

Title

Transgene sequences free of CG dinucleotides lead to high level, long-term expression in the lung independent of plasmid backbone design

Authors

Reto P. Bazzani^{a,†}, Ian A. Pringle^{a,b,†}, Mary M. Connolly^{a,b}, Lee A. Davies^{a,b}, Stephanie G. Sumner-Jones^{a,b}, Martin Schleef^c, Stephen C. Hyde^{a,b}, & Deborah R. Gill^{a,b,*}.

Affiliations

- a. Gene Medicine Research Group, NDCLS, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, OX3 9DU, UK
- b. The UK Cystic Fibrosis Gene Therapy Consortium
- c. PlasmidFactory, Meisenstraße 96, D-33607 Bielefeld, Germany

† These authors contributed equally

*Correspondence: D.R. Gill; E-mail: deborah.gill@ndcls.ox.ac.uk

Tel: +44 1865 221845; Fax: +44 1865 221834

Keywords

Gene therapy; Plasmid DNA; Minicircles; CpG-free; Persistent expression; Cystic fibrosis

Acknowledgements

This work was funded by the UK Cystic Fibrosis Trust

Abstract

Non-viral aerosol gene therapy offers great potential for treating chronic lung diseases of the airways such as cystic fibrosis (CF). Early clinical trials showed that transgene expression in the airways was transient whereas maximal duration of transgene expression is essential in order to minimise the frequency of aerosol treatments. Improved vector design, such as careful selection of the promoter/enhancer, can lead to more persistent levels of transgene expression, but multiple factors affect expression *in vivo*. Following aerosol delivery to the lungs of mice, we measured reporter gene expression from a CpG-free luciferase transgene cassette in the context of both a plasmid and minicircle vector configuration and showed that the vector backbone had no effect on expression. Transgene activity was affected by the vector backbone however, when a similar, but sub-optimal CpG-containing transgene was used, suggesting that aspects of the plasmid backbone had a negative impact on transgene expression. Similar studies were performed in Toll-like receptor-9 (TLR9) knockout mice to investigate a potential role for the TLR9 signalling pathway in detecting CpGs in the vector sequence. Even in the absence of TLR9, persistent expression could only be achieved with a CpG-free transgene. Together, these data indicate that in order to achieve high levels of persistent expression *in vivo*, a CpG-free transgene cassette is required.

1. Introduction

Research into gene therapy continues for many lung diseases, including asthma, cancer, respiratory distress syndrome, α_1 -antitrypsin deficiency, emphysema and cystic fibrosis (CF). A successful gene therapy for chronic lung disease will most likely require long-term or even life-long therapy [1, 2], where persistent transgene expression is required in terminally differentiated airway cells. Studies in human airways have suggested that the half-life of airway epithelial cells is at most 3-5 months [3], making it unlikely that a single dose of vector would achieve life-long transgene expression for therapeutic benefit, thus repeated administration will probably be required.

Non-viral gene transfer agents (GTA) such as the cationic liposome GL67A and polyethylenimine (PEI) have been shown to be suitable for repeated administration to the airways without loss of efficacy [4, 5]. The studies described here use PEI aerosol complexes to deliver plasmid DNA (pDNA) to the lungs of mice. While PEI formulations delivered as a bolus of liquid to the lungs have been associated with acute toxicity [6, 7], when equivalent aerosol formulations were delivered to mice, robust transgene expression with no detectable toxicity was observed [5, 7]. Indeed pDNA/PEI aerosol formulations have been repeatedly administered to airways 10 times with no detectable toxicity [5]. Studies in mice (Supplementary Fig 1) and sheep [8] have shown that the level of transgene expression from pDNA/PEI aerosols is broadly similar to that achieved by GL67A. Furthermore transgene expression has been observed in both ciliated epithelial cells [9] and alveolar type I cells [10] both of which are important cell populations for lung disease.

Early non-viral gene therapy trials for CF showed that transgene expression was only transient, lasting no more than 1 week after a single dose in the nose and lung [11-16]. More recently, such research has focused on the development of pDNA to influence transgene expression levels and duration by careful selection of constituent components [17]. In particular, the selection of an appropriate enhancer/promoter sequence is thought to be a key factor, although few promoters have been assessed in a clinical setting [18]. Studies have shown that short-lived transgene expression in the lung [19], and the liver [20, 21], is not due to the loss of pDNA, but involves transcriptional silencing. Unmethylated CG dinucleotides (CpGs) acting as a ligand for the Toll-like receptor 9 (TLR9) are known for their immunostimulatory potential [22]. Furthermore, the accompanying increased levels of pro-inflammatory cytokines have been linked to attenuation of viral promoters [23-25]. The use of CpG-reduced and CpG-free plasmids in the lung [26-28], muscle [29] and in the liver [27, 28, 30] can be beneficial because they result in reduced levels of pro-inflammatory cytokines. Interestingly, several of these studies also showed a correlation between reduced inflammation and increased levels of sustained transgene expression [26, 28, 30].

We have previously evaluated multiple CpG-free enhancer/promoter combinations for their ability to sustain high-level gene expression in the lungs of mice and identified a hybrid CpG-free sequence based on the human cytomegalovirus enhancer and elongation factor 1-alpha promoter (hCEFI) [26, 31]. Transgene expression from hCEFI resulted in persistent expression for at least 5 months in the mouse lung [4] and at least 1 month in the sheep lung [32]. The hCEFI promoter was selected to express codon-optimised cystic fibrosis transmembrane conductance regulator (CFTR) cDNA in the clinical plasmid pGM169 [26]. Large quantities of the CpG-free pharmacopoeia-compliant pGM169 pDNA

were manufactured [33] and complexed with cationic liposome GL67A for a single application to the airways of CF patients in a safety and dose-ranging Phase 1/2a clinical trial [34]. This trial identified a safe (5 ml) aerosol dose of pGM169/GL67A to take forward for aerosol delivery of 12 monthly doses to the lungs of CF patients in a placebo-controlled Phase 2b clinical trial [35]. This latter trial led to statistically significant changes in lung function compared with placebo-treated subjects and confirmed the utility of CpG-free plasmids and promoters in the human lung. The removal of all CpGs from a plasmid is a significant undertaking however, especially as approximately half of the plasmid molecule (the plasmid backbone) is only required for propagation in bacteria and is therefore unnecessary for the final product. One strategy is to use minicircle DNA (mcDNA) with minimal backbone sequences [36]. An advantage of using minicircles is that, in addition to removing unnecessary bacterial sequences, and reducing the overall number of CpGs, these vectors have a reduced overall size, resulting in a higher effective dose of pDNA delivery [37]. In the mouse liver, minicircles have been shown to improve the duration of expression relative to conventional plasmids, following hydrodynamic injection of naked pDNA [21]. To our knowledge, however, minicircles have not yet been evaluated clinically, probably due to difficulties in producing sufficient quantities of pure minicircle vector suitable for clinical delivery [38]; further investigation of minicircle production processes may ultimately improve yields [37, 39].

In this study we systematically investigated the impact of plasmid design on long-term transgene expression in the mouse lung. We evaluated the impact of various pDNA backbones and minicircles on gene expression following aerosol delivery to the mouse lung and investigated the degree to which transgene CpG content influences duration of expression.

2. Materials & Methods

2.1. Mice

Female BALB/c mice or female TLR9^{-/-} mice on a BALB/c background (Oriental BioService Inc. Kyoto, Japan) aged 6–12 weeks were used throughout the current study. Mice were housed in accordance with UK Home Office ethical and welfare guidelines and fed on standard chow and water *ad libitum*.

2.2. Plasmid DNA

Plasmids used in this study are described in Fig 1. All plasmids contained the hCEFI promoter, a hybrid of CpG-free versions of the human CMV enhancer and the elongation factor 1a promoter [26]. Details of first (G1), second (G2) and fourth-generation (G4) plasmids have been described previously [26] and contain an identical CpG-free 3' untranslated region based on the bovine growth hormone 3'UTR [31]. Transgenes included CpG-rich and CpG-free versions of the Firefly Luciferase (Lux and soLux) [26] and human CFTR cDNA (CFTR and soCFTR2) [26]. Endotoxin-free pDNA was prepared using EndoFree Plasmid Mega Kit (Qiagen, Crawley, UK) or purchased from VGXi (The Woodlands, TX, USA); all plasmid preparations contained < 5 EU/mg.

2.2. Minicircles

Minicircle plasmids devoid of a conventional plasmid backbone were prepared by PlasmidFactory (Bielefeld, Germany) from precursor plasmids containing the hCEFI promoter and CpG-rich and CpG-free versions of the firefly Lux transgenes (see above) [26]. A proprietary production process was then used to produce 2 mg of minicircle DNA containing only the transgene cassette and a 209 bp sequence for chromatography,

affinity and recombination (SCAR) required for minicircle production and purification [37, 40].

2.3. Aerosol delivery of PEI/pDNA to the mouse lung

For aerosol delivery to the mouse lung, plasmid DNA (pDNA, 5 ml @ 0.4 mg/ml) was complexed with 5 ml of 25 kDa branched polymer polyethylenimine (PEI) (Sigma Aldrich) and 10 ml aerosolised to groups of mice as described previously [5]. Mice were sacrificed by cervical dislocation at the indicated time-points post-aerosol.

2.4. Luciferase reporter assay

Mouse lungs were placed in Lysing Matrix D Tubes (MP Biomedicals, Oakbank, UK) containing 300 µl 1XRLB (Promega) and homogenised using a FastPrep FP120 machine (MP Biomedicals) (45 sec at 4.0 m/sec). Lux activity in mouse lung lysates was quantified using the Promega Luciferase assay reagent system as described previously [41, 42]. The level of Lux activity, expressed in arbitrary relative light units (RLU), for each sample was normalised against total protein concentration determined using a detergent-compatible protein assay (Biorad, Larnae, UK). The background detection limit of the mouse lung Lux assay was determined by calculating Lux activity in six untreated mouse lung lysates. To aid in comparison with other studies, standards of recombinant Lux protein (Promega) were also routinely analysed. For comparative purposes 100 RLU per mg total lung protein corresponds to 2467 ng recombinant Lux (Promega) per mg total lung protein.

2.5. Absolute quantification of mRNA

Total RNA was extracted from mouse lung tissue using the RNeasy Mini Kit (Qiagen) as described previously [26]. Absolute levels of vector-derived mRNA and murine *Cfr* mRNA were quantified by TaqMan, as described previously [26, 43].

2.6. Statistics

Group sizes were selected to achieve >0.8 statistical power for the relevant statistical comparisons using G*Power [44]. Differences in transgene activity were determined using Bonferroni corrected Mann-Whitney U tests (MWU) after significant Kruskal-Wallis (KW) analyses where appropriate. Error bars on graphed data represent mean \pm standard error of the mean (SEM) for all data sets. * indicates $p < 0.05$, ** indicates $p < 0.01$.

3. Results

3.1. Comparison of luciferase activity from plasmids and minicircles containing a CpG-rich or CpG-free transgene

Aerosol delivery of pDNA complexed with 25 kDa branched PEI has been shown to result in sustained transgene expression in the murine lung when utilising the CpG-free hCEFI promoter [5]. We wished to further understand the contribution of the hCEFI promoter and the potential role of the CpG-free plasmid in sustained transgene expression. We constructed two Lux reporter expression cassettes, hCEFI-Lux and hCEFI-soLux, containing the hCEFI promoter directing expression of CpG-rich (Lux) and CpG-free (soLux) versions of the Lux transgene, respectively. The hCEFI-Lux and hCEFI-soLux expression cassettes were sub-cloned into the CpG-free G4 plasmid backbone [26] to generate plasmids G4-hCEFI-Lux and G4-hCEFI-soLux (Fig 1). In addition, the hCEFI-Lux and hCEFI-soLux expression cassettes were also manufactured as minicircles (MC) with a minimal 209 bp backbone sequence [40] generating MC-hCEFI-Lux and MC-hCEFI-soLux (Fig 1). All four Lux expression vectors were complexed with 25 kDa PEI at a concentration of 0.2 mg pDNA/ml and delivered to mice via aerosol (10 ml). Lux activity in homogenised lung tissue at 1, 7, 14 & 28 days post-aerosol was determined (Fig 2). The results showed that when the CpG-free hCEFI-soLux transgene was used there was no significant difference between Lux levels obtained from the plasmid and minicircle constructs; in both cases robust expression persisted for the duration of the study (28 days). However, when the hCEFI-Lux transgene, containing 98 CpGs, was used, the corresponding Lux activity was significantly lower at every time-point ($p < 0.01$, MWU). Neither the plasmid nor the minicircle construct exhibited persistent expression, since by day 28 Lux activity had fallen > 5 -fold relative to their respective day 1 levels ($p < 0.05$, MWU). There was significantly higher Lux activity from the minicircle at each time-point

($p < 0.05$, MWU), but by day 28 activity from MC-hCEFI-soLux was 21-fold higher than the MC-hCEFI-Lux, and activity from G4-hCEFI-soLux was 100-fold higher than G4-hCEFI-soLux. Together, these findings showed that the CpG-free hCEFI-soLux expression cassette was most effective at generating high-level and sustained Lux activity in the murine lung, irrespective of the plasmid or minicircle vector format.

3.2. Comparison of luciferase activity from plasmids with varying CpG backbone content containing a CpG-rich or CpG-free Lux transgene

The G4 plasmid backbone, utilised above, is CpG-free (0 CpGs in 1368 bp; Fig 1) and was previously selected as the pharmacopoeia-compliant plasmid backbone for the plasmid pGM169 used in clinical trials [35]. We wished to investigate the contribution of the plasmid backbone in determining the overall transgene expression profile in the murine lung. We therefore utilised the G1 (176 CpGs in 2914 bp) and G2 (51 CpGs in 1919 bp) plasmid backbones [26], to construct plasmids expressing the CpG-rich and CpG-free Lux transgene cassettes comprising pG1-hCEFI-Lux, pG1-hCEFI-soLux, pG2-hCEFI-Lux and pG2-hCEFI-soLux (Fig 1). All six plasmids were complexed with PEI and aerosolised to the lungs of mice. Lux activity was measured in homogenised lung tissue at 7, 14 and 28 days post-dose (Fig 3). Regardless of the plasmid backbone used (G1, G2 or G4), a clear difference was observed between the duration of Lux activity from plasmids carrying the hCEFI-Lux and hCEFI-soLux expression cassettes (Fig 3A). At every time-point Lux activity from the plasmids expressing the CpG-free soLux transgene was greater than their CpG-rich Lux counterpart ($p < 0.005$, MWU), with no observed loss of activity over the course of the study (Fig 3B).

Lux activity from plasmids containing the hCEFI-Lux cassette, however, was transient such that by day 28 activity had fallen to between 30 % and 2 % of their respective day 1 level ($p < 0.05$, MWU) (Fig 3B). However, relative to their (similar) day 1 levels, activity from the G4-hCEFI-Lux plasmid was significantly greater (3.5-fold to 18-fold) than that for plasmids with the G2 or G1 backbone ($p < 0.01$, MWU) at day 28. This suggests that some aspect of the backbone design, most probably the CpG content, can affect the overall Lux activity (Fig 3B), but only when the sub-optimal CpG-containing transgene is used.

3.3. The role of the CpG-specific receptor TLR9 in transgene expression

We wished to determine if the observed differences in duration of expression between the CpG-free and CpG-rich transgenes was dependent on the detection of CpGs *in vivo*. The presence of unmethylated CpGs in DNA can be detected in intracellular compartments via the TLR9 signalling pathway [45]. Therefore we investigated Lux activity following aerosol delivery to TLR9-deficient mice. PEI was complexed with G2 (G2-hCEFI-Lux & G2-hCEFI-soLux) and G4 (G4-hCEFI-Lux & G4-hCEFI-soLux) plasmids and delivered as an aerosol to the lungs of TLR9-deficient mice alongside BALB/c wild type mice as a control. Lux activity was measured at day 1, 7 & 28 days post-aerosol. Figure 4A shows that irrespective of the TLR9 status of the animals, Lux activity from the CpG-rich Lux transgene fell to around 10 % of day 1 activity within 7 days, and remained at those levels for the duration of the experiment (Fig 4A). Importantly, Lux activity from the CpG-free soLux transgene persisted at around 100 % of day 1 activity in both wild type and TLR9-deficient mice (Fig 4B). These data indicate that the effect of the transgene CpGs on the expression profile is independent of the TLR9 pathway. However, even in the absence of TLR9, Lux activity at 28 days post-aerosol was significantly higher (8.5-fold) when using the CpG-free G4 backbone compared with the CpG-containing G2 backbone ($p = 0.009$,

MWU) (Fig 4A). This again indicates that the plasmid backbone can affect overall Lux activity, but only when a CpG-rich transgene is used.

3.4. Duration of mRNA activity from CpG-rich and CpG-free CFTR expression plasmids

The experiments described above demonstrate that the duration of Lux activity from the hCEFI promoter in the mouse lung is dependent upon the use of a CpG-free Lux transgene cDNA. Although the Lux transgene is a widely-used reporter gene it has no clinical relevance. Therefore, we determined if our observations could be extended to the duration of human CFTR transgene expression. We investigated the CpG-free, clinical plasmid pGM169 (G4-hCEFI-soCFTR) [26] and constructed a new plasmid G4-hCEFI-CFTR, containing the standard CpG-rich human CFTR cDNA (4440 bp, 58 CpGs; [26]) (Fig 1). Both plasmids were complexed with PEI and aerosolised to groups of mice as described above. At 1, 7, 14, 28 & 56 days post-aerosol, the mouse lungs were harvested and total mRNA extracted for quantitative RT-PCR analysis of vector-specific CFTR mRNA. Endogenous murine CFTR mRNA was also quantified. Figure 5 shows that at day 1 post-aerosol there was no significant difference in the activity of the two plasmids. However, at all other time-points CFTR mRNA activity from the plasmid containing the CpG-free CFTR (G4-hCEFI-soCFTR) was significantly greater by 2-4 fold ($p = 0.006$, $p = 0.004$, $p = 0.004$, $p = 0.006$, MWU). Furthermore, activity from the plasmid containing the CpG-rich CFTR (pG4-hCEFI-CFTR) was significantly lower at days 7 to 56 (by 3-5 fold) than it was at day 1 ($p < 0.004$, MWU) indicating an overall lack of persistence (Fig 5).

4. Discussion

Control of transgene expression, particularly the maintenance of persistent expression in the target organ, is an important requisite for success in gene therapy applications. It is now well recognised that the selection of the promoter/enhancer sequence required to mediate transgene expression is crucial [42, 46]. However, several other factors can affect the outcome, especially when evaluated *in vivo*. Here, we investigate the impact of the plasmid backbone on transgene expression in the airways. The CpG-free transgene expression cassette hCEFI-soLux, shown to generate high level, persistent luciferase expression [26], was subcloned into four different plasmid backbones: a standard CpG-rich G1 backbone, a CpG-reduced G2 backbone, a CpG-free G4 backbone and finally a minicircle configuration (11 CpGs). In the pDNA/PEI aerosol delivery model, the choice of plasmid backbone (G1, G2 or G4) had no effect on the duration of Lux activity from hCEFI-soLux (Fig 3A) as all of these constructs displayed stable Lux activity for 28 days. Similarly, placing the hCEFI-soLux expression cassette in the context of a minicircle had no effect on activity; such minicircles performed no better than the conventional CpG-free plasmids *in vivo* (Fig 2). By contrast, when a different transgene cassette, containing a CpG-rich Lux gene (hCEFI-Lux) was used in these plasmid backbones, none of the plasmids or the minicircle exhibited persistent stable expression (Fig 2 & Fig 3A). There were, however, significant differences in Lux activity between the three plasmid backbones; G4-hCEFI-Lux, with the lowest CpG content had the highest day 28 activity relative to its day 1 level (Fig 3B).

The acute inflammatory response following delivery of pDNA complexes to the lung has been widely studied [47, 48] and it has been shown that the removal of CpGs from pDNA minimises such responses [26, 28, 49]. It was therefore hypothesised that detection of

CpGs by TLR9 could have some role in both the lack of persistence from hCEFI-Lux plasmids and the difference between the levels of Lux activity observed at day 28. Therefore, duration of Lux activity from a number of plasmid vectors was assessed in TLR9^{-/-} (knockout) mice. The results matched those observed in wild type BALB/c mice. Both G2 and G4 plasmid backbone plasmids resulted in persistent stable Lux activity when hCEFI-soLux was used (Fig 4A) but not when hCEFI-Lux was used (Fig 4B). The absence of CpG detection via TLR9, therefore, did not influence the duration of Lux activity. Once again however, the duration of Lux activity from hCEFI-Lux constructs was greater when the CpG-free backbone was used (Fig 4A). Therefore, while it seems that plasmid backbone can have an effect on the duration of Lux activity this observation was only seen when a CpG-containing Lux transgene was used and in all of these cases the persistence of transgene expression was poor.

It was then determined if the same effect was evident when a clinically relevant transgene was used. Comparison of the duration of expression from hCEFI-soCFTR (CpG-free) and hCEFI-CFTR (CpG-containing) transgene cassettes showed significantly higher activity from the CpG-free transgene as measured by Taqman rt-PCR (Fig 5).

One possible explanation is that in the CpG-containing transgenes the CpGs are being subjected to *de novo* methylation leading to the formation of a more compact chromatin structure that prevents or reduces transgene activity [50]. A previous study looked for evidence of *de novo* methylation of plasmids in the lung but failed to find any evidence of this [19]. However, studying the composition of pDNA in the lung is difficult because although detection of delivered plasmid is straightforward, confirming how much of the recovered pDNA is likely to have been transcriptionally active is not. Studies in the liver

have been more convincing in showing evidence of *de novo* methylation of pDNA *in vivo* [51] and similarly, studies in mouse muscle following adenoviral vector delivery have found evidence of *de novo* DNA methylation within 24 hours of delivery [52]. Therefore, *de novo* methylation of non-integrating DNA vectors does seem to be a process that can occur *in vivo*. Importantly, while the levels of activity from CpG-containing Lux and CFTR transgenes decline over time, they are still active, albeit at a lower level. This would suggest that if *de novo* methylation is responsible for the loss of activity, the transgenes are perhaps only partially methylated [48, 53]. An alternative explanation for the improved persistence when CpG-free transgenes are used is that these sequences are somehow prone to more favourable random integration into the genome. Due to the technical difficulties described above, this is difficult to disprove but it seems unlikely for a number of reasons. Firstly, improved persistence is not a general feature of CpG-free plasmids, as most plasmids designs tested *in vivo* do not result in persistent expression even when these transgenes are used [18, 26, 31]. Secondly, the effect is observed when two separate unrelated transgenes are used and persistence from many other CpG-free transgenes has also been observed *in vivo* (data not shown). Thirdly, when these plasmids are transfected into cell culture models we do not observe persistent expression when the cells divide as one would expect from an integrating vector.

As stated above when the CpG-free hCEFI-soLux construct was used, no difference was observed between the expression profiles of a minicircle and a plasmid (Fig 2). This result is in contrast to other studies that have demonstrated improved vector performance with minicircles [21, 37]. Nevertheless, it is a preferable outcome as despite the theoretical benefits of minicircles and their higher effective dose, the costs of production makes them

unsuitable for widespread use, particularly in studies involving large amounts of vector such as these.

It is also worth noting that there are some size differences between the plasmids and minicircles used in these studies (Fig 1). When conducting PEI aerosols a constant total mass of DNA was used (5 ml @ 0.4 mg/ml) rather than delivering equimolar amounts of material. Although this ensures consistency between aerosols for each plasmid, it results in slight differences in active doses per plasmid. For example, for every copy of G4 plasmid delivered in Figure 2, 1.45 copies of minicircle were delivered. If the activity of the expression construct was the same between the plasmid and minicircle configurations then one might expect to have seen 1.45X greater activity from the minicircles; however for the soLux constructs this was not the case, despite the experiments being powered to detect such an increase. Furthermore, we have observed previously that small (<2-fold) differences in activity can be caused by other confounding variables such as subtle differences between DNA manufacturers, or even batch to batch variations [17]. Therefore we do not believe that the relatively small size differences between the plasmids used in these studies have any significant impact on the data.

Based on observations in previous human clinical trials [11, 54] and preclinical mouse studies, the use of a CpG-free plasmid had advantages for repeated delivery of non-viral vectors to the human lung [35]. The use of CpG-free transgenes and backbone sequences in plasmid development facilitated the identification of the hCEFI promoter, shown to be superior to other promoters in terms of levels and duration of transgene activity in the mouse lung [4, 26]. The immune stimulatory effects of CpGs are also well known [23-25] and the results here are suggestive of CpGs suppressing transgene activity, possibly via

de novo methylation. Therefore, the presence of CpGs in a plasmid vector has the potential to interfere with and obscure transgene expression profiles *in vivo*. The economic and practical costs involved in producing CpG-free or low-CpG plasmids are minimal and we believe the data presented here demonstrate the benefits of such an approach.

References

- [1] Griesenbach U, Alton EW. Expert opinion in biological therapy: update on developments in lung gene transfer. *Expert opinion on biological therapy*. 2013;13:345-60.
- [2] Gill DR, Hyde SC. Delivery of genes into the CF airway. *Thorax*. 2014;69:962-4.
- [3] Teixeira VH, Nadarajan P, Graham TA, Pipinikas CP, Brown JM, Falzon M, et al. Stochastic homeostasis in human airway epithelium is achieved by neutral competition of basal cell progenitors. *eLife*. 2013;2:e00966.
- [4] Alton EW, Boyd AC, Cheng SH, Davies JC, Davies LA, Dayan A, et al. Toxicology study assessing efficacy and safety of repeated administration of lipid/DNA complexes to mouse lung. *Gene Ther*. 2014;21:89-95.
- [5] Davies LA, Hyde SC, Nunez-Alonso G, Bazzani RP, Harding-Smith R, Pringle IA, et al. The use of CpG-free plasmids to mediate persistent gene expression following repeated aerosol delivery of pDNA/PEI complexes. *Biomaterials*. 2012;33:5618-27.
- [6] Davies LA, McLachlan G, Sumner-Jones SG, Ferguson D, Baker A, Tennant P, et al. Enhanced lung gene expression after aerosol delivery of concentrated pDNA/PEI complexes. *Mol Ther*. 2008;16:1283-90.
- [7] Gautam A, Densmore CL, Golunski E, Xu B, Waldrep JC. Transgene expression in mouse airway epithelium by aerosol gene therapy with PEI-DNA complexes. *Mol Ther*. 2001;3:551-6.
- [8] McLachlan G, Davidson H, Holder E, Davies LA, Pringle IA, Sumner-Jones SG, et al. Pre-clinical evaluation of three non-viral gene transfer agents for cystic fibrosis after aerosol delivery to the ovine lung. *Gene Ther*. 2011;18:996-1005.
- [9] Rudolph C, Schillinger U, Ortiz A, Plank C, Golas MM, Sander B, et al. Aerosolized nanogram quantities of plasmid DNA mediate highly efficient gene delivery to mouse airway epithelium. *Mol Ther*. 2005;12:493-501.
- [10] Davies LA, Seguela C, Varathalingam A, Cheng SH, Hyde SC, Gill DR. Identification of transfected cell types following non-viral gene transfer to the murine lung. *J Gene Med*. 2007;9:184-96.
- [11] Alton EFWF, Stern M, Farley R, Jaffe A, Chadwick SL, Phillips J, et al. Cationic lipid-mediated CFTR gene transfer to the lungs and nose of patients with cystic fibrosis: a double-blind placebo-controlled trial. *The Lancet*. 1999;353:947-54.
- [12] Caplen NJ, Alton EFWF, Middleton PG, Dorin JR, Stevenson BJ, Gao X, et al. Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Nat Med*. 1995;1:39-46.
- [13] Gill DR, Southern KW, Mofford KA, Seddon T, Huang L, Sorgi F, et al. A placebo-controlled study of liposome-mediated gene transfer to the nasal epithelium of patients with cystic fibrosis. *Gene Ther*. 1997;4:199-209.
- [14] Hyde SC, Southern KW, Gileadi U, Fitzjohn EM, Mofford KA, Waddell BE, et al. Repeat administration of DNA/liposomes to the nasal epithelium of patients with cystic fibrosis. *Gene Ther*. 2000;7:1156-65.
- [15] Porteous DJ, Dorin JR, McLachlan G, Davidson-Smith H, Davidson H, Stevenson BJ, et al. Evidence for safety and efficacy of DOTAP cationic liposome mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Gene Ther*. 1997;4:210-8.
- [16] Zabner J, Cheng SH, Meeker D, Launspach J, Balfour R, Perricone MA, et al. Comparison of DNA-lipid complexes and DNA alone for gene transfer to cystic fibrosis airway epithelia in vivo. *J Clin Invest*. 1997;100:1529-37.

- [17] Gill DR, Pringle IA, Hyde SC. Progress and prospects: the design and production of plasmid vectors. *Gene Ther.* 2009;16:165-71.
- [18] Pringle IA, Gill DR, Connolly MM, Lawton AE, Hewitt AM, Nunez-Alonso G, et al. Rapid identification of novel functional promoters for gene therapy. *Journal of molecular medicine.* 2012;90:1487-96.
- [19] Pringle IA, Raman S, Sharp WW, Cheng SH, Hyde SC, Gill DR. Detection of plasmid DNA vectors following gene transfer to the murine airways. *Gene Ther.* 2005;12:1206-14.
- [20] Chen ZY, Yant SR, He CY, Meuse L, Shen S, Kay MA. Linear DNAs concatemerize in vivo and result in sustained transgene expression in mouse liver. *Mol Ther.* 2001;3:403-10.
- [21] Gracey Maniar LE, Maniar JM, Chen ZY, Lu J, Fire AZ, Kay MA. Minicircle DNA vectors achieve sustained expression reflected by active chromatin and transcriptional level. *Mol Ther.* 2013;21:131-8.
- [22] Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, et al. A Toll-like receptor recognizes bacterial DNA. *Nature.* 2000;408:740-5.
- [23] Li S, Wu SP, Whitmore M, Loeffert EJ, Wang L, Watkins SC, et al. Effect of immune response on gene transfer to the lung via systemic administration of cationic lipidic vectors. *Am J Physiol.* 1999;276:L796-804.
- [24] Qin L, Ding Y, Pahud DR, Chang E, Imperiale MJ, Bromberg JS. Promoter attenuation in gene therapy: interferon-gamma and tumor necrosis factor-alpha inhibit transgene expression. *Hum Gene Ther.* 1997;8:2019-29.
- [25] Tan Y, Li S, Pitt BR, Huang L. The inhibitory role of CpG immunostimulatory motifs in cationic lipid vector-mediated transgene expression in vivo. *Hum Gene Ther.* 1999;10:2153-61.
- [26] Hyde SC, Pringle IA, Abdullah S, Lawton AE, Davies LA, Varathalingam A, et al. CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nat Biotechnol.* 2008;26:549-51.
- [27] Yew NS, Przybylska M, Ziegler RJ, Liu D, Cheng SH. High and sustained transgene expression in vivo from plasmid vectors containing a hybrid ubiquitin promoter. *Mol Ther.* 2001;4:75-82.
- [28] Yew NS, Zhao H, Przybylska M, Wu IH, Tousignant JD, Scheule RK, et al. CpG-depleted plasmid DNA vectors with enhanced safety and long-term gene expression in vivo. *Mol Ther.* 2002;5:731-8.
- [29] Mann CJ, Anguela XM, Montane J, Obach M, Roca C, Ruzo A, et al. Molecular signature of the immune and tissue response to non-coding plasmid DNA in skeletal muscle after electrotransfer. *Gene Ther.* 2012;19:1177-86.
- [30] Hodges BL, Taylor KM, Joseph MF, Bourgeois SA, Scheule RK. Long-term transgene expression from plasmid DNA gene therapy vectors is negatively affected by CpG dinucleotides. *Mol Ther.* 2004;10:269-78.
- [31] Pringle IA, Hyde SC, Connolly MM, Lawton AE, Xu B, Nunez-Alonso G, et al. CpG-free plasmid expression cassettes for cystic fibrosis gene therapy. *Biomaterials.* 2012;33:6833-42.
- [32] McLachlan G, Davies LA, Gordon CM, Vrettou C, Baker E, Tennant P, et al. Repeat Aerosol Delivery of Concentrated PEI/pDNA to the Sheep Lung. *Molecular Therapy.* 2010;18:S339.
- [33] Cai Y, Rodriguez S, Nelson J, Batten T, Hebel H, Hyde SC, et al. Large-Scale cGMP Manufacture of a Plasmid Vector for Cystic Fibrosis Gene Therapy Clinical Trials. *Molecular Therapy.* 2012;20:S293.
- [34] Alton EW, Boyd AC, Porteous DJ, Davies G, Davies JC, Griesenbach U, et al. A Phase I/IIa Safety and Efficacy Study of Nebulized Liposome-mediated Gene Therapy for

Cystic Fibrosis Supports a Multidose Trial. American journal of respiratory and critical care medicine. 2015;192:1389-92.

[35] Alton EW, Armstrong DK, Ashby D, Bayfield KJ, Bilton D, Bloomfield EV, et al. Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial. The Lancet Respiratory medicine. 2015;3:684-91.

[36] Darquet AM, Cameron B, Wils P, Scherman D, Crouzet J. A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. Gene Ther. 1997;4:1341-9.

[37] Kobelt D, Schleef M, Schmeer M, Aumann J, Schlag PM, Walther W. Performance of high quality minicircle DNA for in vitro and in vivo gene transfer. Molecular biotechnology. 2013;53:80-9.

[38] Mayrhofer P, Schleef M, Jechlinger W. Use of minicircle plasmids for gene therapy. Methods in molecular biology. 2009;542:87-104.

[39] Gaspar VM, Maia CJ, Queiroz JA, Pichon C, Correia IJ, Sousa F. Improved minicircle DNA biosynthesis for gene therapy applications. Human gene therapy methods. 2014;25:93-105.

[40] Mayrhofer P, Blaesen M, Schleef M, Jechlinger W. Minicircle-DNA production by site specific recombination and protein-DNA interaction chromatography. J Gene Med. 2008;10:1253-69.

[41] Pringle IA, McLachlan G, Collie DD, Sumner-Jones SG, Lawton AE, Tennant P, et al. Electroporation enhances reporter gene expression following delivery of naked plasmid DNA to the lung. J Gene Med. 2007;9:369-80.

[42] Gill DR, Smyth SE, Goddard CA, Pringle IA, Higgins CF, Colledge WH, et al. Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1alpha promoter. Gene Ther. 2001;8:1539-46.

[43] Rose AC, Goddard CA, Colledge WH, Cheng SH, Gill DR, Hyde SC. Optimisation of real-time quantitative RT-PCR for the evaluation of non-viral mediated gene transfer to the airways. Gene Ther. 2002;9:1312-20.

[44] Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods. 2007;39:175-91.

[45] Latz E, Schoenemeyer A, Visintin A, Fitzgerald KA, Monks BG, Knetter CF, et al. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. Nature immunology. 2004;5:190-8.

[46] Pringle IA, Hyde SC, Gill DR. Non-viral vectors in cystic fibrosis gene therapy: recent developments and future prospects. Expert opinion on biological therapy. 2009;9:991-1003.

[47] Freemark BD, Blezinger HP, Florack VJ, Nordstrom JL, Long SD, Deshpande DS, et al. Cationic lipids enhance cytokine and cell influx levels in the lung following administration of plasmid: cationic lipid complexes. Journal of immunology. 1998;160:4580-6.

[48] McLachlan G, Stevenson BJ, Davidson DJ, Porteous DJ. Bacterial DNA is implicated in the inflammatory response to delivery of DNA/DOTAP to mouse lungs. Gene Ther. 2000;7:384-92.

[49] Bazzani RP, Cai Y, Hebel HL, Hyde SC, Gill DR. The significance of plasmid DNA preparations contaminated with bacterial genomic DNA on inflammatory responses following delivery of lipoplexes to the murine lung. Biomaterials. 2011;32:9854-65.

[50] Razin A. CpG methylation, chromatin structure and gene silencing-a three-way connection. The EMBO journal. 1998;17:4905-8.

- [51] Argyros O, Wong SP, Niceta M, Waddington SN, Howe SJ, Coutelle C, et al. Persistent episomal transgene expression in liver following delivery of a scaffold/matrix attachment region containing non-viral vector. *Gene Ther.* 2008;15:1593-605.
- [52] Brooks AR, Harkins RN, Wang P, Qian HS, Liu P, Rubanyi GM. Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *J Gene Med.* 2004;6:395-404.
- [53] Hsieh CL. Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol.* 1994;14:5487-94.
- [54] Ruiz FE, Clancy JP, Perricone MA, Bebok Z, Hong JS, Cheng SH, et al. A clinical inflammatory syndrome attributable to aerosolized lipid-DNA administration in cystic fibrosis. *Hum Gene Ther.* 2001;12:751-61.

Transgene sequences free of CG dinucleotides lead to high level, long-term expression in the lung independent of plasmid backbone design

Figure 1. Summary of plasmids used in these studies. Number of CpGs, CpG frequency and overall plasmid size in bp displayed in numbers. Black vertical bars indicate CpGs.

Figure 2. Comparison of Lux activity from plasmids and minicircles containing a CpG-rich or CpG-free transgene. Plasmids G4-hCEFI-soLux, G4-hCEFI-Lux, and minicircles MC-hCEFI-soLux & MC-hCEFI-Lux were complexed with PEI and aerosolised to groups of mice (N=6 per time-point). Lux activity (RLU per mg total lung protein) was determined at the time-points indicated. Error = 1 SEM. * indicates $p < 0.05$, ** indicates $p < 0.01$ by Mann-Whitney U-test.

Figure 3. Comparison of Lux activity from plasmids with varying CpG backbone contents containing a CpG-rich or CpG-free transgene. Plasmids G1-hCEFI-Lux, G2-hCEFI-Lux, G4-hCEFI-Lux, G1-hCEFI-soLux, G2-hCEFI-soLux & G4-hCEFI-soLux were complexed with PEI and aerosolised to groups of mice (N=6 per time-point). Lux activity (RLU per mg total lung protein) was determined at the time-points indicated (A) and displayed relative to their day 01 values (100%) at day 28 (B). Error = 1 SEM. * indicates $p < 0.05$, ** indicates $p < 0.01$ by Mann-Whitney U-test.

Figure 4. The role of the CpG-specific receptor TLR9 in transgene expression. Plasmids G2-hCEFI-Lux, G4-hCEFI-Lux (A), G2-hCEFI-soLux & G4-hCEFI-soLux (B) were complexed with PEI and aerosolised to groups of BALB/c (wt) or TLR9^{-/-} mice (TLR9) (N=6 per time-point). Error = 1 SEM. NS indicates no significant difference: * indicates $p < 0.05$, ** indicates $p < 0.01$ by Mann-Whitney U-test.

Figure 5. Duration of mRNA activity from CpG-rich and CpG-free CFTR expression plasmids. Plasmids G4-hCEFI-soCFTR and G4-hCEFI-CFTR were complexed with PEI and aerosolised to groups of mice (N=6 per time-point). Total mouse lung RNA was purified at the time-points indicated and quantitative RT-PCR used to determine the absolute copy number of vector-derived mRNA normalised to the absolute copy number of mouse CFTR mRNA (% vector mRNA copies / murine CFTR mRNA copies). Error = 1 SEM. NS indicates no significant difference: * indicates $p < 0.05$, ** indicates $p < 0.01$ by Mann-Whitney U-test.

Supplementary Figure 1. Comparison of Lux activity following aerosol administration of G4-hCEFI-soLux complexed with PEI or GL67A. G4-hCEFI-soLux, (5 ml @ 0.4 mg/ml) was complexed with 5 ml of 25 kDa branched polymer polyethylenimine (PEI) (Sigma Aldrich) and 10 ml aerosolised to groups of mice as described previously [5]. G4-hCEFI-soLux/GL67A complexes were formed by mixing 5 ml pDNA (5 ml @ 5 mg/ml) and 5 ml GL67A solution (Genzyme, Framingham, MA, USA) rehydrated in water at 12 mM as described previously [55] Lux activity (RLU per mg total lung protein) was determined at the time-points indicated (N=6 per time-point).

FIGURE 1

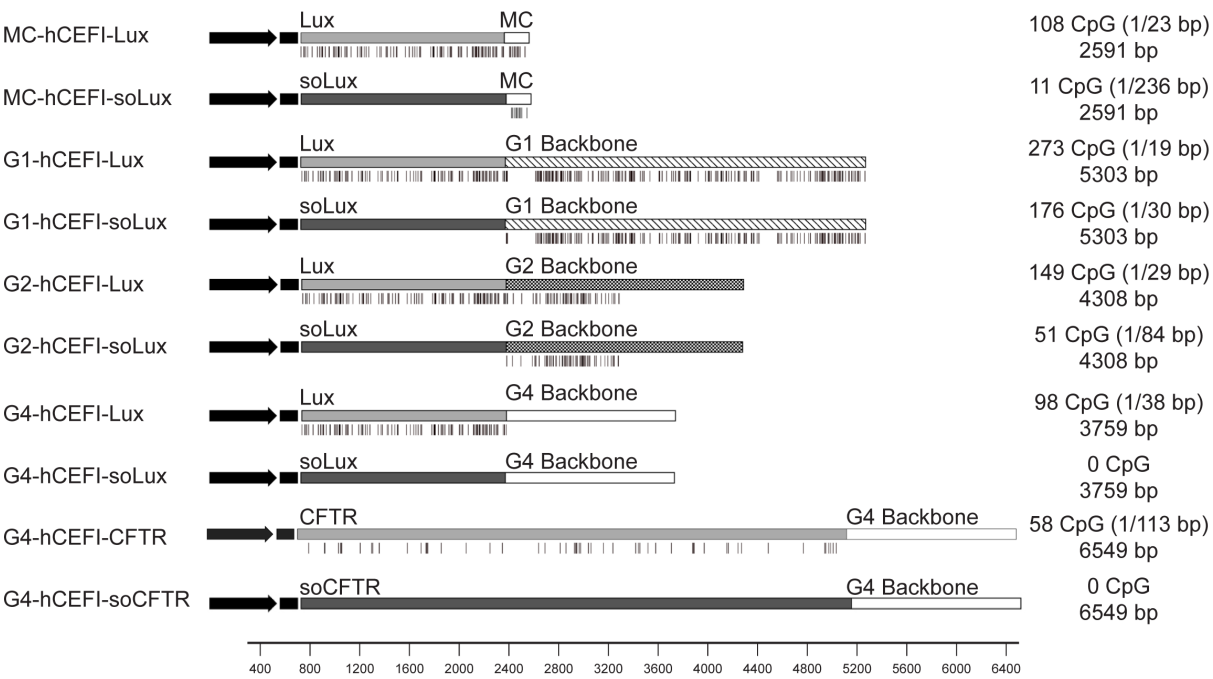


FIGURE 2

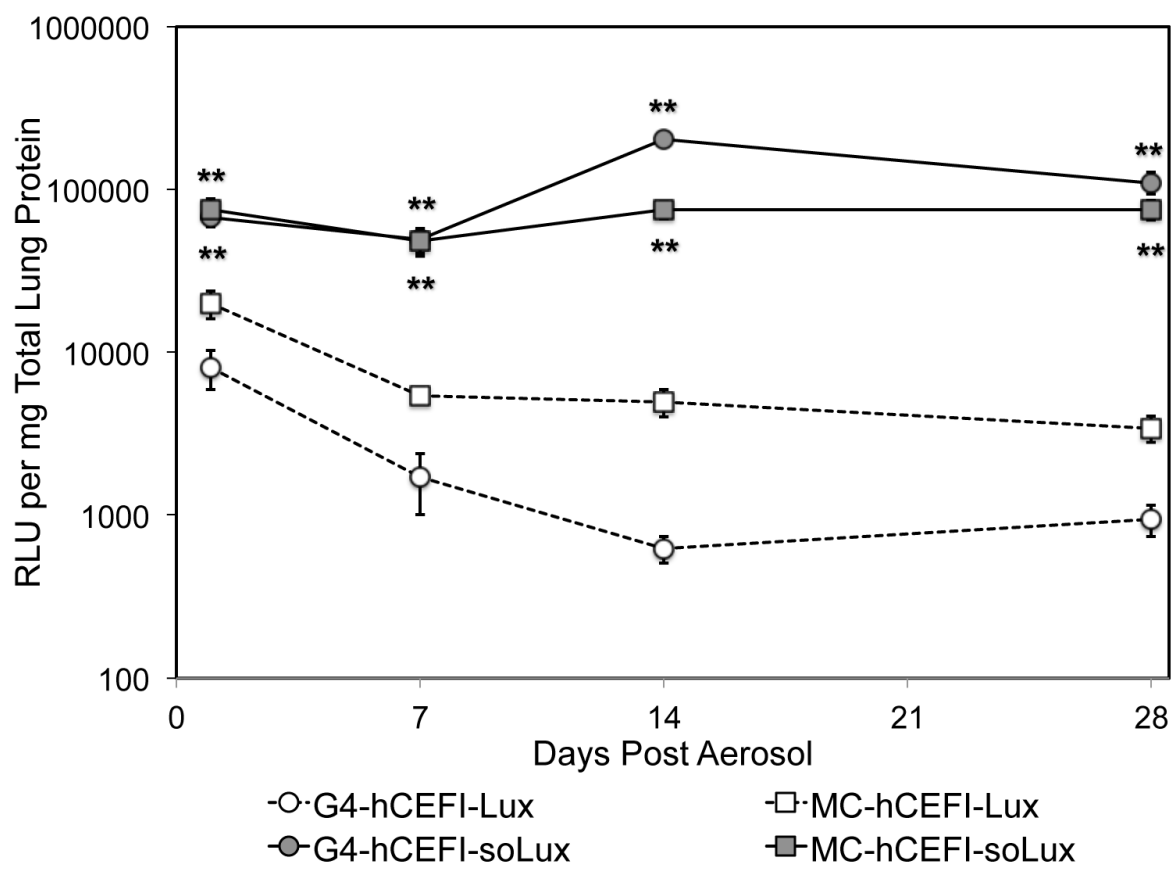


FIGURE 3

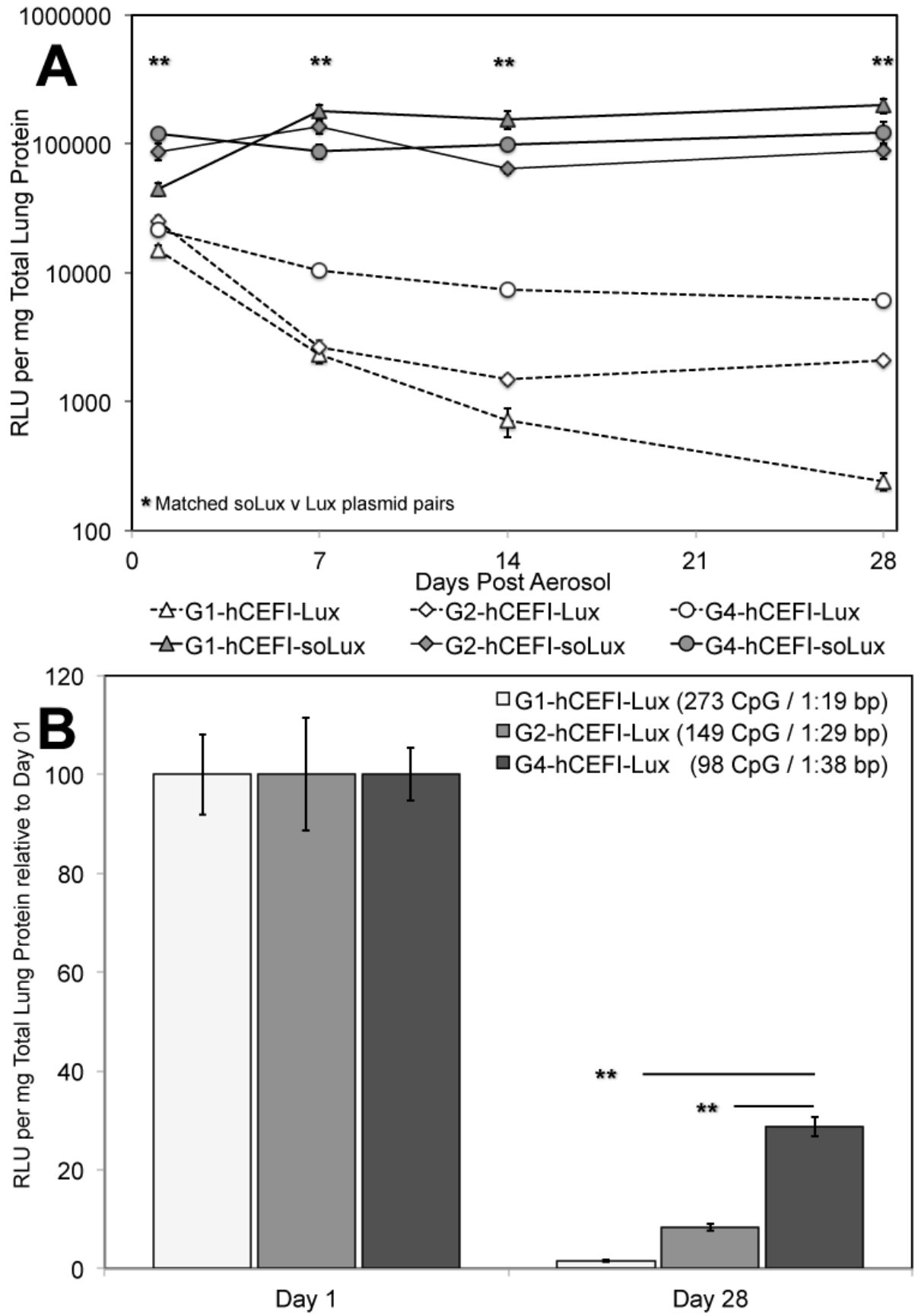


FIGURE 4

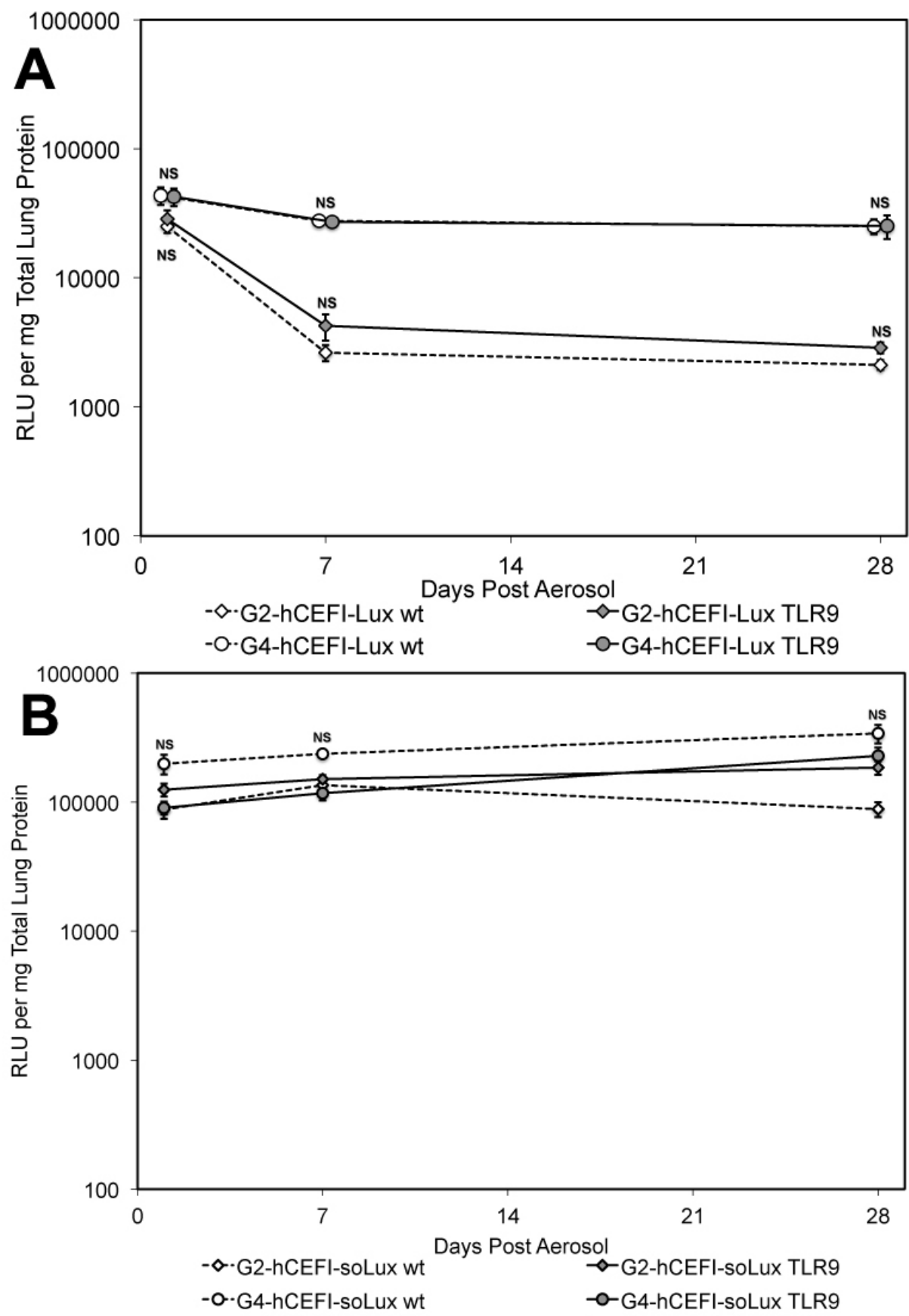
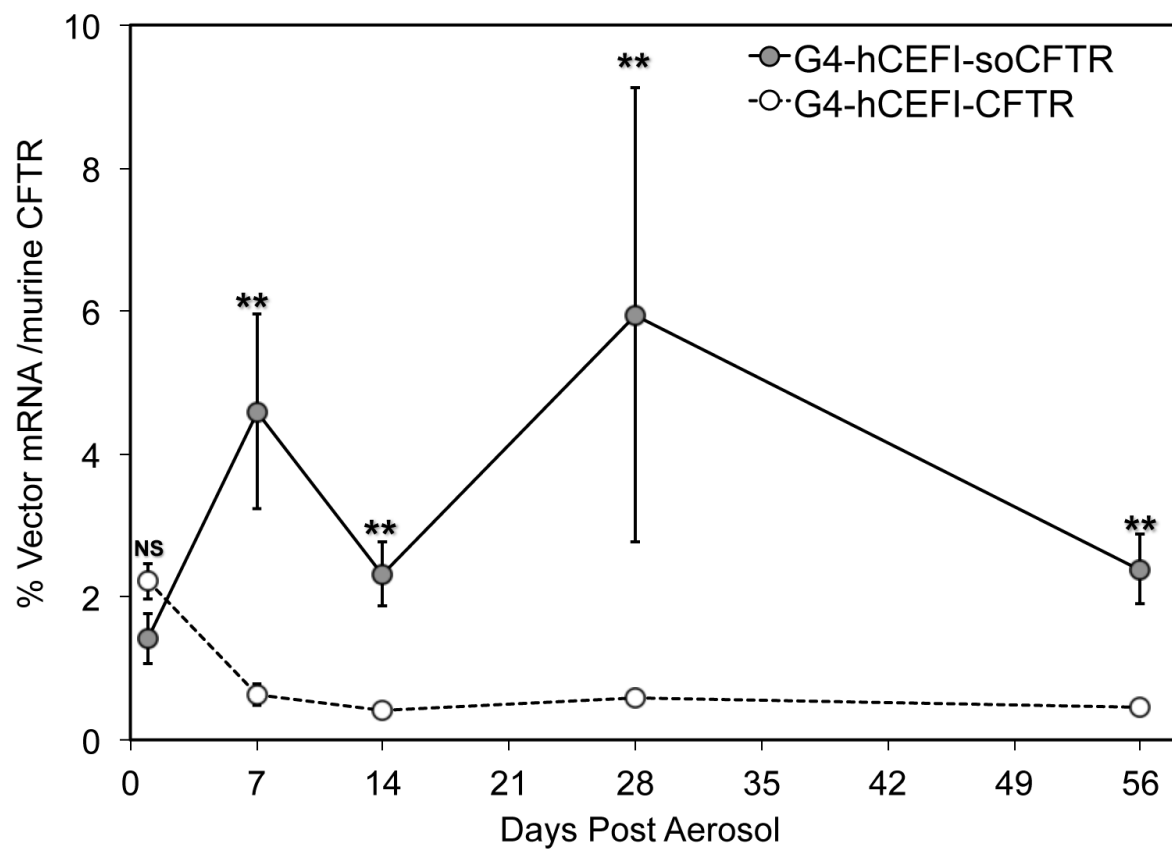


FIGURE 5



SUPPLEMENTARY FIGURE 1

