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	**Skin**
Project title *(max 150 characters)*	**Genetic causes of rare skin disease**

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

1. *Optimise clinical data and sample collection, clinical reporting and data interpretation for return to clinicians and improving diagnostic for patients.*
2. *Perform research that improves our understanding of the implications of the findings in a clinical setting, for example improved phenotype-genotype correlation.*
3. *Discovery of new genes/mutations for skin diseases.*
4. *Discovery of new gene function through a cell biology approach.*
5. *Stratification of disease for new therapies.*
6. *Develop a cohort of clinical training fellows working on rare skin disease-related genomics projects.*

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)*

Skin, as a large and clearly visible organ, is very important in life. Genetic variation, which is inherited from our parents, contributes to the colour and texture of our skin. In some families, rare genetic changes can be inherited, causing problems with the skin which range from dry scaly skin (called ichthyosis) to thickened skin (which we call keratoderma) and inflammation (red, sore or itchy skin) or fragile skin which leads to blistering. The hair, nails are teeth are produced in the same way as skin, so they can also be affected. These diseases are very distressing and can also be life-threatening.

Our group represents a collection of skin specialists (dermatologists) and geneticists who are interested in rare skin diseases. Many of the inherited skin diseases have no effective treatments and no cure. If we can understand the genes which cause these diseases we would be able to explain the diseases to our patients more clearly, for example to advise whether other people in their family could be affected. We also hope that understanding genetic mechanisms – how genes work to cause disease in the skin – will pave the way to developing new and better treatments.

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)*

Rare variants in genes known to be associated with skin disease will be identified by: filtering variants identified in these candidate genes against frequency in the general population; co-segregation analyses (including de-novo variant detection); comparison with databases of genetic variants found in people with the disease and prediction of effect of each variant on protein structure and function. Variants will be classified by likely pathogenicity and either fed back to participants if they can be reliably inferred to be disease-causing or subjected to further analyses, such as comparison of frequency in patients and controls. Investigation of the effects of gene changes on protein function will be performed in laboratories where particular genes/pathways are studied.

Whole genome sequencing of patients with rare genetic skin disorders represent an opportunity for discovering new mechanisms that cause disease in the skin. By coupling deep phenotyping of carefully selected pedigrees with whole genome sequencing at a national scale, we are poised to discover novel pathogenic variants that will inform patient management. A further strength is the ability to investigate putative functional consequences in target tissue, namely skin biopsies from affected patients. Molecular academic dermatologists have an established track record of taking such findings forward to understand disease pathogenesis and inform genetic counselling of patients.

We plan to collect clinical data and samples from well characterised pedigrees with a range of skin disorders (currently 14 conditions). These families have been carefully selected by academic dermatologists that have compiled clinical and molecular diagnostic information but have not been able to demonstrate a genetic cause using existing technologies. Using multiple unrelated individuals will now optimise the ability to discover recurrent pathogenic changes. We will prioritise research strategies on these conditions based on a range of factors, including the number of patients that are enrolled, the impact of the disease, the ability to model and study the disease in collaboration with academic partners, and finally the likelihood of using repurposed therapies based on existing understanding of gene function.

This has immediate implications for patient care. Firstly, patients and families can receive genetic counselling that can influence decisions such as preimplantation genetic diagnosis. In dermatology, there are precedents where patients may also benefit from personalised treatments. In patients with autoinflammatory syndromes for example, the discovery that genetic mutations resulted in abnormal regulation of interleukin signalling led to trials of IL-1 antagonists with dramatic response in the skin (1). Similarly, patients with a rare form of inflammatory skin disease (generalised pustular psoriasis-IL36RN) that is associated with recurrent fevers and pustulation warranting hospital admission, have their health transformed by the administration of anti-TNF therapies (2).

To make the transition from discovery to therapeutic, a programme of research on these findings is planned. Research on newly discovered genes from the current list of genetic skin disorders will inform fundamental processes in the skin from response to UV induced DNA damage, barrier function and structural integrity to cutaneous immunity and inflammation. Novel gene discoveries will be explored through a cell biology approach, and will form part of a training programme to develop a new cohort of clinical training fellows in dermatology. Transitioning the process of identifying and characterising pathogenic changes in the skin from academia to the clinic will need to be underpinned by the development of trainees in this skillset, and will form part of the legacy of the new genomic NHS.

Expected start date	**Autumn 2018**
Expected end date	**Autumn 2020 (analysis), cell biology is likely to be ongoing**

Lead Applicant(s)	
Name	John McGrath (1), Edel O'Toole (2)and Neil Rajan (3)

Post	Professor of Molecular Dermatology and Head of the Genetic Skin Disease Group (1), Centre Lead and Professor of Molecular Dermatology (2), Clinical Senior Lecturer (3)
Department	Genetics and Molecular Medicine (1). Blizard Institute (2), Institute of Genetic Medicine (3)
Institution	Kings College London (1), Queen Mary University of London (2), Newcastle University (3)
Current commercial links	

Administrative Support	
Name	
Email	Leadership team: john.mcgrath@kcl.ac.uk; neil.rajan@ncl.ac.uk; e.a.otoole@qmul.ac.uk
Telephone	

Subdomain leads		
Name	**Subdomain**	**Institution**
E O'Toole, D Kelsell, S Brown, M Simpson	Atopy	QMUL, Dundee, KCL
V Kinsler	Neurocutaneous	UCL
F Capon/M Simpson	Generalised pustular psoriasis	KCL
E O'Toole/D Kelsell/N Rajan	Keratodermas/erythrokeratodermas	QMUL/ Newcastle
J McGrath	Epidermolysis bullosa	KCL
J McGrath	Skin adnexal disorders	KCL
J McGrath/E O'Toole/D Kelsell	ARCI	QMUL/KCL

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**Genetic causes of rare skin disease**

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).*

Rare genetic skin diseases mainly affecting the skin and its appendages but can also be a manifestation of multi-system disease. They frequently occur at birth or early in life, are generally chronic, often severe and may be life-threatening. They can be difficult to diagnose, as healthcare professionals other than specialist dermatologists may be not aware of their clinical presentation and UKGTN accredited diagnostic testing is not available for many skin diseases. Furthermore, the current absence of curative therapies poses significant problems in the clinical management of patients, which frequently requires a costly and time consuming multidisciplinary approach. Finally, the quality of life of both patients and their families may be severely affected by the psycho-social and physical impact of severe skin disease.

Understanding the genetic basis of disease can lead to improved care and drive momentum in translational research. For example, CHILD syndrome (Congenital Hemidysplasia with Ichthyosiform erythroderma and Limb Defects) is a rare X-linked dominant ichthyotic disorder caused by loss of function mutations in the *NSDHL* gene, which leads to inhibition of cholesterol synthesis and accumulation of toxic metabolic intermediates in affected tissues. Topical treatment with lovastatin (inhibiting accumulation of toxic intermediates) and cholesterol (restoring deficiency) restores the skin appearance to normal (3). The finding that filaggrin mutations are the most frequent genetic predisposing factor to atopic eczema (4) and associated allergies has kick-started major academic and industrial research to increase filaggrin expression and improve the skin barrier. Mutations in the type VII collagen gene (*COL7A1*) cause severe generalised recessive dystrophic epidermolysis bullosa. A major worldwide effort is now directed at clinical trials to restore type VII collagen in these patients including gene, cell and protein therapies (5).

We expect that the research outlined will improve clinical care by giving patients with rare skin diseases a precise diagnosis and offering the opportunity for clinical trials. Deeper phenotyping of groups of patients and genotype-phenotype correlation should allow stratification of subgroups of patients for therapy, for example, severe atopic patients with high IgE levels or patients with autosomal recessive congenital ichthyosis.

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.*

Each disease category GeCIP lead/deputy will be responsible for generating research questions that address clinical issues. These questions are expected to cover (a) discovery of new genes/mutations for rare skin diseases; (b) defining or refining genotype-phenotype correlation; (c) stratification for therapy (present/future); (d) feasibility of prenatal testing (if clinically indicated); (e) information for patients relating to the latest information about their condition, current best management, active clinical trials, and access to disease updates. We also expect some of the research questions to involve integration of other – omics data, with development of new functional assays for informative non-coding region variants linked to phenotype or other clinical metrics. Within the skin domain, there are several research groups that are active in rare genetic skin disease research including King's College London, Queen Mary University of London, Institute of Child Health, University College London, University of Dundee and Newcastle University. Each phenotypic subgroup will be linked to a research group or groups, as required. Each research group has linked bioinformatics expertise including Vincent Plagnol at University College London, Michael Simpson at King's College London, Mauro Santibanez-Koref at Newcastle, Michael Barnes at Queen Mary University of London etc.

Discovery of new skin disease genes: The GeCIP is already endeavouring to encourage recruitment of patients with diseases where we know the genetic basis is poorly understood. For example, just 40% of patients with erythrokeratodermas have pathogenic mutations in *GJA1*, *GJB3* or *GJB4*. Grouping together increased numbers of trios should allow identification of novel gene variants associated with disease pathogenesis. Finding new genes will also present the opportunity for functional studies including in vitro siRNA, CRISPR-CAS9, disease modelling using organotypic cultures and in vivo work to understand disease pathogenesis.

Defining or refining genotype-phenotype correlation: The identification of new genes or novel variants in known genes in for example the palmoplantar keratodermas or autosomal recessive congenital ichthyoses will allow better subclassification of subgroups of these categories of

diseases by both expert deep phenotyping and may also allow opportunities for stratification of patients for possible clinical trials.

The timelines for this research will vary depending on factors including access to patient material, existing research group expertise and momentum in the novel candidate gene and technical factors such as the availability of molecular tools such as antibodies and existing transgenic models. A typical example may take 12-18 months from identification of a novel variant to establishing it's role in disease pathogenesis with experimental laboratory data. Publication and dissemination of these discoveries may take a further 6 months. As highlighted earlier, the skin GeCip will prioritise projects that can be rapidly progressed, such that deliverables will appear in the time frame of this project.

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 Skin GeCIP propose to build and maintain close links with related GeCIPs where skin input may be needed eg Immune Disorders, Inherited cancer predisposition, Paediatrics, Musculoskeletal, Validation and Feedback and Stratified health care and therapeutic innovations. International collaborators that provided letters of support for the original Skin GeCIP application included the following: Prof Alain Hovnanian, Inserm, Paris, France (inherited disorders of keratinisation); Prof Alan Irvine, Trinity College Dublin, Ireland (ichthyosis, keratoderma); Prof Julia Lee, National Cheng Kung University Medical College, Tainan, Taiwan (inherited pigmentary and scarring diseases); Dr Arti Nanda, As'ad Al-Hamad Dermatology Center, Kuwait (ectodermal dysplasias, developmental disorders); Dr Julio Salas-Alanis, University of Monterrey, Mexico (inherited blistering diseases, hair disorders); Prof Hiroshi Shimizu, Hokkaido University, Japan (blistering diseases, ecotdermal dysplasias); Prof Rodney Sinclair, University of Melbourne, Melbourne, Australia (inherited hair diseases); Dr Andrew South, Jefferson Medical College, Philadelphia, USA (inherited blistering diseases); Prof Eli Sprecher, University of Tel Aviv, Israel (inherited disorders of keratinisation, keratodermas); Dr Antonio Torrelo, Hospital Infantil Universitario Nino Jesus, Madrid, Spain (inherited vascular and keratinizing disorders in children).

Other possible academic collaborations include collaboration with PSORT (a big MRC/industry funded consortium working on psoriasis) and the A-STAR Registry, a new nation-wide registry of patients with severe atopic eczema requiring disease-modifying therapy.

Collaborations with industry include a collaboration with Nick Lench at Congenica (an existing collaborator with some of the GeCIP researchers) and existing collaborators of the skin research community in the UK including Unilever, GSK, Novartis, Sanofi etc.

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

Training will be an important part of the Rare Skin Diseases GeCIP. **Dermatology Specialty Registrars:** There is an active Dermatology trainee network through the British Association of Dermatologists (BAD) with annual training days and an afternoon meeting at the BAD annual meeting, attended by the majority of trainees. Currently ~10% of Dermatology trainees

countrywide undertake a higher degree. Efforts to improve this have included bursaries to attend the British Society for Investigative Dermatology meeting, NIHR Senior Investigator/BAD-sponsored taster weeks in the laboratory, an annual THESIS meeting (where trainees who are undertaking or have completed a higher degree present), and a British Association of Dermatologists-sponsored Research Techniques course based at the Blizard Institute, Barts and the London School of Medicine and Dentistry led by Edel O'Toole and David Kelsell (members of GeCIP Steering Group). The Research Techniques course includes some bioinformatics and mutation analysis. The Skin Domain leads will also help prospective applicants for Wellcome Trust and Medical Research Council training fellowships by offering practice interviews and taster weeks in the laboratory. The BADGEM group (which includes the leadership of the Skin GeCIP) has proposed a course to the Wellcome Trust on Genomics in Dermatology which has been accepted and the first course will be held in October 2016. Interested Specialty Registrars in Dermatology will be encouraged to undertake a MSc in Genomic Medicine. Bursaries could also be offered for successful clinical training fellowship awardees to attend courses such as the Wellcome Trust Advanced Course on Human Genome Analysis during their studentship.
Consultant Dermatologists, Associate Specialists in Dermatology and Clinical Geneticists: The Skin Domain will present an annual update at the British Association of Dermatologists annual meeting (British Association of Dermatologists Dermatology and Genetic Medicine and British Society for Paediatric Dermatology subgroups) and the Dermatogenetics meeting at the British Society for Genetic Medicine annual meeting. We would envisage that the GEL project will critically inform the further development of the BADGEM network so that this becomes the GEL legacy for delivering Dermatology genomic medicine in the near future. Consultants that are actively involved in recruiting patients will be encouraged to become part of the BADGEM network, if they are not already members. An annual meeting will be hosted by the Skin Domain for educational purposes and to review cases and results of interest. There is already evidence of interest from both dermatology and genetics; an educational meeting held on dermatogenetics recently at the Royal College of Physicians attracted a full house of 270 attendees. Phenotypic subdomains will also discuss cases on a regular basis. Limited bioinformatics training will be provided for Genomic Medicine Centre clinical leads and clinicians actively involved in phenotypic subdomains to improve understanding of genetic data and reports. **Specialist Nurses**: Specialist nurses with an interest in rare skin disease (mainly paediatric dermatology nurses and epidermolysis bullosa nurses) will be updated annually at the British Dermatology Nursing Group meeting. **Online Training:** Members of the Skin Domain will endeavour to update further the Genetics section of Dermatology in E-learning for Health which is accessible to all NHS employees.
Patients: Patients can be updated and educated about developments through reports in patient support group newsletters such as the Ichthyosis Support Group, Pachyonychia Project and DebRA UK.

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

The GeCIP group for Skin includes national and international experts in the field, alongside clinicians in practice in teaching hospitals and general hospitals across the UK:

The leadership team consists of Edel O'Toole, Neil Rajan and John McGrath. Up to 100 individuals have signed up to be part of the team including clinicians and scientists. Dermatology is a relatively small specialty and as such the UK community has a proven track record of working effectively together, evidenced by UK-wide collaborative ventures including the dermatology genetics network 'BADGEM' (http://badgem.org.uk/) as well as an internationally-acclaimed therapeutics register 'BADBIR' (http://www.badbir.org/), and a recently established network for

translational research 'UK-TREND' (http://www.uktrend.org/), each under the umbrella organisation of the British Association of Dermatologists.

Subdomain-leaders will take responsibility for analysis of data from individuals recruited to their domain in collaboration with recruiting clinicians (where this is desired, not all may be academically active).

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 Skin GeCIP has a validation and feedback sub-domain that is composed of molecular geneticists who are already responsible for interpreting genetic variants for currently available genetic tests in various clinically accredited genetics laboratories. A network of international collaborators is also available. By liaison with sub-domains (via respective leads) they will be able to access additional disease-and gene-specific expertise available within the Skin GeCIP in order to assess pathogenicity of variants. This sub-domain will be the point of contact for the Cross-cutting 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?*

The identification of genes, genetic variations and mechanisms resulting in skin disease offers the opportunity for very substantial benefit to patients and the NHS as well as clinical and academic researchers:

Patients and families

It has been clearly established from patient support group surveys (quote the ISG?) that patients place high value on having a molecular diagnosis to 'name' their disease. This allows a fuller explanation by the clinician, regarding disease mechanism and any therapies available, as well as likely prognosis. Families benefit from the understanding of inheritance pattern and risk of recurrence.

The NHS

Genetic skin disease places a life-long burden of care on the NHS are our current treatments are predominantly aimed at relieving symptoms rather than addressing the underlying disease mechanisms. The opportunity for NHS clinicians to collaborate in bringing together patients of similar but unsolved phenotypes offers a unique and powerful opportunity to identify genetic mechanisms. This offer the opportunity for improved molecular diagnoses for these and future patients.

Researchers

Genetic data generated from this project will place the UK at the forefront of the world in identifying genetic mechanisms to improve our understanding of skin and multi-system disease. Skin as organ offers opportunities to reveal genetic mechanisms. Examples of dermatogenetic findings that have improved understanding of multisystem disease include mutations in *IRHOM2* causing keratoderma and oesophageal cancer (6) and loss of function mutations in *FLG,* which encodes the skin barrier protein filaggrin, playing a role in atopic eczema, asthma and food allergy.

Pharmaceutical industry

Novel genetic mechanisms may be amenable to specific therapeutic interventions and are likely to be of considerable interest to the pharmaceutical industry (see below).

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?*

There are several attractive incentives for rare disease drug development. These include licensing advantages for orphan diseases with regulatory authorities such as the FDA and the MHRA, as well as the lifelong need for therapy, making this financially viable to the pharmaceutical industry. Examples include the use of Smoothened inhibitors in patients with Naevoid basal cell carcinoma syndrome, which was licensed in the UK by NICE, at an annual drug cost per patient of approximately £76000. Dermatology is also recognised by the pharmaceutical industry as a specialty where severe phenotypes of common diseases such as psoriasis are treated with expensive antibody therapies. Existing partnerships with commercial partners are in place in large scale projects in psoriasis that aim to develop personalised treatments as part of stratified medicine approach(PSORT), as well as smaller scale interactions in programmes such as the MRC proximity to discovery scheme. As such novel genetic discoveries where a targeted biological intervention is feasible is in dermatology likely to attract commercial interest and exploited through existing links.

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

1: Hüffmeier U, Wätzold M, Mohr J, Schön MP, Mössner R. Successful therapy with anakinra in a patient with generalized pustular psoriasis carrying IL36RN mutations. Br J Dermatol. 2014 Jan;170(1):202-4.

2: Onoufriadis A, Simpson MA, Pink AE, Di Meglio P, Smith CH, Pullabhatla V, Knight J, Spain SL, Nestle FO, Burden AD, Capon F, Trembath RC, Barker JN. Mutations in IL36RN/IL1F5 are associated with the severe episodic inflammatory skin disease known as generalized pustular psoriasis. Am J Hum Genet. 2011 Sep 9;89(3):432-7.

3: Paller AS, van Steensel MA, Rodriguez-Martín M, Sorrell J, Heath C, Crumrine D, van Geel M, Cabrera AN, Elias PM. Pathogenesis-based therapy reverses cutaneous abnormalities in an inherited disorder of distal cholesterol metabolism. J Invest Dermatol. 2011 Nov;131(11):2242-8.

4: Palmer CN, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, Goudie DR, Sandilands A, Campbell LE, Smith FJ, O'Regan GM, Watson RM, Cecil JE, Bale SJ, Compton JG, DiGiovanna JJ, Fleckman P, Lewis-Jones S, Arseculeratne G, Sergeant A, Munro CS, El Houate B, McElreavey K, Halkjaer LB, Bisgaard H, Mukhopadhyay S, McLean WH. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. Nat Genet. 2006 Apr;38(4):441-6.

5: South AP, Uitto J. Type VII Collagen Replacement Therapy in Recessive Dystrophic Epidermolysis Bullosa-How Much, How Often? J Invest Dermatol. 2016 Jun;136(6):1079-81.

6: Blaydon DC, Etheridge SL, Risk JM, Hennies HC, Gay LJ, Carroll R, Plagnol V, McRonald FE, Stevens HP, Spurr NK, Bishop DT, Ellis A, Jankowski J, Field JK, Leigh IM, South AP, Kelsell DP. RHBDF2 mutations are associated with tylosis, a familial esophageal cancer syndrome. Am J Hum Genet. 2012 Feb 10;90(2):340-6.

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)*

Analyses will use phenotype data, filtered variant lists, VCF and BAM files for all individuals recruited under the rare skin disease eligibility criteria. Should skin patients appear to have relatives recruited separately with a non-skin phenotype to a second domain, we will approach the relevant domain to collaborate.

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

Successive filtering steps will include

1. single variant analysis: which will imply the Identification of obvious disease causing variants using Ingenuity and an in-house High Throughput Sequencing Pipeline;
2. haplotype association analysis: for which we will use SVS and PLINK/SEQ;
3. collapsing methods for the identification of rare variants SVS, PLINK/SEQ, in house methods to be developed; data driven (after the identification of possible candidate genes) pathway analysis: starting with the literature generation of pathways (PALM-IST, Pathway Assembly Literature Mining) and continuing with an array of first (ORA), second (FCS) and third generation (PT) pathway analysis tools, as appropriate for the task.

Then:

4. When sufficient cases are available to proceed, video teleconference meetings with relevant stakeholders will take place to discuss candidate variants. These will be ranked and methods to validate and strengthen evidence for pathogenic variants will be reviewed. As appropriate, findings will be channelled to cross cutting teams involved in functional analysis to investigate the biological implications of these changes. Feedback to clinical teams will be made with information of the strength of suspicion of pathogenicity of the newly identified variant.

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)*

Key phenotype data will vary according to the phenotype for which participants are recruited. Examples of key data points by phenotype are as follows:
Severe atopy: serum IgE, recurrent infections, asthma, hay fever.
Palmoplantar keratoderma: Type of keratoderma- focal, diffuse or punctate. Other associated features eg hair or nail phenotypes, follicular hyperkeratosis.
ARCI: Presence of a collodion membrane, other extracutaneous features.
Ectodermal dysplasia: Hair/nail. Hair/teeth/sweating. Teeth/nail.
Etc.

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)*

We plan to use the Genomics England analysis pipeline for this.

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)*

Most of the tools in the Embassy apps list (except the ones dedicated to alignment and mapping).

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)*

BAM files and VCF files from previously investigated rare skin disease patients (WES or NGS).

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

It is anticipated that 20-40 cores will be needed for initial .vcf file-based variant and phenotype data analyses planned. Resources needed will depend on the numbers of patients recruited and whether additional analysis of raw data (.bam) files is needed.

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)*

The initial emphasis will be focused on availability of genome data. Utilizing omics samples will be driven by gene discoveries, sufficient recruits for various diseases, and therefore likely initiated after a significant number of samples have been accrued.

Data access and security	
GeCIP domain name	Skin
Project title *(max 150 characters)*	Genetic causes of rare skin diseases

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*

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

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

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

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

X☐ *Other health use - clinical audit*

X☐ *Public health purposes*

☐ *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.

Other attachments

Attach other documents in support of your application here 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)