GeCIP Detailed Research Plan Form

August 2015

Background

The Genomics England Clinical Interpretation Partnership (GeCIP) brings together researchers, clinicians and trainees from both academia and the NHS to analyse, refine and make new discoveries from the data from the 100,000 Genomes Project.

The aims of the partnerships are:

1. To optimise:

- clinical data and sample collection
- clinical reporting
- data validation and interpretation.

2. To improve understanding of the implications of genomic findings and improve the accuracy and reliability of information fed back to patients. To add to knowledge of the genetic basis of disease.

3. To provide a sustainable thriving training environment.

The initial wave of GeCIP domains was announced in June 2015 following a first round of applications in January 2015. On the 18th June 2015 we invited the inaugurated GeCIP domains to develop more detailed research plans working closely with Genomics England. These will be used to ensure that the plans are complimentary and add real value across the GeCIP portfolio and address the aims and objectives of the 100,000 Genomes Project. They will be shared with the MRC, Wellcome Trust, NIHR and Cancer Research UK as existing members of the GeCIP Board to give advance warning and manage funding requests to maximise the funds available to each domain. However, formal applications will then be needed to individual funders. They will allow Genomics England to plan shared core analyses and the required research and computing infrastructure to support the proposed research. They will also form the basis of assessment by the Project's Access Review Committee, to permit access to data. Some of you have requested a template for the research plan which we now provide herewith.

We are only expecting one research plan per domain and have designed this form to contain common features with funder application systems to minimise duplication of effort. Please do not hesitate to contact us if you need help or advice.

Domain leads are asked to complete all relevant sections of the GeCIP Detailed Research Plan Form, ensuring that you provide names of domain members involved in each aspect so we or funders can see who to approach if there are specific questions or feedback and that you provide details if your plan relies on a third party or commercial entity. You may also attach additional supporting documents including:

- a cover letter (optional)
- CV(s) from any new domain members which you have not already supplied (required)
- other supporting documents as relevant (optional)

Genomics England Clinical Interpretation Partnership (GeCIP) Detailed Research Plan Form

Application Summary	
GeCIP domain name	**Eye and Ear**
Project title *(max 150 characters)*	**Proposed use of genomic data and eye and ear phenotypic data to increase understanding of disorders affecting eye and ear towards better counselling, treatment and trial design.**

Objectives. *Set out the key objectives of your research. (max 200 words)*

1) Known genes

i) The determination of the detailed phenotype of patients with inherited disease related to eye and ear. The cohort generated by the GE study will allow us to confine characterisation to specific molecular subtypes, and to determine natural history data and optimal metrics of progression to inform clinical trials.

ii) Unusual alleles – exome analysis does not detect all alleles of known genes in typical cases (eg one allele only in a typical recessive proband). Intronic and/or regulatory mutations are likely to be responsible; GE data will be invaluable to their identification

iii) Known genes, organ-specific alleles. Genes discovered to cause one disorder are often allelic with others. At least in ophthalmology, examples include many genes that can cause generalised ciliopathies, Usher type II, and CLN3 and 7 that otherwise cause neuronal ceroid lipofuscinosis. There will be more to discover and the phenotype-genotype correlations will inform our understanding of basic biological processes.

2) Unknown genes causative of eye and ear disease will be sought and confirmation made through our network of international collaborators. Mendelian disorders causing hearing and sight loss are highly heterogeneic and we anticipate a valuable cohort of unsolved families for new gene discovery

3) Complex disease. Age-related macular degeneration is the best understood complex neurodegenerative disorder in terms of susceptibility genes. By mining GE we wish to identify adults with high and low-risk genotypes and invite them to undergo detailed phenotypic investigation (i.e. adaptive optics imaging of the retinal pigment epithelium). The identification of a premorbid phenotype will inform our knowledge of the initial pathogenesis and identify a metric to identify and monitor individuals who could be offered preventive treatments. Secondly, specific inflammatory disorders of the eye and ear are likely to involve the auto-immunity towards organ-specific epitopes. We are in parallel performing VDJ sequencing on such disorders (not eligible for GE) and will seek to identify HLA matched individuals and perform VDJ sequencing on their leukocyte DNA as an age HLA-matched control. This will increase the likelihood of finding specific enrichment of TCR and AB variable domains in the affected patients.

4) Modifiers . In the majority of inherited diseases in this domain, there is a great variability in severity despite the same genotype within and between families. For instance, in some genes causative of retinal degeneration mutations can be blinding in some and completely non-penetrant in others (eg RP11 due to mutation in PRPF31). Modifiers matter, and their identification will point towards novel treatments or at least methods of amelioration. Possible strategies are described in the individual research sections.

5) The sensitivity of clinical features. Our proposed program of work will allow the determination of those human phenotype ontology (HPO) terms that are most predictive of the underlying causative gene. For instance with our emerging resource of patients with known gene mutations

we will be able to determine the Probability(HPO term A| gene mutation B). However, what will inform most future genetic studies will be the reverse probability - Probability(gene mutation B | HPO term A) which we will endeavour to infer from our data. Understanding the relationship between specific clinical features and underlying genetic cause will have massive implications for all who seek to determine the likely causation of variants in families. This will also inform clinicians to chose how best to phenotype patients and how best to conduct clinical trials . That is, which clinical features should be gathered to be most predictive of molecular cause when faced with a number of plausible variants in candidate genes? The genetic complexity inherent in seeking the molecular diagnosis in families will grow as WGS becomes mainstream, which will increase still further the importance of a detailed understanding of gene-phenotype association.

Lay summary. *Information from this summary may be displayed on a public facing website. Provide a brief lay summary of your planned research. (max 200 words)*
Eye only so far (Maria please add your part for the ear). Too many words at present. Need to put elsewhere.

Advances in the study of inherited eye and ear disease include the discovery of over 180 causative genes for retinal disease alone. This poses challenges in molecular diagnosis and requires the sequencing of many genes in any one family to diagnose the cause of their specific inherited disorder. High-throughput sequencing surmounts this by determining the sequence of all genes in parallel and the GE project will improve our ability to find previously 'hidden' spelling mistakes in known genes, as well as novel genes. In the eye, for example, the molecular diagnosis can be associated with highly detailed information on its structure and function as it is peculiarly accessible to examination. Direct imaging at the cellular level is possible, non-invasively, in the living patient. Moreover, the accessibility allows the institution of gene-directed therapy. In the ear, the advent of cochlear implantation has already transformed the lives of many with inherited deafness. Already gene-replacement trials, using conventional eye surgery and genes carried by safe viruses, have been undertaken for two blinding retinal disorders, with more under trial. Information from GE will inform ongoing treatments and future trials by identifying relevant subgroups of patients, determining the best clinical measurements that allow optimal trial design and improving our ability to understand why some family members are more or less severely affected than others affected by the same spelling mistake in the same gene.
Complex and more common eye diseases, particularly age-related macular degeneration (AMD) are better understood than any other age-related degenerative disease in terms of their at-risk genes. Mining GE data will allow the identification of young adults with particularly high- and low-risk gene changes. Non invasive investigation of these persons will allow us to determine the first pathological changes that precede vision loss in AMD. Importantly, successful preventative drugs will be most effective if implemented early, before the onset of irreversible degeneration. Only by understanding the earliest changes will we be able to determine best measures to undertake trials of intervention, such as those that effect the complement system (a part of the immune process now known to cause the disorder).

Technical summary. *Information from this summary may be displayed on a public facing website. Please include plans for methodology, including experimental design and expected outputs of the research. (max 500 words)*

The complex genetic landscape of those affected by Mendelian disorders causing blindness and deafness means that the data generated by whole genome sequencing will have a major impact on understanding the alleles of known genes, on the detection of novel genes and on the identification of genetic modifiers that can affect whole systems (eg Usher syndrome versus non-syndromic deafness, Joubert syndrome versus non-syndromic blindness) as well as cause non-

penetrance in some gene carriers. The deafness and blindness investigator communities already form effective networks in the UK, with significant links to international collaborators. In the world of eye genetics, individuals including clinicians, scientists and patient representatives work together within the Eye Genetics Group, which is closely aligned to the 'Sight' part of this GeCIP. Moreover, clinical grade molecular diagnosis is well advanced in retinal and cataract genetics (Manchester Centre for Genomic Medicine) and deafness (North Thames Clinical Genetics Laboratory) and so the interpretation of next-generation sequencing data in these highly heterogeneic disorders is well advanced, which we feel will increase the efficiency of gains from the GE endeavour.
Individual detail is given for six areas of study. Firstly, the investigation of Mendelian disorders eligible for inclusion in Genomics England including deafness, retina, ocular maldevelopment and optic neuropathy. Moreover, we propose two uses of GE data to inform the understanding of two classes of disorders, not eligible for inclusion in GE. This is, i) the identification and phenotyping of presymptomatic persons with high-risk genotypes for age-related macular degeneration and ii) the use of HLA matched individuals to use as controls in the investigation of the antigen-binding repertoire in inflammatory eye disease.

Expected start date	**Start of GE data**
Expected end date	**Through the term of GE study recruitment and beyond!**

Lead Applicant(s)	
Name	Andrew Webster
Post	Chair Molecular Ophthalmology UCL, Consultant Ophthalmologist Moorfields Eye Hospital
Department	Genetics
Institution	Moorfields Eye Hospital, London EC1V 2PD
Current commercial links	None

Lead Applicant(s)	
Name	Maria Bitner-Glindzicz
Post	Professor of Clinical and Molecular Genetics, UCL Institute of Child Health
Department	Genetics and Genomic Medicine Programme, UCL Institute of Child Health, 30 Guilford St, London WC1N 1EH
Institution	UCL Institute of Child Health
Current commercial links	None

Administrative Support	
Name	Samantha Malka
Email	Samantha.lawrence@ucl.ac.uk
Telephone	0207 566 2260
Name	Loraine Daly
Email	loraine.daly@ucl.ac.uk
Telephone	0207 905 2608

Subdomain leads		
Name	**Subdomain**	**Institution**
Bitner-Glindzicz	Hearing	UCL/ GOSH

Yu Wai Man/Votruba	Optic Nerve	Newcastle/Cardiff
Sowden/ FitzPatrick/Ragge	Developmental Disorders	GOSH/ Edinburgh
Webster/ Black	Retina	UCL/Manchester
Tuft/Hardcastle/Moore	Cornea	UCL/Moorfields
Chain/Lee/Dick/Denniston	Inflammatory Eye Disease	London/Bristol/Birmingham
Sivaprasad/Cipriani/Abecasis	Age-related Macular Disease	Moorfields/UCL/Uni Michigan

Data requirements

Data scope. *Describe the groups of participants on whom you require data and the form in which you plan to analyse the data (e.g. phenotype data, filtered variant lists, VCF, BAM). Where participants fall outside the disorders within your GeCIP domain, please confirm whether you have agreement from the relevant GeCIP domain. (max 200 words)*

We plan to request the complete germline genotype data (VCF format) for all samples that fall within the "Eyes and Ears" GeCIP combined with the complete set of HPO terms associated with these samples (in whatever format that is used for these type of data through the openClinica platform). We would be interested in similar data from cases recruited that fall within other GeCIPs in which a syndrome includes an ocular abnormality such as those within Paediatrics, Nephrology or Neurology. We would value direct access to the bam files for visualisation or raw data using such tools as IGV as this is invaluable in the confirmation of variants highlighted through the pipeline.

Data analysis plans. *Describe the approaches you will use for analysis. (max 300 words)*

In the simplest instance, we plan to use R to run basic genotype-phenotype correlations. R can also be used to parse VCF files with the appropriate libraries. The R package developed my Daniel Greene and Ernest Turro, SimReg, associates rare exonic variants in phenotypically similar blood platelet disorders cases for which detailed Human Phenotype Ontology are readily available. Existing tools like Exomiser which use external datasets, such as model organism databases, disease gene network and pathway databases are also capable of prioritising variants. In time, case sharing databases such as PhenomeCentral and Decipher could also be queried to find phenotypically similar patients.

Key phenotype data. *Describe the key classes of phenotype data required for your proposed analyses to allow prioritisation and optimisation of collection of these. (max 200 words)*

We are interested in exploring in depth correspondence between phenotypic and genotypic data. Hence, we would like to download the full set of HPO terms associated with each clinical samples that falls within the "eyes and ears" GeCip. These data will be used to discover novel disease genes but also to identify novel and clinically relevant phenotype/genotype correlations.

Alignment and calling requirements. *Please refer to the attached file (Bioinformatics for 100,000 genomes.pptx) for the existing Genomics England analysis pipeline and indicate whether your requirements differ providing explanation. (max 300 words)*

As far as we can tell, we are fine with the Genomics England analytical pipeline. We would like to confirm that CNV calls from software like MANTA or CANVAS will be made available so that we can make full use of the rich analysis that WGS enables. We would also ideally be able to query BAM files using IGV, but this appears to be readily available on the Genomics England platform. Some important regions of the genome are either intractable to WGS biochemistry or pose challenges to alignment, and a paradigm for this is the ORF15 exon on RPGR. We hope to use specific realignment strategies of this region in some cases where this gene is suspected.

Tool requirements and import. *Describe any specific tools you require within the data centre with particular emphasis on those which are additional to those we will provide (see attached excel file List_of_Embassy_apps.xlsx of the planned standard tools). If these are new tools you must discuss these with us. (max 200 words)*

Our main concern relates to CNV calling algorithm, which we do not see in the lists of embassy apps. We find that these data have a large practical impact and not having such calls is concerning.

Data import. *Describe the data sets you would require within the analysis environment and may therefore need to be imported or accessible within the secure data environment. (max 200 words)*

We would like to import additional reference datasets that are currently being built outside of Genomics England that capture our current knowledge of genotype-phenotype correlations. These data rely on HPO consistent with Genomics England annotations.

Computing resource requirements. *Describe any analyses that would place high demand on computing resources and specific storage or processing implications. (max 200 words)*

We will run correlation analyses between HPO terms and VCF files. We do not anticipate these computations to be substantial and they should fit without much concern within the Genomics England infrastructure.

Omics samples

Analysis of omics samples. *Summarise any analyses that you are planning using omics samples taken as part of the Project. (max 300 words)*

At the time or writing there are no concrete plans to use serum or plasma samples as part of the Eye Ear GeCIP. However, intriguingly, our suite of disorders can often be caused by mutation of ubiquitously expressed genes, such as those involved in metabolism (eg dehydrodolichol diphosphate synthase DHDDS, mevalonate kinase, alpha-beta hydrolase domain-containing protein 12 ABHD12), lysosomal function (CLN3 and MFSD8), fatty-acid metabolism (CYP4V2). In such families, the assay of specific biomarkers would be biologically informative. A more cohesive research plan will depend upon the accumulation of families with these specific disorders as the study proceeds.

Data access and security	
GeCIP domain name	Eye and Ear
Project title *(max 150 characters)*	**Proposed use of genomic data and eye and ear phenotypic data to increase understanding of disorders affecting eye and ear towards better counselling, treatment and trial design.**

Applicable Acceptable Uses. *Tick all those relevant to the request and ensure that the justification for selecting each acceptable use is supported in the 'Importance' section (page 3).*

X *Clinical care*

X *Clinical trials feasibility*

X *Deeper phenotyping*

X *Education and training of health and public health professionals*

X *Hypothesis driven research and development in health and social care - observational*

☐ *Hypothesis driven research and development in health and social care - interventional*

X *Interpretation and validation of the Genomics England Knowledge Base*

☐ *Non hypothesis driven R&D - health*

☐ *Non hypothesis driven R&D - non health*

☐ *Other health use - clinical audit*

☐ *Public health purposes*

X *Subject access request*

☐ *Tool evaluation and improvement*

Information Governance

X The lead and sub-leads of this domain will read and signed the Information Governance Declaration form provided by Genomics England and will submit by e-mail signed copies to Genomics England alongside this research plan.

Any individual who wishes to access data under your embassy will be required to read and sign this for also. Access will only be granted to said individuals when a signed form has been processed and any other vetting processes detailed by Genomics England are completed.

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**i) The determination of the presymptomatic phenotype of age-related macular disease. ii) Disentangling the causal variants on chromosome 10q.**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Age-related macular degeneration (AMD) is a progressive neurodegenerative retinal disease resulting in irreversible visual loss. It is the most prevalent cause of blindness in the Western world affecting over 5% of those over 75 years old (Chakravarthy et al 2010). The disease progresses from an early stage of asymptomatic accumulation of extracellular deposits between the retinal pigment epithelium (RPE) and Bruch's membrane (drusen) to advanced stage of geographic atrophy (GA) caused by central retinal atrophy or the growth of choroidal vasculature that invade the retina (choroidal neovascularisation or CNV). While treatments exist to inhibit the damage caused by CNV, this is likely to delay the onset of visual loss rather than abrogate it. Moreover, no effective treatments are available for the atrophic process.

Genetic factors confer significant AMD risk and the disorder is arguably the best understood in terms of genetic susceptibility of all complex disorders. For a review of the causative genes see Nature Genetics - December 2015; doi:10.1038/ng.3448 and citations therein. This recent genome wide association study, using dense cSNP arrays identified 52 independently associated common and rare variants ($P < 5 \times 10^{-8}$) distributed across 34 loci on a cohort of 16k patients and 17k controls from a number of different centres. It is clear that two of these 34 loci contribute the largest population attributable risk (that is harbour common variants with a high odds ratio) at the region encoding complement factor H (CFH) on 1q and that encoding variants in two genes at chromosome 10q. High risk alleles for 1q are common and comprise 40% of alleles in the UK population. Although, the genetic architecture of CFH is complex, one coding SNP, that of p.Y402H captures much of the risk. Consequently, approximately 16% of the population will be homozygous for a significant risk factor at this locus. At 10q the population attrbutable risk is highest. The risk allele can be unambiguously identified by variants in two neighbouring genes, ARMS2 and HTRA1, which are in high linkage disequilbrium. Because of this LD, it has not so far been possible to attirbute the risk to variation in expression or function of the protease HtrA1 on the one hand, or expression/function of the orphan mitochondrial protein encoded by ARMS2 on the other. However, the risk allele is easily identified and again comprises about 40% of alleles in the UK population. Hence homozygotes will comprise 16% and double homozygotes for high risk alleles at both alleles 2.5%.

Other genes are clearly involved and a person's risk of developing visual loss due to AMD can be more accurately determined through genotyping at these, which include Complement factor 3, complement factor 2/B and apolipoprotien E amongst others (Seddon et al 2015)

The Genomics England data-set will provide an opportunity to find participants with a highest-risk and lowest risk genotype, at an age before AMD has caused symptoms in order to explore the premorbid retinal phenotype.

Secondly, the locus at 10q is confounded by the issue of linkage disequilibrium such that it is not known whether matrix turnover (Htra1) or mitochondrial function (ARMS2 – Fritsche 2008), is the true culprit. This impedes the development of novel therapies. Again by mining the GE data we hope to find those who are discordant for the risk SNPs at the 10q locus, as well as individuals with rare null alleles. Depending on the prevalence of such individuals, we will invite those adults for examine and compare their retinal structure and function.

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate this.*

Participant selection

i) High and low risk individuals – these can be selected on the basis of the top loci for AMD. For instance approximately 2.5% of the entire cohort will be homozygous for risk alleles at both 10q and 1q. Of the GE cohort this will comprise approximately 10^5 x 0.025 = 2560 participants. We will select those in the late 40s, early 50s and prioritise them by their risks at other major loci. We would aim to select gender and age-matched individuals with the opposite, protective genotypes. Overall, depending on available research funds we would hope to select at least 20 highest-risk and 20 lowest-risk willing participants for detailed retinal imaging and psychophysical testing.

ii) Specific genotypes at 10q – a review of GE WGS data will allow the rapid appraisal of those with different risk alleles for ARMS2 and Htra1. Given the near complete LD of this region (this too can be estimated independently in the GE data), the numbers are likely to be small. However, those of late age (seventh decade and over) will be informative for selection and other siblings who might also share the rare discordant allele be genotyped where possible. By selecting those homozyogus for one and heterozygous for another, and phentyping them deeply, one has the power to gain insights into which of the two genes is most likely the culprit for AMD susceptibility. Also, we shall mine the data for those with rare deleterious mutations in either gene, to prioritise, invite and examine in a similar way.

Phenotyping

Selected participants will be invited to attend MEH for one to two daily visits.

Non invasive imaging using optical coherence tomography to assess retinal thickness and en face adaptive optics will allow the detailed retinal structure to be examined. These techniques are available at UCL/MEH and recently techniques to image the RPE, previously hidden by the reflections from photoreceptors have

been determined (Scoles et al 2013). These procedures are entirely non-invasive requiring only mydriatic drops prior to the imaging.

We intend to examine for early and subtle sensitivity loss using scotopic perimetry using a recently acquired Medmont perimeter designed to assess retinal function at precise locations under scotopic conditions. Scotopic sensitivity is known to be affected early in AMD (Steinberg et al 2013 and citations therein) and it will be important to determine the utility of this metric in high-risk individuals.

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

This study intends to mine the whole GE dataset to find participants that will be informative for the further study of the risk factors of this important disease, as detailed above. Hence, no collaborations with other GeCIPs are planned.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*
Sobha Sivaprasad has a track record in AMD phenotyping and has recently acquired the specialist equipment to perform scotopic perimetry
Michel Michaelides has set up a SLO/ adaptic optics facility at IOO.
Valentina Cipriani (UCL) and Goncalo Abacasis have a track record in the investigation of genetic risk factors in AMD.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

Understanding the initial pathology associated with disease risk will have a high impact on the field by
i) improving our understanding of the pathobiology,
ii) directing molecular therapies to the correct pathological process(es) and
iii) gathering key phenotypic data to help determine those at clinical risk and to whom preventative measures should be directed.

References. *Provide key references related to the research you set out.*

Chakravarthy U, Evans J, Rosenfeld PJ. Age related macular degeneration. BMJ. 2010 Feb 26;340:c981. doi: 10.1136/bmj.c981.

Fritsche LG, Loenhardt T, Janssen A, Fisher SA, Rivera A, Keilhauer CN, Weber BH. Age-related macular degeneration is associated with an unstable ARMS2 (LOC387715) mRNA. Nat Genet. 2008 Jul;40(7):892-6. doi: 10.1038/ng.170.

Scoles D, Sulai YN, Dubra A. In vivo dark-field imaging of the retinal pigment epithelium cell mosaic. Biomed Opt Express. 2013 Aug 23;4(9):1710-23. doi: 10.1364/BOE.4.001710.

Seddon JM, Silver RE, Kwong M, Rosner B. Risk Prediction for Progression of Macular Degeneration: 10 Common and Rare Genetic Variants, Demographic, Environmental, and Macular Covariates. Invest Ophthalmol Vis Sci. 2015 Apr;56(4):2192-202. doi: 10.1167/iovs.14-15841.

Steinberg JS, Fitzke FW, Fimmers R, Fleckenstein M, Holz FG, Schmitz-Valckenberg S Scotopic and Photopic Microperimetry in Patients With Reticular Drusen and Age-Related Macular Degeneration. JAMA Ophthalmol. 2015 Jun;133(6):690-7. doi: 0.1001/jamaophthalmol.2015.0477.

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	The determination of the antigen-binding repertoire using massively parallel sequencing of lymphocyte V(D)J segments in patients presenting with specific forms of inflammatory eye disease – the parallel analysis of HLA-matched controls amongst the GE cohort.

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Inflammatory eye disease is a significant cause of visual loss in all populations. Present treatments consist of immunosuppressive therapies which have limited effectiveness, low specificity and significant ocular and systemic side effects. Improved, less toxic treatments are therefore an important goal. An autoimmune basis for these diseases is suspected and is supported by the fact that persons with specific proteins expressed on cells that interact with the immune system (MHC proteins) are particularly susceptible. For instance Birdshot choroidopathy is consistently associated with HLA-A29 and acute anterior uveitis with HLA-B27.

A collaboration involving investigators at UCL, Moorfields Eye Hospital, Bristol and Birmingham has recently secured funding from fight for sight to analyse the antigen binding repertoire of blood lymphocytes in patients presenting with three specific forms of uveitis, Birdshot choroidopathy, Acute Zonal Occult Outer Retinopathy (AZOOR) and idoipathic anterior uveitis. This study will harness the experience of Professor Benny Chain, who is at the vanguard of the technology and has successfully defined the repertoire in HIV patients before and after treatment as well as healthy volunteers (Heather et al 2016). This study will compare the T-cell receptor binding repertoires from blood and inflamed eye fluid (the anterior chamber of uveitis patients) to controls and also between different times in the course of the diseases. This is done through the selective amplification of RNA only from the transcripts derived from TCR lymphcyte loci. Because lymphocytes are unique in their genomic rearrangement, this technique will only amplify RNA from the two major specific loci involved in the formation of T-Cell receptors (TRA chr 14 and TRB chr 7). Moreover, it will determine the many thousands of such sequences, unique to the individual concerned, and provide an inventory of previous antigen exposure. For a review of the loci involved in antigen-recognition see Schatz et al 2011.

As a further resource of control individuals we propose to mine the whole GE dataset for those individuals with similar HLA A, B, C genotypes and then perform identical VDJ sequencing on RNA from those individuals. An aim is to identify consistent differences in the TCR binding repertoire in affected patients that will indicate binding to specific epitopes that predispose to the disease. This will not only allow the identification of biomarkers in patients presenting with uveitis but also lead to the characterisation of specific pathogens that trigger inflammation in susceptible patients, an important requirement for the advancement of the field.

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

For the main project the following cohorts will be recruited. For each case of uveitis we intend to mine the GE data for the best matches against the MHC class I genes, and for each propose a prioritised lists of participants that match each participant. This by definition will select HLA-A29 controls for those presenting with Birdshot retinopathy and HLA-B27 for those presenting with acute anterior uveitis. There is no known HLA association of Acute Zonal Occult Outer Retinopathy. Where possible we will apply for ***500ng of RNA*** from the PAXgene samples from these participants (usually 1 tube of 2.5ml of blood yields 3 ug of RNA, so we would be asking for 1/6 of the total stored at GE for each participant).

We intend to recruit 60 cases and so will be looking for the best match for each of these with GE, hoping then to obtain 500ng of RNA from at least ***60 GE participants***.

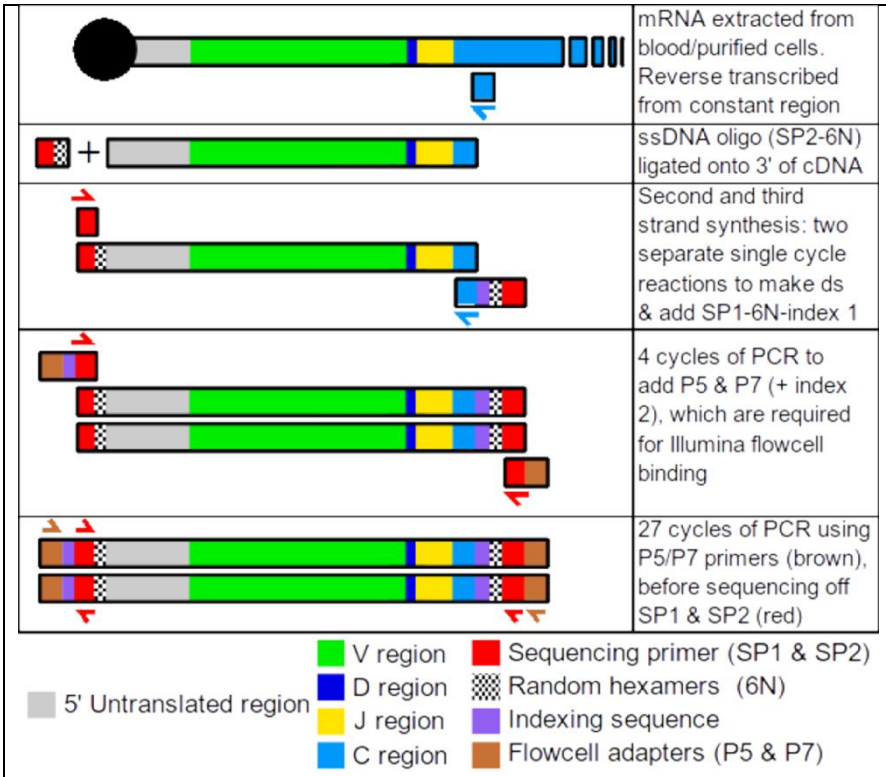
Cohorts from main study (outwith GE):

1) Cohorts of cases previously presenting with BCR and AZOOR and age-matched controls. We are confident that, given the clinical activity of the three units and existing databases, to recruit 15 patients with each disorder (total 30 cases). We hope to recruit unaffected siblings where possible (25% of which will be HLA identical).

2) Cases presenting with the first acute presentation of BCR and AZOOR, a total of 10 patients each, at three time-points – presentation, 3 months and 6 months. (total 20 cases). Again we hope to contemporaneously recruit siblings as above.

3) It is likely that TCRs that are associated with the inflammatory disorder will sequester in the organ that is inflamed. Acute anterior uveitis, a relatively common disorder in clinical practice will allow us to examine TCRs from the anterior chamber and compare with blood lymphocytes. From previous work (AK Denniston), in a series of 16 patients with AAU, in which between 50-100ul of aqueous humour was taken, a mean of 3870 (SD - 2645) cells were recovered of which 35.1 (SD 18.8)% were lymphocytes. This will allow us to harvest RNA from these small samples using the protocol described previously (Dash et al 2011) before further processing and sequencing. We intend to recruit at least 10 patients, taking paired contemporaneous anterior chamber and blood samples from the patients, (10 cases).

The library preparation for next-generation sequencing is well developed in the Chain lab and represented by the schematic below:



The fastq files generated from sequencing experiments will be analysed using the Decombinator program (Thomas et al. 2013 - http://www.innate2adaptive.com/software) and global statistics of complexity, richness and inequality of abundance calculated for the specific samples. Clonotypes that differ consistently from patient and control and show a correlation with disease regression in longitudinal samples will be identified. Of these, CDR3 sequences and V segments (contain CDR1 and 2) that show homology between patients will be sought. These principals have already been used to show the decrease in the diversity of the TCR repertoire in untreated HIV patients and their change over time with anti-retroviral therapy (Heather et al 2016)

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

The aim is to select those with matched MHC Class I genotypes for TCR sequencing irrespective of the reason for their inclusion in GE. Those participants that have a primary disorder affecting T-Cell function would be excluded from recruitment although this would be apparent on a reduced TCR diversity upon VDJ sequencing.

However, using the WGS data and RNA to archive the T-cell binding repertoire of participants would be an informative use of the samples, and give an unprecedented database of human T-cell binding diversity. This would allow correlations to be made with HLA genotype, and clinical metrics such as immunisation and significant systemic infection. A similar strategy can be done for B-cell derived Ig heavy and light loci. We would welcome cross-GeCIP collaborations in this regard, particularly given the expertise of the Chain group at UCL.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

Andrew Webster, Andrew Dick, Richard Lee and Alistair Denniston are clinical members of the Eye-Ear GeCIP and command access to an unprecedented population of patients presenting with the three forms of uveitis being investigated in this study within three large ophthalmology centres (Moorfields, Bristol and Birmingham). The procurement of aqueous humour samples from patients with HLA-B27 AAU, and blood samples from people with this disorder, birdshot choroidopathy and AZOOR falls within the remit of an approved and active multi-centre ethics proposal which will cover the necessary samples for the patient samples. This study, 'Immune mechanisms in the ocular environment', (REC ref O6/Q2702/63) is portfolio adopted (UKCRN4654) and the named investigators, based at the University of Birmingham includes Alastair Denniston.

Benny Chain has an expertise in sequencing T-cell receptor genes, at the RNA level in mouse and human. His lab has refined the biochemistry of RNA-Seq of the alpha and beta chains from blood leukocytes as well as the analysis of the RNA reads. A schema of the biochemistry is shown above, and the method described in detail in Heather et al 2016 and the underlying analysis in Thomas et al 2013.

References. *Provide key references related to the research you set out.*

Dash P wt al. Paired analysis of TCR[alpha] and TCR[beta] chains at the single-cell level in mice. 2011. J Clin Invest 121(1):288

Heather JM, Best K, Oakes T, Gray ER, Roe JK, Thomas N, Friedman N, Noursadeghi M, Chain B. Dynamic Perturbations of the T-Cell Receptor Repertoire in Chronic HIV Infection and following Antiretroviral Therapy. Front Immunol. 2016 Jan 11;6:644. doi: 10.3389/fimmu.2015.00644

Thomas N, Heather J, Ndifon W, Shawe-Taylor J, Chain B. Decombinator: a tool for fast, efficient gene assignment in T-cell receptor sequences using a finite state machine. Bioinformatics. 2013 Mar 1;29(5):542-50

D.G. Schatz, Y Ji, Recombination centres and the orchestration of V(D)J recombination. Nat Rev Immunol. 11:251-63 (2011)

Detailed research plan

Full proposal (total max 1500 words per subdomain)

Title *(max 150 characters)*	**Optic neuropathies**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Inherited optic neuropathies affect 1 in 10,000 people in the UK and currently, a genetic diagnosis is only achieved in 40-50% of cases. This is a major limitation for accurate genetic counselling and the development of therapeutic strategies. They are a group of blinding genetic disorders in which optic atrophy secondary to loss of retinal ganglion cells is a clinical key feature. The commonest causes world-wide are mutation in mitochondrial DNA (causing Leber's Hereditary Optic Neuropathy) and OPA1 mutations (causing Autosomal Dominant Optic Atrophy: ADOA). 60-80% of patients with autosomal dominant optic atrophy have mutations in the OPA1 gene. There are a slowly increasing number of additional genes that may also be implicated. However, 20-40% of patients remain without a molecular diagnosis. We need to identify the missing disease genes in these patients in order to be able to offer realistic and accurate genetic counselling and in order to start to identify potential therapeutic targets.
For every one patient given a new molecular diagnosis the clinical impact can be huge in terms of the possibilities for genetic counselling for them and their wider family. If we want all patients to eventually benefit from the genomic revolution and new opportunities for personalised medicine we need to find the disease genes and understand the disease mechansism. A molecular diagnosis for our patients will be a key step in their management, since without this all doors into our understanding of the disease pathophysiology and the rational design or targeting of novel therapeutic interventions are closed.

In order to realise and harness these opportunities, molecularly typed cohorts of patients need to be determined efficiently. Moreover, natural history data and precise phenotypic measures are required on such cohorts to inform the design and interpretation of such trials.

There is still a need to establish robust data on natural history and on long-term prognosis.

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

1. Phenotyping data and genotype-phenotype correlations.
Patients collated and entered in the study would be very suitable to investigate further with additional phenotyping strategies and novel techniques. So-called deep phenotyping is required to generate more biologically meaningful end-points for clinical trials and as biomarkers of disease staging and progression. This additional phenotyping will allow us to further stratify the patients and use the genetic data to compare and contrast genes and phenotypes.
2. Finding new genes.
The need to find new disease causing genes and alleles in these cohorts of patients without variants in known genes remains a priority. There may be patients with significant familial history of the disease and these would form the first tranche of patients suitable to identify disease causing genes. Nevertheless, the new technologies will enable a considerable prospect of novel disease gene discovery, which will need to be validated =in other patient cohorts.

In addition to analysing the recruited cohort as a whole, we will also sub-categorise patients with inherited optic neuropathies into the following groups: (i) suspected mode of inheritance: dominant, recessive and mitochondrial; (ii) isolated and syndromic optic atrophy; and (iii) specific ethnic groups.

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

We will collaborate with the NIHR RD-TRC, and have obtained funding for the initial deep phenotyping of a small cohort of patients which helps determine natural history data and sensitive metrics for

Further suggested collaborations: with charity and patient groups
LHON Patient Group
Fight for Sight
Rare Disease Alliance UK

Collaboration and data sharing with Neurology GeCIP

We will be interacting closely with other key members of the EURION (European Project on Inherited Optic Neuropathies) network. As a group, we have recently identified *RTN4IP1* as a new causative gene for autosomal recessive optic atrophy (Angebault et al).

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*

Dr Ana Majendar, a NIHR RD-TRC Clinical Research Fellow, 2014-2016, has been working to establish information relating to the deep phenotyping of patients with the mitochondrial optic neuropathies, Leber's hereditary optic atrophy and dominant optic atrophy. It is planned that further opportunities for research training will arise with the continued funding from NIHR- RD-TRC and other research arising from these initiatives.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

Mr Patrick Yu-Wai-Man is a Clinical Senior Lecturer and Consultant Neuro-ophthalmologist with a specialist interest in inherited optic neuropathies and mitochondrial eye disorders. He is affiliated with both the Newcastle and Moorfields NIHR BRCs.

Professor Marcela Votruba is a Clinical Ophthalmologist at the University Hospital Wales, Cardiff and Head of the Cardiff University School of Optometry & Vision Sciences. Her specialist interest is in mitochondrial optic neuropathies and inherited retinal dystrophies.

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

The data will be analysed in the context of the patient's phenotype and the wider cohort of patients with inherited optic neuropathies that have been recruited nationally into the GeL study. While the genomic landscape causing optic neuropathies is less well developed than say that causing retinal degeneration, both Yu Wia Man and Votruba have unrivalled experience as well as access to an international network of colleagues with similar expertise.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

The clinical and genetic data will provide immediate benefit to the counselling of patients and families with optic neuropathy by determining segregation and a likely prognosis. Importantly, two gene-directed trials have started in the field which we expect to grow during the lifetime of the GE project.

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

GenSight – phase III gene therapy trial for patients with Leber hereditary optic neuropathy (LHON) harbouring the m.11778G>A mitochondrial DNA mutation.

Edison – phase II trial to evaluate the efficacy of EPI-743 in dominant optic atrophy secondary to pathogenic *OPA1* mutations

References. *Provide key references related to the research you set out.*

Yu Wai Man P, Votruba M, Moore AT, Chinnery PF. Treatment strategies for inherited optic neuropathies – past, present and future. Eye. 2014;28(5):521-537. [Review]

Angebault C, Guichet P, Talmat-Amar Y, Charif M, Fares-Taie L, Gueguen N, Halloy F, Moore D, Amati-Bonneau P, Manes G, Hébrard M, Boquet B, Quiles M, Piro-Mégy C, Teigell M, Delettre C, Rossel M, Meunier I, Priesing M, Lorenz B, Carelli V, Chinnery PF, Yu-Wai-Man P, Kaplan J, Roubertie A, Barakat A, Bonneau D, Reynier P, Rozet J, Bomont P, Hamel CP, Lenaers G. Recessive mutations in *RTN4IP1* cause isolated and syndromic optic neuropathy. American Journal of Human Genetics. 2015;97(5):754-60.

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**Retina**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Inherited retinal dystrophies are the most common cause of blindness in working aged adults (Liew G, et al BMJ Open. 2014 Feb 12;4(2)). Precise diagnosis is confounded by the vast genetic and allelic heterogeneity of the disorder, with over 170 implicated genes. However, unlike most Mendelian disorders, retinal dystrophies are tractable to novel treatments such as gene-replacement and cell-therapy. The UK has been at the vanguard of the emerging technology being involved in two seminal gene-replacement trials (Bainbridge N Engl J Med. 2008 May 22;358(21):2231-9, MacLaren R et al Lancet. 2014 Mar 29;383(9923):1129-37) and further trials are planned. In order to realise and harness these opportunities, molecularly typed cohorts of patients need to be determined efficiently. Moreover, natural history data and precise phenotypic measures are required on such cohorts to inform the design and interpretation of such trials.

As well as its tractablity to treatment, the eye and retina are uniquely accessible to investigation, both by psychophysical testing, electrophysiology and imaging. Coupling underlying gene with the results of such testing in patients gives insights into important areas of retinal biology. For instance, RPE and photoreceptor maintenance is perturbed by mutation in a growing suite of genes, some retinal specifc and others ubiquitously expressed. Secondly, key components of visual phototransduction are effected by human knockouts identified through studies such as this, and the precise effect on retina function and structure can be determined. A further suite of genes effects signal processing through bipolar cells, and another the regeneration of visual pigment. The investigation of such patients informs the basic biology of these processes more accurately than can be achieved through animal models or cell culture.

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

1) Collating phenotypic data and investigating with further tests those patients and families with disease due to known genes. This will allow us to understand each molecular subtype of inherited retinal disease, to establish age-related severity, degree of variability, phenotype-genotype correlations. At this stage the GeCIP can reach out to the collaborators below to strengthen the data from each molecular subgroup. Phenotype-genotype correlations remain important and continue to throw up unexpected insights. A recent example is the finding of RP and cone-rod dystrophy in adults with specific alleles of CLN3 and CLN7 respectively. Most mutations cause a severe, childhood onset neurological failure (Battens disease) yet these patients are otherwise well. At least in the case of CLN3 the retinal phenotype is distinct from that seen in the syndromic disorder. This requires more study to understand the mechanism and it is likely that other unexpected correlations emerge from the data from GE.

2) Finding new genes. There will be a cohort of patients and families without tier 1 or 2 variants in whom novel genes are suspected. HPO terms and imaging data will be used to group patients into similar phenotypes to improve the efficiency of novel gene detection

3) Modifier alleles. Many examples exist in which affected member so the same family are vastly different in their age-of-onset and severity. An extreme example is dominant retinitis pigmentosa dues to splicing factor genes (particularly RP11 and RP13 caused by heterozygous mutation of PRPF31 and PRPF8 respectively). GE data gives an unprecedented opportunity to identify those modifiers that are genetic. Hypotheses to be explored involved variants in the non-mutated allele in dominant disease, and SNPs in those genes known to be binding partners or in the same biological pathway as the primary gene. One hypothesis is that there are modifiers that affect rod-cone dystrophy (that is RP) off many different molecular causes and this can be investigated robustly by pooling all such cases together. To confound this analysis of course, the number of individual comparisons incurred by searching the genome for potential modifiers would be intractable, but with focus on those loci with a high prior probability this, hopefully, can be successfully managed.

4) HPO terms and their predictive power for specific genotypes. As clinical and genotypic data is accumulated, differences in clinical presentation will emerge between different subtypes of inherited retinal disease. The emerging data and that from previous studies (> 600 IRD cases from BRIDGE-NIHR for instance) can allow the determination of the likelihood of a specifc HPO term given a known diagnosis. Using Bayes, this allows estimates of the reverse probabilities, that is the likelihood of an underlying gene given a specific HPO term. These probabilities will be powerful in determining HPO terms that are relevant in making a molecular diagnosis and for prioritising the many candidate variants that can occur in single families. The analysis will use age-related terms, and this will complicate the analysis, but we anticipate make the results more powerful.

The GeCIP is well represented in the application of cell biological assays to test the effects of coding and non-coding variation. For example, heterologous expression of tagged proteins and variants can be used to determine protein folding and stability and effects on sub-cellular localization and protein:protein interactions (e.g Davidson AE et al Am. J. Hum. Genet. 2013 93(2): 321–329). The localization of novel candidate disease proteins can be ascertained in the retina using immunohistochemistry (Evans, Rj et al Hum. Mol. Genet 2010. 19(7):1358-67). The effects of variants on mRNA processing (i.e. alternative splicing) and stability will be probed using RT-PCR on blood RNA or using mini-genes or patient derived cells.
iPSC can be differentiated to RPE and/or photoreceptors to study potential disease mechanisms and test potential therapies. For example, Mike Cheetham and colleagues have recently differentiated cells from an X-linked RP patient with a premature stop codon to RPE and observed trafficking defects that can be rescued by using translational readthrough drugs (Schwarz N et al Hum Mol Genet. 2015 24(4):972-86).

Finally, Clustered regularly-interspaced short palindromic repeats (CRISPR)/Cas9 technology for genome editing and/or morpholino antisense oligonucleotide knockdown can be performed to examine the phenotypes related to specific mutations in the zebrafish model (Shu X et al. Invest Ophthalmol Vis Sci. 2011 5;52(6):2960-6).

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

The Eye genetics group and inherited retinal disease consortium provide an available working network for the dissemination of results to the national inherited eye disease community. We intend to collaborate on cases in which syndromic disease involves the retina (eg GeCIPs of paediatrics, nephrology, neurology). Much data has been generated already as part of NIHR-

BRIDGE SPEED project (Raymond as PI) which will be immediately available to add to the emerging GE data.

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*
Mariya Moosajee is an academic clinical lecturer and Panos Sergouniotis an academic clinical fellow who will oversee the involvement of trainees in the GE projec t generally. Both have an interest in inherited eye disease, with high impact publications. A fellowship in Genetic Eye disease is available in Manchester and the Special Trustees of Moorfields have invited an application to provide seed-funding for a dedicated fellow in genomic ophthalmology based at MEH/IOO.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

Investigation of the perturbation of retinal disease genes is well advanced in both iPSCs and animal models such as the zebrafish. Professor Mike Cheetham in collaboration with Professor Coffey has led the development of an iPSC bank of 20 different retinal disease gene alleles and we will develop more cell lines to complement these functional investigations exploiting our ability to reprogram uroepitelial cells, which is non-invasive and applicable to children. Mariya Moosajee leads a zebrafish group devoted to eye disorders and works closely with Professor Steve Wilson at UCL. Black, Ramsden, Webster, Arno have unrivalled experience in variant interpretation in the context of inherited retinal disease.

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

Retinal disease is well advance in the interpretation of genetic results (see above). The cited networks should also ensure that specific families have access to relevant gene-related observational or intervention trials whereever from England they were recruited.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

Given the advances in variant interpretation and the onset of gene-directed treatments we envisage the GE resource to be a pipeline from discovery to treatment trials for the benefit of those affected with blinding disorders.

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

References. *Provide key references related to the research you set out.*

See inline.

Detailed research plan

Full proposal (total max 1500 words per subdomain)

Title *(max 150 characters)*	**Developmental disorders of the eye and ear (DDEE)**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Congenital eye defects causing childhood visual impairment affect around 5 per 10,000 children. Malformations affect the anterior segment (e.g. corneal opacity/angle abnormalities/congenital glaucoma), the posterior segment (e.g. congenital cataract, chorioretinal coloboma) and/or cause whole globe abnormalities (e.g. anophthalmia, coloboma and microphthalmia). These developmental disorders often present as a combination of clinical features. Syndromic cases with additional non-ophthalmic disorders are prevalent (>75% of cases are syndromic; 10% of children die within year of diagnosis). Congenital deafness and permanent childhood hearing loss affects around 2 in 1000 children (~ 30% of cases of congenital or early onset hearing loss is syndromic - part of a rare recognisable syndrome). Deafblind conditions represent about 5% of all children who are profoundly deaf.

The majority of cases of developmental disorders causing childhood blindness/deafness are without a genetic diagnosis, and most cases are managed without knowledge of the cause of the condition and underlying molecular pathology. Syndromes may, in addition to eyes and ears, affect any body system including the nervous, cardiovascular, renal, integumentary. These conditions need specialist paediatric cross disciplinary management and this needs to be supported by a coordinated and systematic approach for integration of genomic and clinical data, including knowledge of associated risks. For example, early genetic diagnosis of metabolic cataract or SOX2 gene-associated endocrine abnormalities can prevent longer term morbidity.

There is a need for research to complete the discovery and functional characterisation of all genes within the genome that contribute to the aetiology of childhood developmental disorders of the eye and ear. Knowledge of these genes is likely to contribute to the interpretation of genomic data for late onset sensory system disorders, as well as improving management, and opening new avenues for improved care and therapy development for these rare conditions. Genetic analysis of these conditions lags considerably behind analysis of inherited retinal disease, though indications are of high genetic and allelic heterogeneity.

The ability to study trios and determine de novo mutations in affected probands, and discover new genes will be a key opportunity within the Genomics England project (this is illustrated by the success of the similar approach, with relatively limited genetic data within the DDD project http://www.ddduk.org/). The DDEE subdomain partners (Sowden, Bitner-Glindzicz, Black, Lloyd, Fitzpatrick and Ragge and other GeCIP members) have made significant contributions to the genetic analysis of developmental eye and ear disorders and are all currently engaged in research in this area. For example, Sowden and Bitner-Glindzicz are chief investigators on multicentre studies recruiting children with sight and hearing impairment for genetic analysis (UKCRN IDs 11800 & 12004).

(425 words)

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

GE provides a significant opportunity to advance and apply understanding of the molecular pathology of developmental ocular and auditory disorders. Proposed research studies in paediatric genetic hearing loss and blindness include the following:

- Identification of novel genes for non-syndromic and syndromic forms of congenital deafness and blindness and development of new diagnostic test batteries and care pathways
- Creation of cell models of developmental disorders for therapeutic manipulation (iPSC models)
- Establishment of national longitudinal studies of developmental disorders and outcomes linked to genotype (patient stratification for future trials, for example for congenital glaucoma and cataract) and creation of international collaborative networks.
- Creation and study of novel mouse models of human disease genes to understand their role in hearing and vision; use of zebrafish model to assay variant protein function.
- Development of regenerative medicine therapies for treatment of ear and eye malformations and approaches using TRID, antisense knock down and gene editing.
- Development of diagnostic and prognostic tests combining genomics with biomarker/OMICs analysis

Work streams will be developed to analyse variants of uncertain significance (Tier 3), and for novel disease genes, for example using the UK Gene Function Pipeline. The DDEE subdomain partners (Sowden, Bitner-Glindzicz, Black, Lloyd, Fitzpatrick and Ragge and other GeCIP members) and their UK-wide and international collaborators provide a comprehensive range of expertise for the proposed research. The major challenges anticipated are (i) the small numbers of patients affected by each rare condition; this will be addressed by developing strategies for national and international data sharing. (ii) The challenge of identifying non-coding variants of relevance, and of assessing whether genes that segregate in a rare disease population are significant within the wider GE population. We will engage with cross cutting themes for advanced analytics and with our local Genomics expertise. The bioinformatics teams from UCL Genetics Institute, GOSgene, Birmingham and Manchester are vital for the proposed interpretation, validation and reporting of genomics results. GOSgene, the West Midlands Regional Genetics laboratories and the Manchester laboratories have experience of NGS data analysis and novel gene discovery using bespoke pipelines.
(338 words)

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

The Developmental Disorders of the Eyes and Ears Subdomain, will cross link and interact closely with Paediatrics led by Professors Tim Barrett/Phil Beales and with Endocrine and Metabolism led by Prof Stephen O' Rahillly; Professor Bitner Glindzicz is NHS GMC representative for Paediatrics. In particular this subdomain will have close links with the Pediatric Developmental Disorders Subdomain led by Matt Hurles with cross representation by Professor David Fitzpatrick. MBG is lead for Rare Disease.

Subdomain members are part of the Eye Genetics Group (UKEGG; http://www.ukegg.com) set up in 2001 to enable specialists in the field to meet regularly and exchange experience and share expertise regarding genetic ocular conditions. The steering group and membership are involved in disseminating information about ongoing research in the field, guidelines for clinics and is taking a role in promoting training and development as well as research and education in the speciality of ocular genetics. Its inclusive membership represents a broad base for this GeCIP. In addition we plan to have a twice yearly meeting with all members of the collaborative network to discuss new, collaborative, multidisciplinary projects arising from this GeCIP sub-domain.

We have partnerships with patients groups, such as Genetic Disorders UK and the Micro and Anophthalmia Children's Society (MACS).
(205 words)

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*

Great Ormond Street Hospital and other national pediatric centres in Manchester, Leeds, and Birmingham see an unrivalled mix of children who have both ocular/ auricular and systemic

disease. These children often have many other significant medical illnesses. A key objective is to improve the diagnosis and management of these of these rare conditions. Trainees will be involved in the research at each site and we will seek support from NIHR research funding streams for clinical and non clinical and allied health research training.
(82 words)

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

The subdomain comprises a group of investigators with an excellent track record of working in developmental eye disorders, in laboratory based experimental medicine and clinical studies. In addition to the planned submission of applications to individual funders for separate components of the outlined work we will seek support from the Eye Theme, part of the NIHR Rare Diseases Translational Research Collaboration led by Webster and Black to establish a national registry for patients with these Rare Disease that can be integrated with Genomics England data and with existing genomic analyses from national initiatives. For example, Fitzpatrick has analysed a cohort of coloboma patients as part of the UK10K project, Sowden, Bitner-Glindzicz and Black have developed NGS panels screening for developmental eye disorders within the Regional Genetics laboratories and other collaborators have undertaken exome analysis within local facilities.
(138 words)

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

We expect this research to provide patient benefit by improving the level of genetic diagnosis provided nationally for developmental eye and ear conditions. We expect to discover novel disease genes and syndromes, and to provide sufficient evidence to support re-assignment of “amber” and “red” genes to become actionable “green” genes for clinical genetic diagnosis.
(54 words)

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

In addition to the aim of improving pathways for clinical management, the knowledge generated will be of high importance for the wide range of investigators working on sensory system development in model organisms; we will continue to expand collaborations using animal models to validate human functional variants and using existing animal models to guide interpretation of newly discovered disease genes, for example with PhD student training within the cross disciplinary Sensory Systems Technologies and Therapies programme at UCL (SenSyT).

(79 words)

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

We have strong links with potential industrial partners via UCL Business and the Translational Research Office and as part of NIHR Biomedical Research Centre activities. Specific areas for development will be discussed as appropriate and as national cohorts and registries are established for trial readiness, particularly in the area of orphan drug indications.
(53 words)

References. *Provide key references related to the research you set out.*

MiR-204 is responsible for inherited retinal dystrophy associated with ocular coloboma.
Conte I, Hadfield KD, Barbato S, Carrella S, Pizzo M, Bhat RS, Carissimo A, Karali M, Porter LF, Urquhart J, Hateley S, O'Sullivan J, Manson FD, Neuhauss SC, Banfi S, **Black GC**.

Proc Natl Acad Sci U S A. 2015 Jun 23;112(25):E3236-45.

Heterozygous loss-of-function mutations in YAP1 cause both isolated and syndromic optic fissure closure defects.
Williamson KA, Rainger J, Floyd JA, Ansari M, Meynert A, Aldridge KV, Rainger JK, Anderson CA, Moore AT, Hurles ME, Clarke A, van Heyningen V, Verloes A, Taylor MS, Wilkie AO; UK10K Consortium, **Fitzpatrick DR**.
Am J Hum Genet. 2014 Feb 6;94(2):295-302.

Next-generation Sequencing in the Diagnosis of Metabolic Disease Marked by Pediatric Cataract.
Gillespie RL, Urquhart J, Anderson B, Williams S, Waller S, Ashworth J, Biswas S, Jones S, Stewart F, **Lloyd IC**, Clayton-Smith J, **Black GC**.
Ophthalmology. 2016 Jan;123(1):217-20.

Mutation of SALL2 causes recessive ocular coloboma in humans and mice.
Kelberman D, Islam L, Lakowski J, Bacchelli C, Chanudet E, Lescai F, Patel A, Stupka E, Buck A, Wolf S, Beales PL, Jacques TS, **Bitner-Glindzicz M**, Liasis A, Lehmann OJ, Kohlhase J, Nischal KK, **Sowden JC**. ***Hum Mol Genet***. 2014 May 15;23(10):2511-26.

Mitochondrial m.1584A 12S m62A rRNA methylation in families with m.1555A>G associated hearing loss. O'Sullivan M, Rutland P, Lucas D, Ashton E, Hendricks S, Rahman S, **Bitner-Glindzicz M**. ***Hum Mol Genet***. 2015 Feb 15;24(4):1036-44.

Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, King DA, Ambridge K, Barrett DM, Bayzetinova T, Bevan AP, Bragin E, Chatzimichali EA, Gribble S, Jones P, Krishnappa N, Mason LE, Miller R, Morley KI, Parthiban V, Prigmore E, Rajan D, Sifrim A, Swaminathan GJ, Tivey AR, Middleton A, Parker M, Carter NP, Barrett JC, Hurles ME, **FitzPatrick DR**, Firth HV; DDD study. ***Lancet***. 2015 Apr 4;385(9975):1305-14.

ALDH1A3 mutations cause recessive anophthalmia and microphthalmia.
Fares-Taie L, Gerber S, Chassaing N, Clayton-Smith J, Hanein S, Silva E, Serey M, Serre V, Gérard X, Baumann C, Plessis G, Demeer B, Brétillon L, Bole C, Nitschke P, Munnich A, Lyonnet S, Calvas P, Kaplan J, **Ragge N**, Rozet JM. ***Am J Hum Genet***. 2013 Feb 7;92(2):265-70.

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**Disorders of Hearing and the Ear**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Data from Newborn Hearing Screening has shown that **congenital deafness** affects more than 1/1000 children with a similar number losing their hearing before maturity. In industrialised countries >80% are thought to have a genetic cause. In 70% of cases deafness is the only phenotype ; 30% is syndromic. Thus deafness is incredibly heterogenous, with over around 170 genes mapped. More than half of these loci remain unidentified meaning current gene screening is informative in less than half of cases.[1]

In adults **progressive sensorineural hearing loss** is the most common sensory deficit, affecting more than half of the over 70s, and is associated with depression and cognitive decline. Hearing aids are the only current treatment but do not restore normal frequency selectivity, temporal processing, or ability to follow a single speaker in a noisy background. Mutation detection for diagnostic purposes is low in such cases underlining the unmet clinical need.

Familial otosclerosis is an early adult onset monogenic hearing loss consistent with autosomal dominant inheritance which exhibits variable penetrance and affects 1/1000 individuals in the UK. So far linkage analysis and sequencing has failed to identify causal variants. Consequently, the molecular basis of the disorder remains unclear although it is likely to be heterogeneous[2, 3]. Through the recruitment of familial cases to the Programme (proposal imminent), we will identify the causal genes, thereby aiding early diagnosis, stratifying patients according to cause and outcomes, and in due course inform therapeutic approaches to treat or prevent the bone dys-regulation that causes the hearing loss.

Over 800 children across Europe are born every year with severe ear abnormalities **(anotia/microtia).** These malformations have a significant functional and psychological impact on quality of life and an onerous long-term socio-economical impact. Although genes for syndromic forms of microtia have been identified (*SIX1, EYA1, HOXA1, TCOF1* etc) very few for non-syndromic microtia have been identified to date. If identified they would reveal the basis of ear abnormalities and whether they are intrinsic to chondrogenic precursors (e.g. proliferation/migration/differentiation ability) or extrinsic (e.g. patterning cues/vascularization defects/extracellular matrix defects), and how this may affect the behavior of patient-derived stem cells with chondrogenic potential. Development of autologous tissue-engineered cartilage to treat microtia/anotia in the Ferretti lab aims reduce number of operations and morbidity, helping to define optimal parameters for successful engineered cartilage for reconstruction[4].

There is a huge knowledge gap in terms of personalised genetic counselling for families, stratification of patients and families for long term natural history and outcome studies. For example, defining appropriate candidates for best cochlear implantation outcome, or surgical intervention otosclerosis, by genetic profiling. Knowledge of biological pathways for intervention and treatment for progressive hearing loss is in its infancy.

With access to 100,000 Genomes datasets we will be able to address a number of outstanding research areas (see research plans below):

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

1. **Discovery of novel genes in Hearing and Ear disorders and their characterization in animal models.**

Data from patients without mutations in known genes will be analysed using expanded 'virtual gene panels' based on our combined knowledge of inner ear biology, ear development, known mouse mutants and comparison with the Morton human cochlear cDNA library.
Our GeCIP has access to new candidate genes identified through large-scale mouse mutant screens Several mutants have recessive congenital deafness, with mutations affecting inner hair cells and their afferent innervation (similar to human auditory neuropathy). Mice with dominant inheritance of progressive hearing loss may represent orthologues of genes involved in deafness in relevant patients groups in the 100,000 Genomes Project. These genes will all be part of the 'Virtual panel'.[5]

Mouse models will be sought for novel deafness genes [2]; over 13,000 mouse genes targeted in ES cells are available for selection and microinjection to generate new mouse lines, and a growing number through co-ordinated public archives, as well as a large database of predicted pathogenic mutations archived as frozen sperm from exome-sequenced, ENU-mutagenised male mice. These include single base changes including missense, stop lost/gained and splice site mutations.
Those genes considered strong candidates and for which there is no animal model will be submitted for consideration by the proposed UK Gene Function Pipeline or to the Genomics England/Harwell collaboration for the generation of mouse models.

2. **Familial otosclerosis** (proposed for inclusion in the Programme).

The 100,000 Genomes data will be compared to Dr Dawson's WES data from 24 familial cases and findings validated in a larger cohort of otosclerosis patients (n=620 of familial and sporadic cases). Similar approaches to those above will be used to find and investigate animal models of new genes.

3. **WGS combined with Newborn Hearing Screening** and comparison with SEQaBOO Boston.

The SEQaBOO project (SEQuencing a Baby for an Optimal Outcome), involves rapid exome sequencing within 6 weeks in babies who fail Newborn Hearing Screening at Brigham and Women's Hospital, Boston. Aims are to determine whether this approach aids clinicians and families in reducing additional investigations and in decisions on auditory management. Access to the 100,000 Genomes dataset will allow us to determine
i) what proportion of those with congenital profound deafness have reportable findings
ii) whether WGS increases mutation detection rate compared with WES in SEQaBOO Boston
iii) whether a similar study would be feasible in England where cost-benefit analysis would differ.
The first two questions could be rapidly answered using 'diagnostically reportable' results found by this GeCIP.

4. **Stratification by genotype and response to Cochlear Implant (CI).**

Outcomes such as speech production, language and speech perception vary considerably between individuals. This may be due to differences in the number of surviving spiral ganglion neurons, aetiology of hearing loss (whether the lesion is pre or post synaptic), or other factors such as age of implantation. Genetic heterogeneity of patients is a confounding factor when assessing other variables related to outcome. Genetic stratification correlated with CI outcomes, through international collaborations, will tease out these complex variables. For example

ANSD caused by *OTOF* mutations, (synaptic), generally has a good outcome in contrast to other types of ANSD. Identification of novel genes, their expression pattern, and functional studies of animal models will enable us to determine location of the lesion and correlation of this with outcomes[6, 7].

Challenges to these questions will be the genetic heterogeneity of deafness in relation to numbers of subjects recruited. We will mitigate against these problems by domain members engaging with all local GMCs/clinicians to ensure maximal recruitment. Each GeCIP member will give presentations on the Programme to their local Cochlear Implant/Audiology teams to raise awareness and stimulate recruitment.

5. Overall Mutational Burden and genotype-phenotype studies
We will investigate whether overall mutational burden is greater in patients than in the rest of the cohort ie. oligogenic or digenic causes of hearing loss? Investigations of whether mutational burden correlates with increased severity.

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

PIs have extensive international collaborations.
GeCIPs:

- Paediatric (Beales /Barrett) - children with unique disorders including deafness
- RNA/transcriptomics subdomain (Baralle) of Machine Learning.., (verification of splicing)
- Validation and Feedback (Newman)

We have engaged
NHS clinicians working in paediatric and adult audiology and Cochlear Implant teams
Action on Hearing Loss, Sense and Genetic Disorders UK.

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*

Trainees in Audiovestibular medicine, ENT and Genetics will be engaged in the Project including subject ascertainment, recruitment, the research outlined above, and feedback. We will provide opportunities for clinical and basic scientists to train in genomic medicine (BSc, MSc and PhD projects). We will seek support from NIHR for clinical and non-clinical and allied health research training.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*
This domain has extensive expertise in basic science and clinical studies. We have a track record in deafness gene identification, functional studies and population approaches. Resources across the domain include electrophysiology, structural/ultrastructural techniques, expression analysis including RNAseq, ChIPSEq, genomic interpretation, bioinformatics and developmental studies. Ferretti's expertise is in tissue repair and regeneration, stem cell biology, normal and abnormal skeletogenic development.

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

Members of our GeCIP have wide experience in functional verification of variants through molecular/cell biology, animal modelling, or splicing effects and in diagnostic interpretation by Clinical and Basic scientists. We specialise in translational genomics (Le Quesne Stabej) and UKAS/CLIA diagnostic laboratory services (Cullup; Jenkins, Morton – clinical laboratory directors). Bitner-Glindzicz is part of the V&F GeCIP.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

Patients and the NHS will benefit from more comprehensive and rapid testing, accurate genetic counselling, and better information about genotype/phenotype correlations and long-term outcomes. Treatment will become personalised.
Researchers will benefit from better biological knowledge of pathways which may enable druggable targets and collaborations with **industry**.

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

We have links with GSK, Johnson and Johnson, small and large biotech companies, and cochlear implant providers. Although no commercial partners are in place, it is clear that pharmaceutical companies have great interest in both congenital and adult onset hearing loss.

References. *Provide key references related to the research you set out.*
1. Morton CC, Nance WE. Newborn hearing screening--a silent revolution. N Engl J Med 2006;354:2151-64.

2. Sommen M, Van Camp G, Liktor B, Csomor P, Fransen E, Sziklai I, Schrauwen I, Karosi T. Genetic association analysis in a clinically and histologically confirmed otosclerosis population confirms association with the TGFB1 gene but suggests an association of the RELN gene with a clinically indistinguishable otosclerosis-like phenotype. Otol Neurotol. 2014 Jul;35(6):1058-64.

3.Priyadarshi S, Ray CS, Biswal NC, Nayak SR, Panda KC, Desai A, Ramchander PV. Genetic association and altered gene expression of osteoprotegerin in otosclerosis patients. Ann Hum Genet. 2015 Jul;79(4):225-37.

4. Guasti L, Vagaska B, Bulstrode NW, Seifalian AM, Ferretti P. Chondrogenic differentiation of adipose tissue-derived stem cells within nanocaged POSS-PCU scaffolds: a new tool for nanomedicine. Nanomedicine. 2014 Feb;10(2):279-89.

5. Chen J, Ingham N, Kelly J, Jadeja S, Goulding D, Pass J, Mahajan VB, Tsang SH, Nijnik A, Jackson IJ, White JK, Forge A, Jagger D, Steel KP (2014) Spinster homolog 2 (*Spns2*) deficiency causes early onset progressive hearing loss. PLoS Genetics 10:e1004688.

6. Santarelli R, Rossi R, Scimemi P, Cama E, Valentino ML, La Morgia C, Caporali L, Liguori R, Magnavita V, Monteleone A, Biscaro A, Arslan E, Carelli V. OPA1-related auditory neuropathy: site of lesion and outcome of cochlear implantation. Brain. 2015 Mar;138(Pt 3):563-76.

7. Santarelli R. Information from cochlear potentials and genetic mutations helps localize the lesion site in auditory neuropathy. Genome Med. 2010 Dec 22;2(12):91.

Eppsteiner RW, Shearer AE, Hildebrand MS, et al. Using the phenome and genome to improve genetic diagnosis for deafness. Otolaryngol Head Neck Surg 2012;147:975-7.

8. Eppsteiner RW, Shearer AE, Hildebrand MS, et al. Using the phenome and genome to improve genetic diagnosis for deafness. Otolaryngol Head Neck Surg 2012;147:975-7.