SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
IGA NEPHROPATHY
GENETICS RESEARCH

Received
13 August 2014

Accepted
19 November 2014

Published
9 December 2014

Correspondence and requests for materials should be addressed to A.J.M. (a.j.mcknight@qub.ac.uk)

Next-generation sequencing of the mitochondrial genome and association with IgA nephropathy in a renal transplant population

Adam P. Douglas[1], Dwaine R. Vance[1], Elaine M. Kenny[2], Derek W. Morris[2], Alexander P. Maxwell[1,3] & Amy Jayne McKnight[1]

[1]Centre for Public Health, Queen's University Belfast, United Kingdom, [2]TrinSeq, Trinity Genome Sequencing Lab, Department of Psychiatry, Institute of Molecular Medicine, Trinity College, Dublin, Ireland, [3]Regional Nephrology Unit, Belfast City Hospital, Belfast, United Kingdom.

Kidneys are highly aerobic organs that are critically dependent on the normal functioning of mitochondria. Genetic variations disrupting mitochondrial function are associated with multifactorial disorders including kidney disease. This study sequenced the entire mitochondrial genome in a renal transplant cohort of 64 individuals, using next-generation sequencing, to evaluate the association of genetic variants with IgA nephropathy and end-stage renal disease (ESRD, n=100).

Mitochondrial DNA was amplified in two fragments using long-range PCR and sequenced on an Illumina Genome Analyzer II platform using TruSeq, with variants confirmed using bidirectional Sanger sequencing. A total of 427 differences from the mitochondrial DNA reference sequence were identified, of which 113 had a minor allele frequency >5% in this population. Five common SNPs revealed evidence of significant association ($P \leq 10^{-4}$) with ESRD where patients had a primary diagnosis of IgA nephropathy. Sixty-four percent of common SNPs were present in individuals with IgA nephropathy.

There is an excess burden of common mtDNA SNPs in renal transplant recipients with IgA nephropathy compared with individuals who have no evidence of renal disease. Genetic risk profiles may assist clinical stratification of renal transplant patients, and we have validated a cost-effective approach to genotype larger, carefully phenotyped cohorts.

Glomerulonephritis (GN) is the most common primary renal diagnosis in patients receiving renal replacement therapy (RRT) in the UK, accounting for 18.8% of the prevalent end-stage renal disease (ESRD) cohort in 2012[1]. Kidneys are highly aerobic organs, and are critically dependent on mitochondrial activity. Mitochondria possess DNA that is discrete from the nuclear genome in the form of a circular, double-stranded molecule (mtDNA). The 16,569 base pair mitochondrial genome encodes 13 respiratory chain proteins, two rRNAs, and 22 tRNAs[2]. Genetic variations disrupting mitochondrial function are associated with increased molecular stress and multifactorial disorders including kidney disease[3–11].

Genetic variation in mtDNA may be inherited or acquired (termed somatic mutations) throughout life. MtDNA is particularly susceptible to acquiring somatic mutations as they both generate and respond to potentially mutagenic reactive oxygen species (ROS) and mtDNA lacks the protective and efficient repair systems associated with nuclear DNA. Mutations in mtDNA range from single base changes to large scale deletions. The 4977 bp "common deletion" has been associated with a number of disorders, including the renal diseases of focal segmental glomerulosclerosis (FSGS) and immunoglobin-A nephropathy (IgA nephropathy)[12], the most common form of glomerulonephritis in the UK. Another frequently studied mutation is the 3243A>G single base change which is associated with MELAS (mitochondrial encephalomyopathy with lactic acidosis and stroke like episodes) and FSGS[13,14]. MELAS can have renal involvement in the form of FSGS, but is uncommon in White European populations[15].

This project sought to comprehensively investigate mitochondrial genetic variants for association with IgA nephropathy and ESRD, using a next-generation sequencing approach for the entire mitochondrial genome.

Results

There was an average of 1 million reads per DNA sample (range 36,492–3,327,485 reads) with 98% of 71 base reads aligned to the mitochondrial genome. There was >99.9% coverage of each mtDNA base with mapping quality score >40 and average coverage depth of 11,092x (Supplementary Table 1).

A total of 427 differences from the reference sequence were identified (Figure 1, Supplementary Figure 2, Supplementary Table 2), of which 55 had a minor allele frequency (MAF) of zero, leaving 372 SNPs in this population; 113 SNPs had a MAF greater than 5%. Five common SNPs revealed evidence of significant association with ESRD where patients had a primary diagnosis of IgA nephropathy (Table 1; 6419A>C, $P=3.67 \times 10^{-6}$; 199T>C, $P=3.00 \times 10^{-4}$; 150C>T, $P=4.00 \times 10^{-4}$; 8251G>A, $P=4.00 \times 10^{-4}$; 16565C>G, $P=9.00 \times 10^{-4}$). The association of 6419A>C was maintained following permutation testing (n=100,000 permutations) based on 427 variants with resultant P=0.01. Ten SNPs with MAF less than 5% and 65 common SNPs were present only in the case group of individuals with IgA nephropathy. Strong linkage disequilibrium was not observed (Figure 2).

Discussion

This study of 32 cases and 32 controls provides ~60% power to identify a risk variant with an odds ratio of 2.0 at a 30% minor allele frequency. The relatively small number of persons assessed in this study means it is not sufficiently powered to identify significant association with rare mtDNA variants, nor risk variants with small effect sizes, however we have identified mitochondrial SNPs that are significantly associated with IgA nephropathy, confirmed significant association following correction for multiple testing, and shown there is an excess burden of common mtDNA variants in the case group with 75 SNPs present only in individuals with IgA nephro-

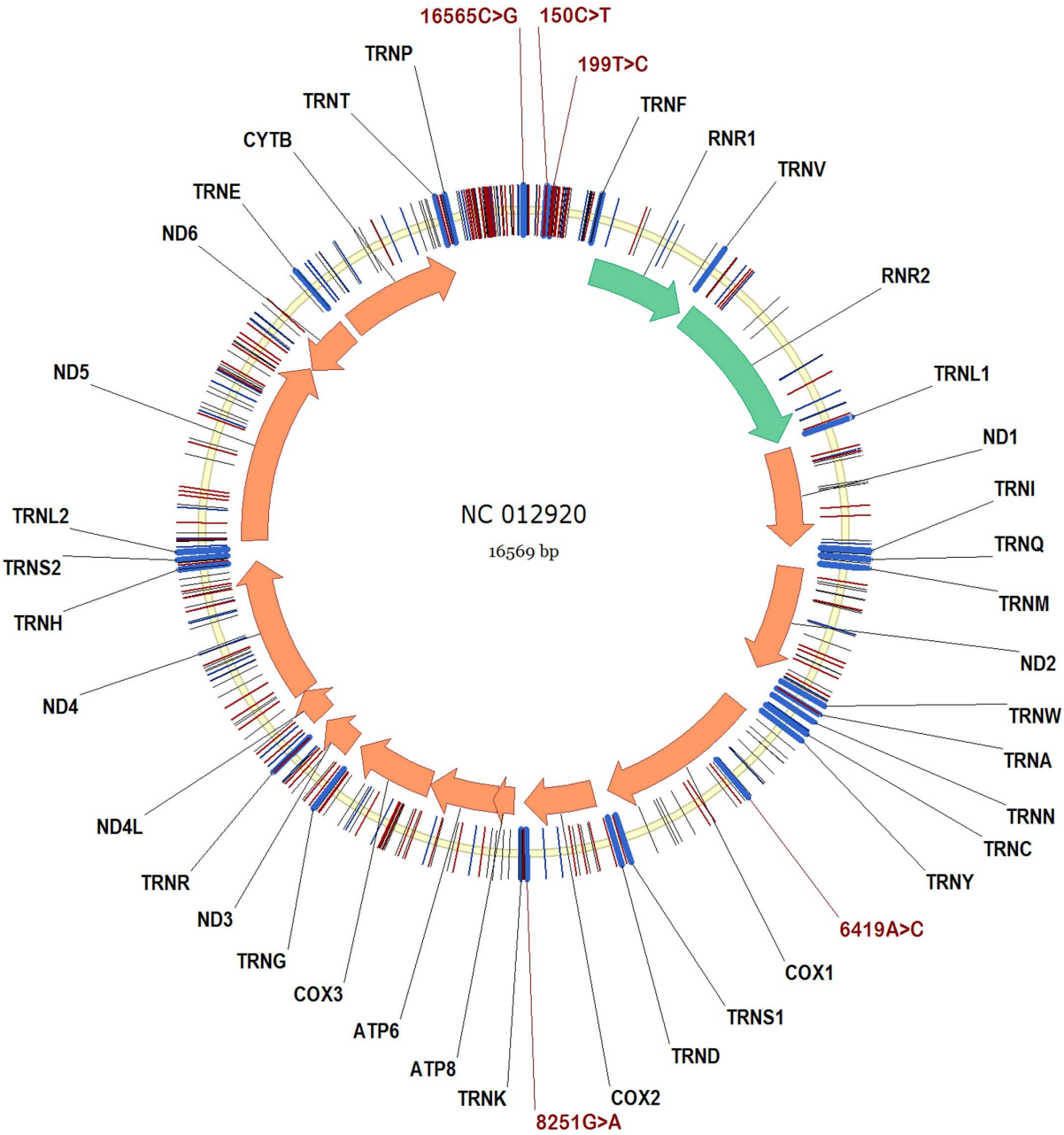

Figure 1 | Vector NTI image of all SNPs on mitochondrial genome. All genes are named in black text, with protein coding genes indicated by orange arrows and ribosomal RNA genes indicated by green arrows. Top five most significantly associated SNPs shown labelled in dark red. Other significant SNPs represented by dark red lines, common SNPs by blue lines, and rare variants by black lines.

Table 1 | SNPs demonstrating significant association with IgA nephropathy. HV2, Hypervariable segment 2; CO1, Cytochrome c oxidase subunit I; CO2, Cytochrome c oxidase subunit II; MAF, minor allele frequency

SNP base position	Location	MAF (%)	Alleles	Case frequency (%)	Control frequency (%)	P value for IgA nephropathy (n=32)	P value for ESRD (n=100)
150	MT-HV2	14	C:T	25.0	3.1	4.00×10^{-4}	>0.05
199	MT-HV2	9	T:C	18.8	0	3.00×10^{-4}	>0.05
6419	MT-CO1	22	A:C	40.7	3.7	3.67×10^{-6}	>0.05
8251	MT-CO2	14	G:A	25.0	3.1	4.00×10^{-4}	>0.05
16565	Control region	17	C:G	27.6	2.4	9.00×10^{-4}	>0.05

pathy. Additionally, we have developed an efficient, cost-effective protocol for comprehensive analysis of the mtDNA sequence with extensive technical validation of the experimental approach. SNPs in the mitochondrial genome are often excluded from genome-wide association studies so this study provides complementary information to ongoing genetic epidemiology research of autosomes. We have provided frequency data, the exact location of all identified SNPs, and details of the relationship between common SNPs.

From the 113 SNPs observed with a minor allele frequency greater than 5%, five common SNPs showed significant association with IgA nephropathy where $P \leq 10^{-4}$. Association of 6419A>C with IgA nephropathy was maintained following permutation testing to adjust for multiple testing bias. This transversion SNP 6419A>C is relatively common with a minor allele frequency of 22% and is located in codon 172 of the mitochondrial DNA encoded cytochrome c oxidase subunit 1 (*MT-CO1*) gene, creating an amino acid change from lysine to asparagine. The functional result of this change in amino acid is unknown. Cytochrome c oxidase, a component of the respiratory chain, catalyses the reduction of oxygen to water; CO1 is the catalytic subunit of the enzyme. *MT-CO1* expression has been shown to be increased in patients with chronic kidney disease and hemodialysis[16]. This increase may be part of a compensatory response to the increased ROS generation in the uremic state, as ROS can inhibit respiratory chain enzyme activity[17].

Regions of the mitochondrial genome demonstrate homology with human autosomes in the nuclear genome. It is possible that 'SNPs' identified in this study are derived from amplified regions not in the mitochondrial genome, but sequence alignment and SNP calling was performed (automated by the software and visual confirmation of results by two researchers) blinded to case versus control status so should have affected case and control samples equally. Off-target and duplicate reads were excluded from analysis; the coverage depth and quality of alignment to the mitochondrial genome was excellent, providing confidence in the results of this study. Additionally, we explicitly enriched for mtDNA based on long range PCR, called SNPs in duplicate for 32 samples that were independently amplified, and validated SNPs (>10% heteroplasmy) using Sanger sequencing with PCR primers designed to specifically amplify mtDNA and prevent co-amplification of nuclear DNA based on the human genome reference sequence GRCh37.

This study identified a number of novel variants significantly associated with IgA nephropathy using a case-control approach. Additionally, we have genotyped an independent cohort of 50 renal transplant recipients without IgA nephropathy, compared to 50 age and gender matched controls with no evidence of kidney disease, which supports genetic association with IgA nephropathy specifically. Ideally, future studies will be conducted using next generation sequencing to evaluate a larger cohort of patients with ESRD attributed to IgA nephropathy compared to healthy controls, utilising multi-centre collaborations.

Survival analysis was not performed as this study is not optimally designed to assess patient or graft survival. However, seventeen patients returned to dialysis following graft failure in this cohort; 48 SNPs, all with MAF >5%, were present only in those individuals whose kidney transplant failed.

Mitochondria also depend on a number of nuclear genome products for their function, and further study in this area might examine both mtDNA and nuclear genes associated with mitochondrial disorders. For example, a next generation sequencing, clinically targeted assay has been developed using RainDance and Illumina approaches to examine the mitochondrial genome and 108 selected nuclear genes implicated in mitochondrial disease[18]. Of note, mtDNA sequences can also be extracted from non-targeted mtDNA amplification, such as exome sequencing data, which provides substantial opportunities for further mtDNA research[19–21].

This study utilised a cost-efficient, high throughput approach to identify novel genetic variants in mtDNA associated with IgA nephropathy. Additionally, we revealed an excess burden of mtDNA variations in renal transplant recipients with ESRD, secondary to IgA nephropathy, compared with individuals who have no evidence of

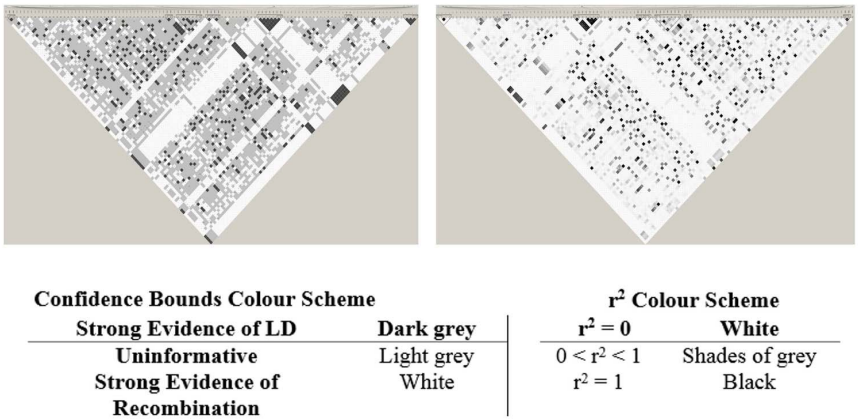

Figure 2 | Linkage disequilibrium plots for confidence intervals on D' and r^2 respectively where all SNPs had a minimum minor allele frequency of 5%.

renal disease. Building up genetic risk profiles may assist risk stratification of renal transplant patients, and aid the development of treatment strategies. These results are encouraging, confirming that further research is warranted in larger, carefully phenotyped cohorts.

Methods

Participants. The Regional Nephrology Unit at Belfast City Hospital is the only centre providing kidney transplant procedures in Northern Ireland. This single-centre site facilitates comprehensive analyses since DNA is available for all transplant recipient-kidney donor pairs in first deceased donor kidney transplants since 1986 with extensive prospective clinical follow-up[22,23]. DNA was extracted from blood samples obtained prior to kidney transplantation using the salting out method[23]. All experiments were performed in accordance with relevant guidelines and regulations. This study had full ethical approval from the Office of Research Ethics Committees, Northern Ireland (ORECNI reference: 08/NIR03/79). Sixty-four individuals were selected for mtDNA next-generation sequencing (NGS) of the entire mitochondrial genome. This consisted of transplant recipients (cases, n=32) with age and gender matched donors (controls, n=32). All transplant recipients had ESRD secondary to primary IgA nephropathy proven by immunofluorescence staining of renal biopsies. A further 50 renal transplant recipients with ESRD, but without IgA nephropathy (cases), were compared to 50 age and gender matched kidney donor (control) individuals with no evidence of kidney disease from Belfast.

Next-generation sequencing. The mtDNA of all DNA samples were individually amplified using long-range PCR (Qiagen, Crawley, UK LongRange PCR Kit), in two overlapping fragments (fragment 1 = 9,289 bp; fragment 2 = 7,626 bp) [15]. Agarose gel electrophoresis (0.7%) was used to confirm the presence of PCR product. PicoGreen quantitation (Life Technologies, Warrington, UK) was employed to quantitate PCR products for fragments 1 & 2, which were then pooled in equimolar amounts for each individual and barcoded for downstream analysis. Barcoding each sample facilitates inexpensive pooling of DNA to generate sequence data, which can be later extracted to analyse each individual separately using bioinformatic tools. Samples were prepared using the Illumina TruSeq NGS protocol according to manufacturer's instructions, enriched DNA was quantitated using Qubit (Life Technologies) and sequencing was performed using an Illumina Genome Analyzer II platform. DNA from the 32 individuals in the case group was separately prepared and sequenced in duplicate for optimisation and validation of the technique; once in a 32 sample pool, and again using two pools comprising 16 samples each. Targeted next generation sequencing data was analysed for SNPs associated with IgA nephropathy in an independent case-control cohort for ESRD.

Validation by Sanger sequencing. Four samples, two renal transplant recipients and two kidney donors, underwent comprehensive Sanger sequencing of the entire mitochondrial genome to validate the NGS results derived from Illumina. Sixty-two internal sequencing primers were employed alongside nine overlapping primer pairs to amplify adequately sized fragments for Sanger sequencing[24,25]. Bidirectional sequencing was performed on an ABI 3730 Genetic Analyser (Life Technologies)

Analysis. Initial alignment and quality control was performed using Partek Flow version 2.2.3, 64 bit architecture for pair-end FASTQ files (Partek Inc, Missouri, USA). Post-alignment quality control and analysis (QC/QA) was performed in Partek Genomics Suite version 6.6 (Partek Inc), in order to visualise the quality of aligned reads. Coverage across the mitochondrial genome was visualised using Partek Genomics Suite (Supplementary Figure 1). Partek Genomics Suite was used to call SNPs against the revised Cambridge Reference Sequence for mtDNA (rCRS, NC_012920), and an allele percentage test against the rCRS was also undertaken to evaluate heteroplasmy. Sequence data was analysed for quality control and the presence of SNPs blinded to case and control status. Concordance was analysed between duplicate samples. Case-control analysis was performed using the open source software Haploview[26]. Linkage disequilibrium plots for the IgA nephropathy cohort were generated using D' and r^2 measures in Haploview.

1. Caskey, F. *et al.* UK Renal Registry 2013: 16th Annual Report of the Renal Association. **125**, 54 (2013).
2. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* **23**, 147 (1999).
3. Alston, C. L. *et al.* A novel mitochondrial MTND5 frameshift mutation causing isolated complex I deficiency, renal failure and myopathy. *Neuromuscul Disord* **20**, 131–135 (2010).
4. Bai, Y. *et al.* Single nucleotide polymorphisms in the D-loop region of mitochondrial DNA is associated with renal cell carcinoma outcome. *Mitochondrial DNA* [Epub ahead of print] (2013).
5. Belostotsky, R. *et al.* Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. *Am J Hum Genet* **88**, 193–200 (2011).
6. Chen, J. B. *et al.* Sequence-based polymorphisms in the mitochondrial D-loop and potential SNP predictors for chronic dialysis. *PloS one* **7**, e41125 (2012).
7. Feigenbaum, A. *et al.* Novel mitochondrial DNA mutations associated with myopathy, cardiomyopathy, renal failure, and deafness. *Am J Med Genet A* **140**, 2216–2222 (2006).
8. Hatunic, M. *et al.* The Leu262Val polymorphism of presenilin associated rhomboid like protein (PARL) is associated with earlier onset of type 2 diabetes and increased urinary microalbumin creatinine ratio in an Irish case-control population. *Diabetes Res Clin Pract* **88**, 316–319 (2009).
9. Kollberg, G. *et al.* A novel homozygous RRM2B missense mutation in association with severe mtDNA depletion. *Neuromuscul Disord* **19**, 147–150 (2009).
10. Xu, J. *et al.* Single nucleotide polymorphisms in the mitochondrial displacement loop and age-at-onset of renal cell carcinoma. *Sci Rep* **3**, 2408 (2013).
11. Bai, Y. *et al.* Association of sequence polymorphism in the mitochondrial D-loop with chronic kidney disease. *Ren Fail.* **36**, 781–784 (2014).
12. Yamagata, K. *et al.* Mitochondrial DNA mutations in focal segmental glomerulosclerosis lesions. *J Am Soc Nephrol* **13**, 1816–1823 (2002).
13. Doleris, L. M. *et al.* Focal segmental glomerulosclerosis associated with mitochondrial cytopathy. *Kidney Int* **58**, 1851–1858 (2000).
14. Manwaring, N. *et al.* Population prevalence of the MELAS A3243G mutation. *Mitochondrion* **7**, 230–233 (2007).
15. Majamaa, K. *et al.* Epidemiology of A3243G, the mutation for mitochondrial encephalomyopathy, lactic acidosis, and strokelike episodes: prevalence of the mutation in an adult population. *Am J Hum Genet* **63**, 447–454 (1998).
16. Granata, S. *et al.* Mitochondrial dysregulation and oxidative stress in patients with chronic kidney disease. *BMC Genomics* **10**, 388 (2009).
17. Han, Z. *et al.* Shear-induced reactive nitrogen species inhibit mitochondrial respiratory complex activities in cultured vascular endothelial cells. *Am J Physiol Cell Physiol* **292**, C1103–C1112 (2007).
18. Dames, S. *et al.* The development of next-generation sequencing assays for the mitochondrial genome and 108 nuclear genes associated with mitochondrial disorders. *J Mol Diagn* **15**, 526–534 (2013).
19. Guo, Y., Li, J., Li, C. I., Shyr, Y. & Samuels, D. C. MitoSeek: extracting mitochondria information and performing high-throughput mitochondria sequencing analysis. *Bioinformatics* **29**, 1210–1211 (2013).
20. Samuels, D. C. *et al.* Finding the lost treasures in exome sequencing data. *Trends Genet* **20**, 593–599 (2013).
21. Ye, F., Samuels, D. C., Clark, T. & Guo, Y. High-throughput sequencing in mitochondrial DNA research. *Mitochondrion* [Epub ahead of print] (2014).
22. Courtney, A. E., McNamee, P. T. & Maxwell, A. P. The evolution of renal transplantation in clinical practice: for better, for worse? *QJM* **101**, 967–978 (2008).
23. Moore, J. *et al.* Association of caveolin-1 gene polymorphism with kidney transplant fibrosis and allograft failure. *JAMA* **303**, 1282–1287 (2010).
24. Ramos, A., Santos, C., Alvarez, L., Nogues, R. & Aluja, M. P. Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification. *Electrophoresis* **30**, 1587–1593 (2009).
25. Ramos, A. *et al.* Validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification: a revision based on the New Human Genome Reference Sequence (GRCh37). *Electrophoresis* **32**, 782–783 (2011).
26. Barrett, J. C. Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc.* **10** (2009).

Acknowledgments

DRV was supported by a PhD studentship from the NI Department of Education and Learning. We are grateful for financial support from the Northern Ireland Kidney Research Fund.

Author contributions

D.R.V., E.M.K., D.W.M. performed Illumina next generation sequencing. A.P.M. provided phenotyping expertise. A.P.D. & A.J.M. conducted sequencing and all analysis. A.P.M., A.P.D. & A.J.M. co-wrote the manuscript. All authors critically reviewed the manuscript.

Additional information

Extensive data is provided with this article and further information is available from the authors on request.

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Douglas, A.P. *et al.* Next-generation sequencing of the mitochondrial genome and association with IgA nephropathy in a renal transplant population. *Sci. Rep.* **4**, 7379; DOI:10.1038/srep07379 (2014).

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/