. mRNA was prepared as previously described[23]. Briefly, the GPI-anchored VHH sequence was ordered as a DNA gBlock from Integrated DNA Technologies (IDT) including a 5′ untranslated region (5′-UTR) with Kozak sequence, a 3′-UTR derived from the mouse α-globin sequence and extensions to allow for Gibson assembly. The sequence was human codon-optimized using the IDT website. The gBlock was then cloned into a PCR-amplified pMA7 vector through Gibson assembly using NEB Builder with three molar excesses of insert. Gibson assembly reaction transcripts were purified using a 0.8% agarose gel prior to the assembly reaction. Subsequent plasmids from each colony were Sanger-sequenced to confirm sequence identity. Plasmids were digested into a linear template using NotI-HF (New England Biolabs, NEB) overnight at 37 °C. Linearized templates were purified using ammonium acetate (Thermo Scientific) precipitation and resuspended with nuclease-free water. In vitro transcription was achieved overnight at 37 °C using the HiScribe T7 kit (NEB), following the manufacturer's directions (full replacement or uracil with N^1-methylpseudouridine). The RNA product was then treated with Dnase I (Aldevron) for 30 min, removing the template, and purified using lithium chloride precipitation (Thermo Fisher Scientific). RNA transcripts were heat-denatured at 65 °C for 10 min and capped with a Cap1 structure using guanylyl transferase (Aldevron) and 2′-*O*-methyltransferase (Aldevron). The transcripts were then polyadenylated enzymatically (Aldevron). Finally, mRNA was purified by lithium chloride precipitation, treated with alkaline phosphatase (NEB) and purified once more. Concentrations were measured using a NanoDrop (Thermo Scientific), with mRNA stock concentrations in the range 2–4 mg ml^{-1}. The mRNA stock solutions were stored at −80 °C. Purified RNA products were analysed by gel electrophoresis to confirm purity.

In vitro assay. A549 cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum and 5% penicillin–streptomycin. The A549 cells were plated in 24-well plates and transfected with 1 μg aVHH mRNA per well or left untransfected. mRNA was formulated with Lipofectamine MessengerMax at a ratio of 1.5 μl μg^{-1}. After incubation for 24 h, the cells were fixed with 4% paraformaldehyde and immunostained. Briefly, the cells were permeabilized with 0.2% Triton X-100 in 1× PBS, followed by non-specific blocking with 5% bovine serum albumin. The cells were then incubated with primary rabbit anti-VHH 647 antibody (Genscript, cat. no. A01994, 1:1,000) for 30 min. The cells were washed with PBS, and the nuclei were counterstained with DAPI before being mounted with Prolong Gold antifade reagent. Finally, the cells were imaged using a Perkin Elmer UltraVIEW spinning disk confocal microscope with a Zeiss ×63 numerical aperture (NA) 1.4 Plan-Apochromat objective lens with Volocity software. The images were linearly contrast-enhanced equally across all images in Volocity.

Nanoparticle formulation. Nanoparticles were formulated with a microfluidic device as previously described[19]. Nucleic acids (DNA barcodes and mRNA) were diluted in 10 mM citrate buffer (Teknova). Lipid-amine compounds (7C1, cKK-E12 and cKK-E15), PEG-lipids ($C_{14}PEG_{2000}$ and $C_{18}PEG_{2000}$), cholesterols (cholesterol and 20α-hydroxycholesterol) and helper lipids (DOPE, DOTAP and DOTMA) were diluted in 100% ethanol. For the mRNA screens, aVHH mRNA and the DNA barcodes were mixed in a 10:1 mass ratio. All PEGs, cholesterols and helper lipids were purchased from Avanti Lipids. The citrate and ethanol phases were combined in a microfluidic device using syringes (Hamilton Company) at a flow rate of 3:1.

DNA barcoding. Each LNP was formulated to carry its own unique DNA barcode. The DNA barcodes were designed rationally with several characteristics as previously described[20]. All DNA barcodes were 91-nucleotide-long, single-stranded DNA sequences purchased from IDT. Briefly, the following modifications were made to all barcodes: (1) the nucleotides on the 5′ and 3′ ends were modified with phosphorothioates to reduce exonuclease degradation, (2) universal forward and reverse primer regions were included to ensure equal amplification of each sequence, (3) seven random nucleotides were included to monitor PCR bias, (4) a ddPR probe site was included for ddPCR compatibility and (5) a unique eight-nucleotide barcode was inserted. An eight-nucleotide sequence can generate over 4^8 (65,536) distinct barcodes. We used only the eight-nucleotide sequences designed to prevent sequence bleaching and reading errors on the Illumina MiniSeq sequencing machine.

Nanoparticle characterization. High-throughput dynamic light scattering (DynaPro Plate Reader II, Wyatt) was used to measure LNP hydrodynamic diameters. LNPs were first diluted in 1× PBS. LNPs were included only if they met two criteria: a diameter between 20 and 200 nm, and a correlation function with one inflection point. Particles that met these criteria were pooled and dialysed in a 20 kD dialysis cassette (Thermo Scientific), followed by a second dialysis in a 100 kD cassette (Spectrum) in 1× PBS. The nanoparticle concentration was determined using NanoDrop (Thermo Scientific).

Animal experiments. Liver xenograft FRG mice were generated and provided by Yecuris (Portland, OR), as previously described[41]. FRG KO on BL/6 mice were repopulated with cryo-preserved hepatocytes from either BL/6 mice (murinized), rhesus macaques (primatized) or a human donor (HHM01008, a male 1-year-old donor). Balb/C (BALB/cJ), BL/6 (C57Bl/6J) and NZB (NZB/BlNJ) mice were acquired from Jackson Laboratories. All animal experiments were performed in accordance with the Georgia Institute of Technology Institutional Animal Care and Use Committee's policies and procedures. All animals were housed in the Georgia Institute of Technology Animal Facility. All mice were injected intravenously in the lateral tail vein with LNPs or 1× PBS.

Cell isolation and staining. Cells were isolated 24 h after LNP injection unless otherwise noted. Mice were perfused through the right atrium with 20 ml of 1× PBS. Tissues were minced and placed in a digestive enzyme solution containing collagenase type I (Sigma Aldrich), collagenase XI (Sigma Aldrich) and hyaluronidase (Sigma Aldrich) at 37 °C at 550 r.p.m. for 45 min. The cell suspension was filtered through 70 μm mesh and red blood cells were lysed. The cells were stained and sorted using the BD FacsFusion cell sorter at the Georgia Institute of Technology Cellular Analysis Core. The antibody clones used were anti-CD31 (390, BioLegend), anti-CD45.2 (104, BioLegend), anti-CD47 (CC2C6, BioLegend) and MonoRab anti-camelid VHH (96A3F5, GenScript). Representative flow gates can be found in Supplementary Fig. 1. PBS-injected mice were used to gate aVHH positive populations.

Biodistribution assay. Biodistribution assays were executed using ddPCR as previously described[20]. To summarize, DNA samples were prepared with 10 μl

of ddPCR with ddPCR Supermix for Probes (Bio-Rad), 1 µl of primer and probe mix (solution of 10 µM target probe and 20 µM reverse/forward primers), 1 µl of template–Tris-EDTA buffer and 8 µl of water. Next, 20 µl of each reaction described and 70 µl of Droplet Generation Oil for Probes (Bio-Rad) were loaded into DG8 cartridges (Bio-Rad) and covered with DG8 gaskets (Bio-Rad). Water–oil emulsion droplets were created using a QX200 droplet generator (Bio-Rad). The cycle conditions for PCR analysis were as follows: 1 cycle of 95 °C for 10 min followed by 40 cycles of 94 °C for 30 s and 60 °C for 1 min, and 1 cycle of 95 °C for 10 min. For each biological repetition, two to three technical repetitions were completed. Unless stated otherwise, technical repetitions were averaged. Technical repetitions were only excluded if saturation was detected or there were inconsistent positive event amplitudes. The QX200 Droplet Digital PCR System (Bio-Rad) was used to analyse all ddPCR results.

PCR amplification. All samples were amplified and prepared for sequencing using nested PCR. More specifically, 1 µl of primers (5 µM for final reverse/forward) were added to 5 µl of Kapa HiFi 2X master mix (Roche) and 4 µl of template DNA–water. During the second PCR Nextera XT chemistry, indices and i5/i7 adapter regions were added. Dual-indexed samples were run on a 2% agarose gel to ensure that the PCR reaction had occurred before being pooled and gel-purified.

Deep sequencing. PCR samples were purified using AMPure XP beads. Final library quality control was conducted using an Agilent Bioanalyser 2100 system. Illumina deep sequencing was conducted using the Illumina MiniSeq system. Primers were designed based on Nextera XT adapter sequences.

Nanoparticle data analysis and statistics. A custom Python-based tool was used to process sequencing results and extract raw barcode counts for each sample. Normalization of these raw counts was achieved with R script (https://github.com/Jack-Feldman/barcode_count) prior to further analysis. Counts for each particle, per tissue, were then normalized to the barcoded LNP mixture injected into the mouse. Statistical analysis was performed using GraphPad Prism 8. Data were plotted as mean ± s.e.m. unless stated otherwise.

Bulk RNA sequencing preparation. A total of 10,000 cells were FACS-sorted based on tissue and species-specific markers. Immediately after sorting, samples were flash-frozen in an acetone–dry ice bath. Once all samples had been collected, they were thawed on ice, lysed with 10× lysis buffer and vortexed for 30 s. Then, 1,000 cell equivalents of each biological replicate were pooled together and 2 µl (1,000 cells of the pool) of this mixture was used following the Smart-seq2 protocol as previously described[42]. Each sample was normalized and pooled using Qubit, and the library size was determined using a Bioanalyser. Samples were sequenced using a NextSeq 550 instrument with a 150-cycle paired-end high-throughput cartridge.

Bulk RNA sequencing analysis and statistics. Fastq files were trimmed and filtered based on quality and then mapped using the Spliced Transcripts Alignment to a Reference aligner, and exon count tables were created using htseq-count. Transcripts were quantified with DESeq2. The genomes used for mapping were human (GRch38.9), mouse (GRCm38) or rhesus macaque (MMul10) based on sample species type. Control humanized/M and primatized/M were also mapped to the GRCm38 genome to compare native cell populations. Count tables were then used in iDEP.90 for normalization and analysis. Data were normalized through regularized log (rlog) calculations (Supplementary Fig. 4). Due to the high number of samples, biological replicates were pooled before sequencing, thereby increasing the total number of reads per condition. To ensure the accuracy of our conclusions, we performed a separate analysis using CORNAS, a tool specifically developed for analysing gene expression data without biological replicates. CORNAS uses the sequencing coverage information determined from the RNA concentration of each sample to estimate the posterior distribution of a true gene count[37].

mRNA fluorescence in situ hybridization analysis. Animals were injected in the lateral tail vein with 0.15 µg of aVHH mRNA LNPs. At 16 h post-delivery, the animals were sacrificed, perfused with 1× PBS and their livers removed and incubated in 4% paraformaldehyde overnight at 4 °C. The livers were incubated in 30% sucrose at 4 °C for 24 h, embedded in optimal cutting temperature (OCT) medium and processed into 10 µm frozen sections. The delivered aVHH mRNA and endogenous mRNA were visualized using an RNAscope Multiplex Fluorescent Reagent Kit v2 (Advanced Cell Diagnostics, 323136), according to the manufacturer's instructions. A custom probe set was designed against the synthetic aVHH mRNA sequence (ACD, 879561). Species-specific probes were used to distinguish engrafted human and rhesus cells within the mouse livers. Human and mouse beta-2 microglobulin mRNA were used to positively identify human cells (ACD, 484661-C3) and mouse cells (ACD, 415191-C2), respectively. Rhesus ubiquitin C mRNA was used to identify rhesus cells (ACD, 521081-C3). Images for cell-type staining were acquired using a Zeiss Plan-Apo ×20 0.8 NA air objective in an UltraVIEW spinning disk confocal microscope equipped with a Hamamatsu Flash 4.0 v2 CMOS camera. Images were captured and processed using Volocity software (PerkinElmer).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All RNA sequencing data have been deposited online at GEO (GSE178313). The scripts used to analyse barcodes are available at Github (https://github.com/Jack-Feldman/barcode_count). All other data are shown in the figures.

References

41. Dobrovolskaia, M. A., Shurin, M. & Shvedova, A. A. Current understanding of interactions between nanoparticles and the immune system. *Toxicol. Appl. Pharmacol.* **299**, 78–89 (2016).
42. Azuma, H. et al. Robust expansion of human hepatocytes in *Fah*$^{-/-}$/*Rag2*$^{-/-}$/*Il2rg*$^{-/-}$ mice. *Nat. Biotechnol.* **25**, 903–910 (2007).

Acknowledgements

The authors thank K. E. Tiegreen, S. Durham and R. Hughley at Georgia Institute of Technology. The work was funded by the National Institutes of Health (R01-GM132985, awarded to J.E.D., and UG3-TR002855, awarded to J.E.D. and P.J.S.) and DARPA (PREPARE, grant no. HR00111920008, awarded to P.J.S. and J.E.D.).

Author contributions

M.Z.C.H., C.N.D., P.J.S. and J.E.D. conceived the experiments. P.J.S. and J.E.D. obtained funding for, and oversaw, the research. All authors performed the experiments. M.Z.C.H., C.N.D. and J.E.D. wrote the initial manuscript, which was edited by the authors.

Competing interests

J.E.D. is a consultant for GV. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41565-021-01030-y.

Correspondence and requests for materials should be addressed to James E. Dahlman.

Peer review information *Nature Nanotechnology* thanks David Morrissey and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

nature research

Corresponding author(s): James Dahlman

Last updated by author(s): Oct 11, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	In vivo cell populations were analyzed and isolated using the BD FacsFusion in the Georgia Institute of Technology Cellular Analysis Core. Library QC for deep sequencing was conducted using the Agilent Bioanalyzer 2100 and sequencing was conducted on an Illumina MiniSeq. Library QC for bulk RNA seq was conducted using the Agilent Bioanalyzer 2100 and sequenced using a NextSeq 550 instrument using the 150-cycle paired end high-throughput cartridge
Data analysis	Data were analyzed and plotted in GraphPad Prism 8. The specifics of the data analyses are described in the text. Flow cytometry data was analyzed using FlowJo v10. Bulk RNA seq data were mapped using STAR aligner 2.7.8a and sequencing data were analyzed using iDEP.90 and CORNAS 1.0. Scripts used to analyze barcodes are available at Github (https://github.com/Jack-Feldman/barcode_count).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All RNA sequencing data have been deposited online at GEO (GSE178313). Genomes used for mapping were human (GRch38.9), mouse (GRCm38), or rhesus macaque (MMul10) based on sample species type. A control humanized/M and primatized/M were also mapped to the GRCm38 genome to compare native cell populations. All other data are shown in the figures.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For in vivo experiments: N = 2 for PBS negative control and 3 mice for experimental control. Those samples size were chosen to ensure accuracy of the data and accurate statistics. The 2 mice per PBS control were chosen to limit the number of mice.
Data exclusions	None
Replication	Three biological replicates were pooled before sequencing. Data were replicated successfully by performing the whole experiment a second time keeping the biological replicates separated.
Randomization	Mice were randomly selected, no algorithm was used.
Blinding	Some of those performing the experiments were blinded (e.g., during mice injection and cell isolation). We did not blind the authors performing the analysis due to limited access to infrastructure.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-CD31 (390, BioLegend, 102427), anti-CD45.2 (104, BioLegend, 109831), anti-CD47 (CC2C6, Biolegend, 323106), and MonoRab™ Anti-Camelid VHH (96A3F5, GenScript, A01860). Antibodies were used as purchased and diluted in 1:300 for staining in vivo and 1:1000 for staining in vitro.
Validation	anti-CD31 (BioLegend, 390): "Anti-mouse CD31 clones 390 and MEC13.3 bind to their respective non-overlapping epitopes in IgD2 of CD31. CD31 is a 130-140 kD glycoprotein, also known as platelet endothelial cell adhesion molecule (PECAM-1) and EndoCAM". anti-CD45.2 (BioLegend, 104): "CD45.2 is an alloantigen of CD45, expressed by Ly5.2 bearing mouse strains (e.g., A, AKR, BALB/c, CBA/Ca, CBA/J, C3H/He, C57BL, C57BR, C57L, C58, DBA/1, DBA/2, NZB, SWR, 129). CD45, a member of the protein tyrosine phosphatase (PTP) family, is a 180-240 kD glycoprotein expressed on all hematopoietic cells except mature erythrocytes and platelets." Anti-human CD47: "CD47 also known as Rh-associated protein, gp42, integrin-associated protein (IAP), and neurophilin, is a 42-52 kD member of the immunoglobulin superfamily containing a five-pass transmembrane attachment." Reactivity species were indicated as the following: Human, African Green, Baboon, Cynomolgus, Rhesus. Application references: 1) Seiffert M, et al. 1999. Blood 94:3633. 2) Leclair P, et al. 2018. Cell Death Dis. 5:544 (Block). MonoRab™ Anti-Camelid VHH (96A3F5, GenScript): "A Camelid VHH Antibody (also called single-domain antibody, sdAb or Nanobody by Ablynx) is a peptide chain of about 110 amino acids long, comprising one variable domain of a heavy-chain antibody, these are called VHH fragments. It is produced from a hybridoma resulting from the fusion of partner and B-lymphocytes obtained from a rabbit immunized with llama sdAb."

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	A549 cell lines were purchased at ATCC.
Authentication	The cell lines were not authenticated by us but were by ATCC (https://www.atcc.org/products/ccl-185).
Mycoplasma contamination	The cell lines were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	A549 are not included in the misidentified lines.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Liver-xenograft FRG mice were generated and provided by Yecuris (Portland, Oregon). FRG KO on C57Bl/6 mice were repopulated with cryopreserved hepatocytes from either C57Bl/6 mice (murinized), rhesus macaques (primatized), or from a human donor (HHM01008, a male 1-year-old donor). Balb/C (BALB/cJ), BL/6 (C57Bl/6J), and NZB (NZB/BlNJ) mice were acquired from Jackson Laboratories. All mice used were males between 7 and 8 month old. Light cycle of mice housing room are from 7 am to 7 pm. Housing room are kept at ~ 70 F with ~ 30% average humidity.
Wild animals	None.
Field-collected samples	None.
Ethics oversight	All animal experiments were performed in accordance with the Georgia Institute of Technology's Institutional Animal Care and Use Committee's animal care and services policy.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were isolated 24 hours after injection with LNPs. Mice were perfused with 20 mL of 1X PBS through the right atrium. Tissues were finely minced, and then placed in a digestive enzyme solution with Collagenase Type I (Sigma Aldrich), Collagenase XI (Sigma Aldrich) and Hyaluronidase (Sigma Aldrich) at 37 ºC at 550 rpm for 45 minutes. Cell suspension was filtered through 70µm mesh and red blood cells were lysed. Cells were stained to identify specific cell populations and sorted using the BD Facs Fusion cell sorter at the Georgia Institute of Technology Cellular Analysis Core.
Instrument	BD Facs Fusion cell sorter
Software	The data were analyzed using FlowJo (BD Biosciences)
Cell population abundance	Greater than 1000.
Gating strategy	Cells were gated on FSC-A / SSC-A. Singlets were gated FSC-A / FSC-H. Cells types were gated by CD31 / CD45.2. Human and monkey cell types were gated by CD47. See supporting figures 1 and 2 for additional information.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.