

Facial Genetics Workshop

London 2016



Thursday 7th
and Friday 8th
April

First meeting

A meeting of Facial Genetic Experts

Purpose

An invited group of international researchers met in London on 7th/8th of April 2016. This is the first meeting of a multi-specialty group of facial genetic experts around the world. The Facial Genetics meeting was organized by Stephen Richmond (Cardiff University) and Peter Claes (KU Leuven) as a direct result of the bilateral agreement between the Universities promoting shared research objectives not only to bring two Universities closer together but linking other like-minded researchers worldwide. The meeting was held in London, halfway between Cardiff and Leuven.

Evening get-together

The format of the meeting was to arrange a get-together on the Thursday evening to "break the ice" over a meal. You can always tell a successful evening when nobody wants to leave,



19 guests enjoy an evening in the restaurant

Presentations

Friday began with a series of presentations from Peter Claes (Belgium) and Mark Shriver (Penn State, USA) showing novel statistically robust numeric methods to explore distinctive facial features and gene associations. Mark did a stoic presentation over the noise of road re-surfacing outside the window! This was followed by Seth Weinberg (Pittsburgh, USA) and Mary Marazita (Pittsburgh, USA) who also presented soon to be published data on new replicated genes associated with facial features. Peter Hammond (Oxford) described the opposite effects on facial morphology resulting from gene dosage sensitivity. Lavinia Paternoster (Bristol) re-traced the GWAS process of the 1st discovery and replication of the *PAX3* gene associated with the nasion – mid-intercanthal distance and methods for managing findings that are just below the 10⁻⁸ threshold. Christoffer Nellåker (Oxford) described machine-learning techniques to extract phenotypic features from routine photographs. Marcus de Jong (Leiden, NL) described his techniques and enhancements to reduce errors by 0.5mm in automated landmarking. Stephen Richmond (Cardiff) completed the morning again with unpublished data related to heritability in a study of father and offspring and female twins as well as highlighting automated systems to capture the shape of lips and noses. Caryl Wilson-Nagrani (Cardiff) present unpublished data on gene associations with lip traits and the increase in prevalence of certain lip traits in parents of cleft children compared to controls.

Open workshop

After lunch an inclusive and interactive meeting (European style-all sitting in a large rectangle facing each other) was moderated by Peter Claes to gain opinions and directions on the following topics.

- A) Data acquisition
 - i) systems – accuracy/validity
 - ii) processing and registration of images
 - iii) managing errors
- B) Phenotyping techniques
 - i) categorical subjective / automated
 - ii) facial anthropometrics
 - iii) dense surface mesh
 - iv) preserving asymmetries
- C) Genotyping techniques
 - i) samples
 - ii) arrays
 - iii) imputation
- D) Study types/ data bases
 - i) population
 - ii) twins (heritability)
 - iii) trios (heritability)
- E) Combining data sets
 - i) admixtures
 - ii) age
 - iii) ancestry
 - iv) ethnicity
- F) Management of data
 - i) Statistical management/processing

Closing remarks

There was a very open and honest atmosphere throughout the day with researchers willing to collaborate and share data and software. The participants expressed how much they enjoyed the meeting and there was a strong proposal from the group to organize a Facial Genetics meeting annually in London with two new coordinators each year. Those who couldn't make it this year will fix their diaries for the next meeting in May 2017.



Attendees - Front row: Yulia Hicks, Lisette Jong, Carine Carels, Seth Weinberg, Hawraa Abbas, Alexei Zhurov, Raija Lahdesmaki Caryl Wilson-Nagrani, Middle row: Lisette Jong, Hilde Peteers, Lavinia Paternoster, Stephen Richmond, David Marshall, Mary Marazita, Markus De Jong, Back row: Pertti Pirttiniemi, Jason Hodgson, Mark Shriver, Peter Hammond, Christoffer Nellåker, Peter Claes.

We are looking forward to an exciting future of shared advancements in the field of facial genetics.

Stephen Richmond



Peter Claes



Contents Page

Overview	1
In Focus	2-4
Peter Claes	5
Mark Shriver	6
Seth Weinberg/ Mary Marazita	7
Peter Hammond	8
Lavinia Paternoster	9
Christoffer Nellåker	10
Markus de Jong	11
Stephen Richmond/ Caryl Wilson-Nagrani	12
Acknowledgements	13

Cardiff University®
KU Leuven®

Any material taken from this document must be fully referenced.

Facial Genetics Workshop – In Focus

The morning presentations showed clear homogenous and overlapping research goals and interests amongst all participants. In contrast, the research equipment, datasets and approaches were very heterogeneous, ensuing different research results and outcomes. In an attempt to make future work reproducible and to stimulate collaborations, different aspects of a typical facial genetics research pipelines were discussed. During the afternoon session participants were able to relate their experiences, opinions and express willingness to share datasets as well as software.

A) Facial Data acquisition

Various acquisition systems for capturing 3D faces were identified. The Konica/Minolta colour laser systems (Vivid 900/910) are now obsolete. No one present had routine experience in using the Dimensional Imaging system. Other systems such as the [3dMD](#) and [Vectra H1](#) were discussed in more detail. The Vectra H1 is a very transportable system and can be purchased for around £/€8k. The flexibility of the system will enable more facial images to be taken compared to a static system. There was a general consensus that the Vectra H1 camera resolution was as good as the 3dMD system. The acquisition of very young individuals however, can prove to be challenging with the Vectra H1. The Vectra H1 and 3dMD systems were being compared in Pittsburg using standardized mannequin heads. The Vectra H1 and Konica Minolta 910 were also being compared in Cardiff.

It was reiterated that standardization of facial posture with lips together in a neutral pose is key to acquiring a reproducible image. A standardized way to quantify spurious expressions is currently missing. A good quantification of spurious expressions allows for the correction of image data, prior to subsequent analysis.

B) Phenotyping techniques

Various facial phenotyping approaches exist, going from conventional morphometric analysis using distances, angles... (facial anthropometrics) to geometric morphometrics using complete landmark configurations typically followed by a principal component analysis (PCA). The amount of facial landmarks used are either sparse or spatially-dense, manually or automatically indicated. Additionally, complex facial traits can be identified subjectively and/or automatically, resulting in labeled/categorical data.

Each phenotype approach has its advantages and disadvantages. Facial anthropometrics for example creates the ability to measure the same facial traits across different datasets and can be related to historical population cohorts. However, subsequent analysis is typically limited to be univariate and combined interpretations of the results are challenging. PCA of complete landmark configurations provides multivariate analyses, but the extraction of principal components is dependent on the underlying dataset. Categorical phenotyping of faces allows the ability to capture complex shape variations, but requires the determination of possible categories and extensive training to generate labels for all data samples.



Facial Genetics Workshop – In Focus (Continued)

All phenotyping approaches can benefit from automated procedures, to enhance repeatability on large datasets. Some automated landmarking and labeling techniques are becoming available.

In some facial evaluations facial asymmetries are preserved and in others averaged for the left and right sides of the face. The two techniques require further evaluation.

Currently phenotypic traits are typically defined by the investigators prior to subsequent analysis. Another focus would be to learn interesting phenotypic traits by a genotype driven data analysis (heritability study e.g.). Such an inverse focus can be further empowered by the use of machine learning. However, warnings have been raised that machine learning is not to be overestimated, as it requires a large amount of learning data and a good knowledge of the underlying algorithms to avoid over-fitting and finding spurious phenotypic traits of interest.

C) Genotyping techniques

Like phenotyping, genotyping techniques are also different, using different sampling material (saliva, blood...), customized genotyping arrays versus personal genomics companies...

A promising way forward is the ability to impute missing genotypes following the Hapmap and/or 1000 Genome imputation protocols. Available software includes Minimac3 and IMPUTE2. Genotyping and imputation usually follows a standard approach as used in other complex trait consortia (e.g. [GIANT](#)).

A different importance is given to possible SNP associations when the SNP was actually genotyped or imputed, with a stronger belief in the former.

Current state-of-the-art on GWAS in a variety of traits, typically reports on both genotyped and imputed SNP analysis. Hardly any differences are observed, confirming that imputed SNP data is a valuable tool to use.

A suggestion was made to think about phenotype imputation to similarly obtain comparable phenotyping across datasets.

Reference:

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet.* 2014 Feb;46(2):100-6.

D) Study types/ data bases

All study types are valid and contribute to the knowledge of facial form.

Two publically available data repositories exist. The [FaceBase](#) data contain Caucasian and Tanzanian samples (approximately 2500 of each). The [ALSPAC](#) sample consists of 4747 15 year old children and 995 fathers with 465 sons and 530 daughters.

Other data bases are accessible through bilateral research agreements. For example, Twin data bases are available through [Visigen](#).

Heritability can be evaluated using [GCTA](#), an approach that gained a lot of interest and that uses genomic information of (mixed) related and unrelated individuals. The study design always depends on the question to be asked. The range of study designs provides the opportunities to ask different questions about heritability (twin/families/ populations), familial segregation of rare traits (trios) as well as genetic association for population level variance in the phenotype (populations). The relative phenotype/ genotype contributions of parents to offspring is currently the least studied in normal facial variation.



Facial Genetics Workshop – In Focus (Continued)

E) Combining Datasets:

The aim is to increase the number of samples and power in association and heritability investigations.

One approach is to perform a meta-analysis on the results of each dataset separately. When using facial anthropometrics this is certainly possible, under the constraint that the same landmarking array was used, which is generally not the case. In contrast, when using PCA based facial traits a meta-analysis becomes less straightforward to execute.

Another approach is to actually pull the raw images of different datasets together and to ensure for a consistent phenotyping. It may be appropriate to circulate a facial calibration dataset to ensure reliability and validity of manual and automated landmarking techniques. Effects of different scanning equipment and operators are to be investigated as previously noted.

It becomes more challenging when datasets from populations with different backgrounds and age ranges are combined. Special attention was given during the discussion about dealing with very young individuals in a sample. Facial growth and development is a complex and non-linear process, making correction and/or condition of facial variation on age questionable and open to further improvement. Linear correction methods are inherently flawed and a way forward might be the use of kernel-based correction methods. For example, using the concept of “facial signatures” in the work of Peter Hammond. The key is that each facial sample is matched against an average face from a group of faces sharing the same confounding values, such as age, background, sex...

F) Management of data

Data management is not to be underestimated, and publically available data repositories, such as FaceBase, can provide the (existing) necessary support.

Restrictions on making data publically available are mostly due to signed informed consents that are too restrictive. Older datasets have this problem in which the participants did not consented for their data to be publically available. Future sampling efforts, preferably employ the informed consents as used in FaceBase. Respective investigators involved in the FaceBase project, can easily be contacted to give advice on future ethics applications and sampling protocols.

The creation of publically available software repositories was greatly supported and could be an interesting extension to FaceBase.

The meeting was brought to a close followed by drinks in the bar and or journey home!

The weather and road works did not detract from a very successful meeting.

Feedback from the meeting:

“I enjoyed it very much and my impression was that everyone else did too! Look forward to meeting up with you all again”.



Hierarchical Modular Facial Phenotyping

Peter Claes and Dzemila Sero

Genome Wide Association Studies (GWAS) focus on “ $p < 5 \times 10^{-8}$ ”, which is simply the minimum cut-off point to claim a SNP discovery. Pairing relevant genetic SNPs with relevant phenotypic traits can be challenging.

The current trend in GWAS, is to increase the number of SNPs (up to 10M) improving the chance of including relevant SNPs. Investigating such a multitude of SNPs is in contrast to the relatively few facial features e.g. nose and mouth width or principal component scores of landmark configurations. The phenotypic complement to genomics is phenomics which aims to obtain high-throughput and high-dimensional phenotyping (Houle et al., 2010). Advances in facial data acquisition techniques with increasing resolutions can provide detailed surface facial morphology to the level needed for phenomics. Facial phenotyping needs to be complete and compact. Completeness - captures all relevant phenotypic traits/variations. Compactness - minimum dataset needed to reduce the computational burden when analyzing each SNP in a GWAS and to avoid compromising the study-wide significance ($p < 10^{-8}$). However, at first sight completeness and compactness appears to be contradictory. For example, spatially-dense landmarking is complete but lacks compactness, having to deal with thousands of variables. Principal Component Analysis (PCA) provides compactness but may not be comprehensive as local facial variations are often diluted or even missed.

We are currently investigating the use of modularity in facial morphology in combination with multivariate statistics to obtain a complete and compact phenomic facial phenotyping for GWAS. The human face is made up of many parts that function as a whole. Individuals display diverse and heterogeneous facial units due to their specific role. The face is organized into multiple facets, each to perform a specific task in accordance to its embryological origin, anatomy and function.

In the past 20 years of research in molecular and biology systems, the integration into subunits has been categorized under the heading of modularity, which implies the division of a structure into several parts that display high integration within the variables (modules), but few and weak interactions across the modules.

Following the constructs of modularity in geometric morphometrics we were able to hierarchically subdivide the face into a series of global to local modules as shown in Figure 1, for 6 hierarchical levels resulting in 63 modules. All the landmarks in each of these modules separately undergo a Generalized Procrustes Alignment (GPA) followed by a PCA reduction of the dimensionality.

The result is a complete description of a global and local perspective of facial variation that is compactly described in detail for each of these modules. SNP associations are tested using a multivariate canonical correlation analyses against all of the PCs simultaneously in each of the modules, resulting in 63 dependent tests.

Using 1490 European subjects in a discovery panel, 25 SNPs, out of 100000 genome-wide selected SNPs in low LD, achieved genome-wide significance. An example SNP, rs2765529 located in *TBX15*, is shown in Figure 2. The SNP shows a highly localized effect in the forehead, which is consistent with the facial characteristic of frontal bossing in the Cousin Syndrome.

Peter Claes & Dzemila Sero
KU Leuven, ESAT/PSI
Medical Imaging
Research Center, UZ
Gasthuisberg
Herestraat 49 - box
7003
3000 Leuven
Belgium



Figure 1 Hierarchical modularization of facial morphology.

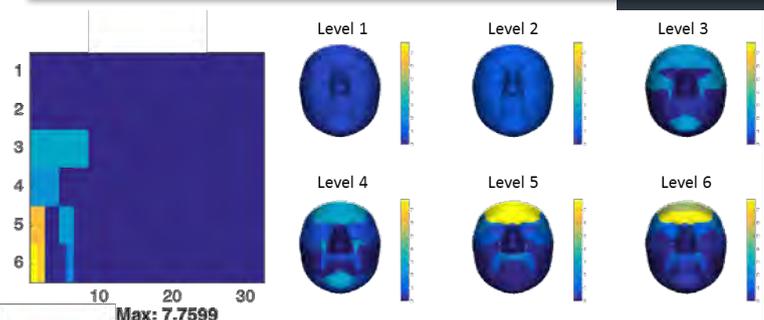


Figure 2 $-\log_{10}$ p-values with rs2765529, left: compact visualization (vertical axis = level index, horizontal axis = module index), Right: visualization on the face using the modules at each level separately.

Conclusion

Facial morphology comprises of many traits which implies high-dimensional and multivariate quantitative characteristics. Future facial phenotyping techniques are necessary to deal with these issues.

Reference

Houle D, Govindaraju DR, Omholt S. Phenomics: the next challenge. Nat Rev Genet. 2010 Dec;11(12):855-66.

Modeling Genes and Other Factors Affecting Variation in the Human Face

Mark Shriver

Over the last 5 years there has been little concerted effort focused on applying modern human genetic methods to explore the biological basis of facial features.

Difficulties in addressing the genetic underpinnings of human genetic variation include;

- 1) the complexity of the human face – arguably, the face is a composite of hundreds of connected developmental modules
- 2) the small effect sizes expected for many alleles affecting typical-range variation,
- 3) the lack of a sufficient human genetic analytical framework for mapping highly multivariate traits.

Although there are indeed well studied facial metrics (inter-landmark distances, angles, and parameters from data reduction methods), it is important to be as agnostic as possible with respect to the definition of what sort of facial “traits” might be affected by variables like SNP genotypes. The approach, exemplified by bootstrapped response-based imputation modeling (BRIM; Claes et al., 2014) defines the trait of interest simultaneous with the association test. Instead of asking, does this gene affect a specific facial trait, we ask do any facial features vary with variation in this gene? This approach to phenotypic definition is fundamental and there are similarities between this approach and the work of Kun Tang and colleagues and Peter Hammond and colleagues.

Overcoming the issue of small average effects of alleles requires not only efficient statistical methods, but large sample sizes. For over 15 years we have employed an innovative participant recruitment approach which provides genetic test results to participants. Over the past three years the “direct-to-consumer” based test we have provided to the participants is also the primary data source for GWAS and other gene mapping approaches. We have also been sampling across the range of human population and our current collection numbers about 8,500 with a plan to collect an additional 5,000 new participants over the next four years. Regardless, the effort we or other individual groups invest in collecting participants, there will always be a greater benefit by working together. The Facial Genetics Workshop provides the opportunity of pooling data sources and expertise. There is a real need for, and happy to help coordinate and recruit for an open public dataset.

Reference

Claes P, Liberton DK, Daniels K, Rosana KM, Quillen EE, Pearson LN, McEvoy B, Bauchet M, Zaidi AA, Yao W, Tang H, Barsh GS, Absher DM, Puts DA, Rocha J, Belezza S, Pereira RW, Baynam G, Suetens P, Vandermeulen D, Wagner JK, Boster JS, Shriver MD. Modeling 3D facial shape from DNA. *PLoS Genet.* 2014 Mar 20;10(3).

Our most recent facial models that include over 3,500 participants from many different human populations help illustrate approaches to discovering both the genetic and non-genetic effects on facial variation, modeling the important variables, and validating the results. We are able to model quite effectively the effects of a number of factors affecting typical-range variation in the human face. These include age, body size (height and weight), ancestry, and sex (Figures 1 and 2). Although, strictly speaking, we may not need to understand the effects of these variables to detect the genes affecting variation in the face, very few are interested in the narrow question of which are these genes. Not only do growth and development have fundamental effects on the human face, but these processes along with the factors affecting genetic and biological background are too easily dismissed as simply sex and ancestry.

Finally, an important source of data on human faces is humans themselves as observers who can be asked to match, rate, and judge facial identities and characteristics. The usefulness of human observers in our studies of human facial femininity/masculinity and facial ancestry highlighting both the accuracy of the observer assessments as well as the inherent biases some of which are compellingly consistent with simple, but unexpected, evolutionary expectations.

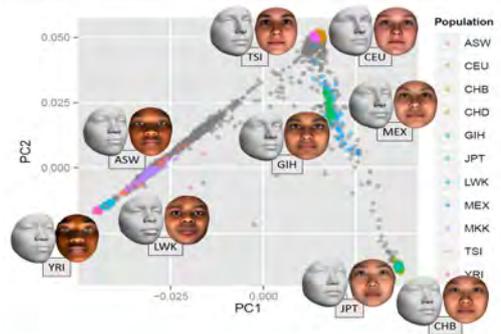


Figure 2 Population variation

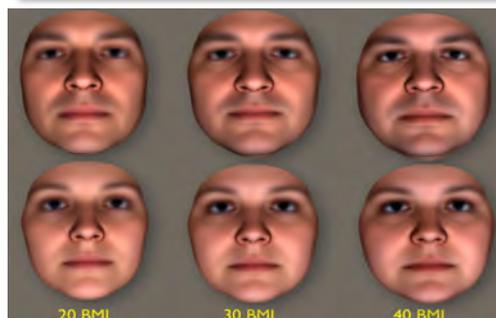


Figure 2 Facial effects of BMI

Mark D. Shriver
Professor of
Anthropology and
Genetics
Penn State University

Using FaceBase resources to investigate the genetics of normal human facial variation

Seth Weinberg and Mary Marazita

Recent genetic studies of normal human facial traits using data available through the National Institute of Dental and Craniofacial Research (NIDCR) FaceBase Consortium were presented.

The first of these studies was a genome-wide association analysis of 20 facial linear distances derived from 3D facial surface images of 3118 healthy participants. Our major findings included genome-wide significant associations ($p < 5 \times 10^{-8}$) for cranial base width at 14q21.1 and 20q12, intercanthal width at 1p13.3 and Xq13.2, nasal width at 20p11.22, nasal ala length at 14q11.2, and upper facial depth at 11q22.1. Of note, several genes in the associated regions are known to play roles in craniofacial development or in syndromes affecting the face: *MAFB*, *PAX9*, *ALX3*, *HDAC8*, and *PAX1*.

In our next study, we tested for genomic associations with facial shape phenotypes (principal components), derived from a geometric morphometric analysis of 3D facial landmarks. We observed a genome-wide significant association with principal component 10, which explained just over 3% of the shape variation and primarily involved upper lip, philtrum and nose landmarks.

The top signal involved a SNP within 500b downstream of the gene *RPGRIP1L*, which when mutated results in Meckel syndrome (type 5). This syndrome and the mouse knockout share significant mid-facial phenotypes including cleft lip.

Finally, we presented some initial results of a candidate SNP analysis of cranial vault shape involving genes implicated in craniosynostosis. We identified several suggestive hits, the most interesting involving SNPs in the *IHH* gene, which in our data was associated with cranial vault length. This is notable because CNVs in *IHH* are associated with a syndromic form of craniosynostosis involving the sagittal suture and resulting in excessive lengthening of the head.

Overall, our results provide evidence that common variants in regions harboring genes of known craniofacial function contribute to normal variation in craniofacial features. Our work may provide insights into the pathways and mechanisms controlling normal and abnormal craniofacial morphogenesis.

Reference

Hochheiser, Aronow BJ, Artinger K, Beaty TH, Brinkley JF, Chai Y, Clouthier D, Cunningham ML, Dixon M, Donahue LR, Fraser SE, Hallgrimsson B, Iwata J, Klein O, Marazita ML, Murray JC, Murray S, de Villena FP, Postlethwait J, Potter S, Shapiro L, Spritz R, Visel A, Weinberg SM, Trainor PA. The FaceBase Consortium: a comprehensive program to facilitate craniofacial research. *Dev Biol.* 2011 Jul 15;355(2):175-82.



Seth Weinberg and Mary Marazita
Center for Craniofacial and Dental Genetics
Department of Oral Biology
School of Dental Medicine
University of Pittsburgh
100 Technology Drive, Suite 500
Pittsburgh, PA 15219

Face Shape And Gene Dosage Sensitivity

Peter Hammond

Opposing variation in growth, head size, cognition and behaviour is known to result from deletions and reciprocal duplications of some genomic regions. Normative inversion of face shape, opposing difference from a matched norm, was recently proposed as a basis for investigating the effects of gene dosage on craniofacial development. Using dense surface modelling techniques to match any face to a facial norm of controls of matched age, sex and ethnicity, the individual's face shape differences from the matched norm are reversed to produce a *normative inversion*.

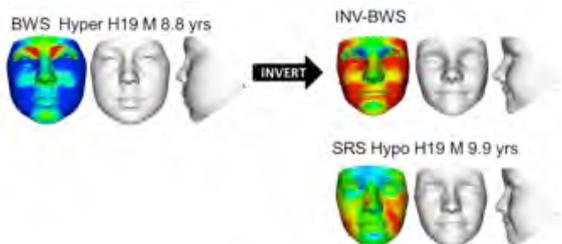
The talk demonstrated that for five genomic regions, 4p16.3, 7q11.23, 11p15, 16p13.3 and 17p11.2, such inversion for individuals with a duplication or (epi)-mutation produces facial forms remarkably similar to those associated with a deletion or opposite (epi)-mutation of the same region, and vice versa. For example, in A), a face of an individual with Beckwith-Wiedemann syndrome (BWS) is inverted to INV-BWS resulting in a face shape that is very close to the face (immediately below) of an individual with Silver-Russell syndrome (SRS). Similarly, in B), the face of an individual with a duplication of 4p16.3 is inverted to produce facial features very like the individual below with Wolf-Hirschhorn syndrome (WHS), associated with a deletion of that region.

Peter Hammond,
Nuffield Department
of Obstetrics and
Gynaecology,
Oxford University,
UK.

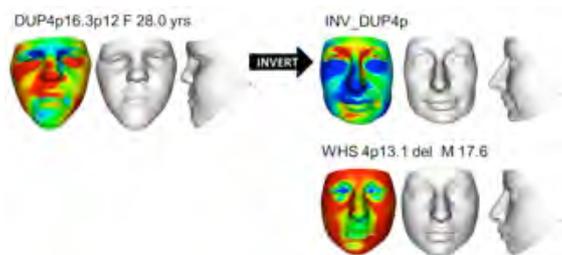
Other genomic regions with similar opposing face shape-copy number variation (CNV) relationships have been identified but remain unpublished due to the lack of reliable or published cases in the literature.

For example, C), to the right, facial inversions of individuals with Prader-Willi syndrome due to maternal uniparental disomy (UPD) of 15q11q13 have remarkable similarity to published cases of 15q11q13 paternal duplication.

Normative inversion inverts every feature of a face relative to the matched norm whereas a genomic CNV is often a localized deletion or duplication. Therefore, more realistic comparisons are faces associated with reciprocal copy number variation.

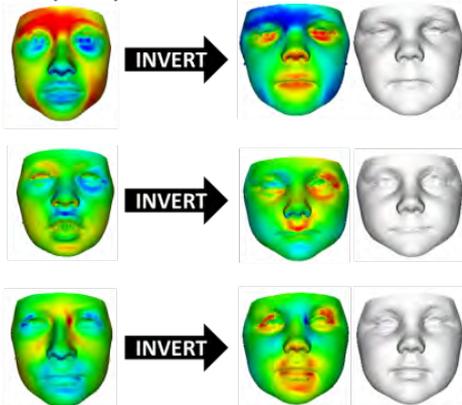


A) Inversion of Beckwith-Wiedemann to Silver-Russell



B) Inversion of dup4p16.3 to Wolf-Hirschhorn

PWS
mat UPD
15q11-q13



pat dup
15q11-q13



Mohandas et al,
Am J Med Genet
1999, 82:294-300;
Urraca et al,
Aut Res, 2013

Reference

Hammond P, McKee S, Suttie M, Allanson J, Cobben JM, Maas SM, Quarrell O, Smith AC, Lewis S, Tassabehji M, Sisodiya S, Mattina T, Hennekam R. Opposite effects on facial morphology due to gene dosage sensitivity. Hum Genet. 2014 Sep;133(9):1117-25.

Clinical geneticists who have patients with reciprocal CNVs of known dosage sensitive genes should contact the author at peter-hammond@obs-gyn.ox.ac.uk to investigate opposing effects on face shape.

C) Inversions of individuals with Prader-Willi syndrome

Genetic Aetiology Of Facial Traits

Lavinia Paternoster

Various statistical methods can be used to explore genetic aetiology of complex traits; Genome-wide association studies (GWAS), SNP heritability estimates, genetic correlations and Mendelian randomization. These have all been successful for many other complex traits and we expect facial genetics to follow the same path.

Lavinia Paternoster,
MRC Integrative
Epidemiology Unit at
the University of
Bristol, Oakfield
House, Oakfield
Grove, Bristol,
BS8 2BN, UK

Genome-wide association studies (GWAS), SNP heritability estimates, genetic correlations and Mendelian randomization have all been successful for many other complex traits and it is expected that facial genetics will follow the same path. In the ALSPAC study we have 3D facial images from 4747 offspring at age 15 (and 995 fathers). All these facial shells have been manually landmarked for 21 easily identifiable facial features (Figure 1). An additional 7 mid-points were created notably the mid-endocanthion point. Imputed genetic data is also available for 8327 offspring.

A GWAS has previously been conducted on the ALSPAC data, using 2185 individuals in the discovery analysis and 1622 in the replication.

We identified four variants in the discovery that met genome-wide significance ($p < 5 \times 10^{-8}$). One of these replicated ($p = 4 \times 10^{-7}$), this was rs7559271 in the *PAX3* gene. The prominence of the nasal bridge (mid-endocanthion to nasion) is influenced by the *PAX3* gene (Figure 2).

This study had a fairly small sample size and so had limited power to detect associations. GWAS of facial traits also suffers from a huge multiple testing problem, further limiting power.

Following the model of GWAS studies in other complex traits, it is also important that meta-analysis is conducted on all available data to increase sample size.

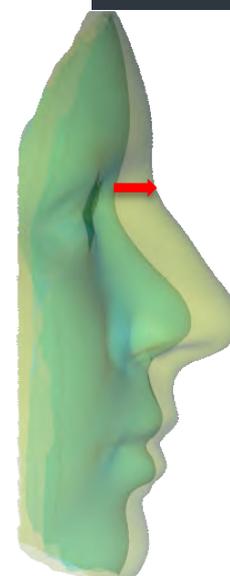


Figure 2: The effect of the *PAX3* gene on the prominence of the nasal bridge

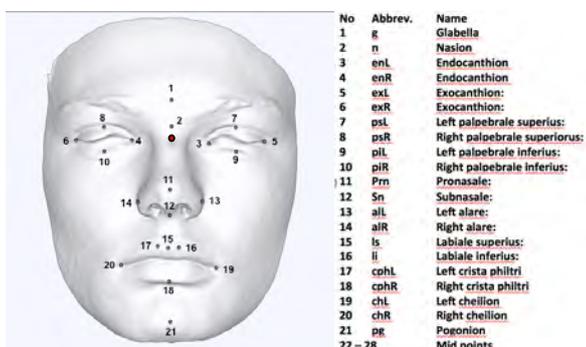


Figure 1: 21 facial landmarks identified and a further 7 mid-points generated.

To have success in future facial GWAS studies, we must be able to reduce the number of tests carried out. This can be achieved in different ways:

- use dimension reduction techniques, such as principle components analysis (PCA), however this was not fruitful in the previous GWAS and other dimension reduction methods may need to be explored.
- select only those phenotypes for which there is evidence of high heritability. This information could come from family studies or using whole genome approaches using unrelated individuals.

Reference

Paternoster L, Zhurov AI, Toma AM, Kemp JP, St Pourcain B, Timpson NJ, McMahon G, McArdle W, Ring SM, Smith GD, Richmond S, Evans DM. Genome-wide association study of three-dimensional facial morphology identifies a variant in *PAX3* associated with nasion position. *Am J Hum Genet.* 2012 Mar 9;90(3):478-85.

Computer Vision For Phenotyping In Rare Diseases

Christoffer Nellåker, Quentin Ferry, Mohsan Alvi, Mitchell Dawson, Julia Steinberg, Caleb Webber, Chris Ponting, David R FitzPatrick, Andrew Zisserman

Using computer vision and machine learning approaches to analyse facial photographs allows us to perform computational phenotyping to narrow the search space for rare disease diagnoses.

Christoffer Nellåker
Department of Physiology, Anatomy and Genetics, Oxford University, UK.

Facial genetics research is currently pursuing 3D imaging modalities, however we work with using normal 2D photographs. The vast volumes of digital photography imaging being collected in globally allow us to apply the latest computer vision and machine learning at the scale of big data. Ordinary photographs are highly variable and typically spurious (non-phenotype relevant) signal dominate the appearance of an image.

We developed a computational pipeline to automatically analyse images (Figure 1). The pipeline finds a face in an image, annotates key facial feature points and extract feature vectors. Next we embed faces in a multidimensional space of possible faces and apply machine learning. The learning reshapes the space to cause spurious signal dimensions to compress and the disease phenotype vectors to expand. In the resulting Clinical Face Phenotype Space (CFPS) patients cluster together, allowing a narrowing of the search space for a disease diagnosis. Crucially the space generalises to unseen, untrained disease diagnoses (Figure 2).

The implication being that CFPS has generalised to reflect aspects of the underlying space of disease phenotype variation. This is supported by a significant correlation with functional genetic pathway analyses for genetic associations in CFPS.

We are building a global consortium of clinical genetic researchers to progress this research. As part of this we are keen to interact with the 3D imaging facial genetics community to integrate 2D and 3D analyses approaches. To this effect a pilot project with Prof. Peter Hammond has begun but we hope that the Facial Genetics Workshop community will be a good forum to build collaboratively on this promising research avenue.

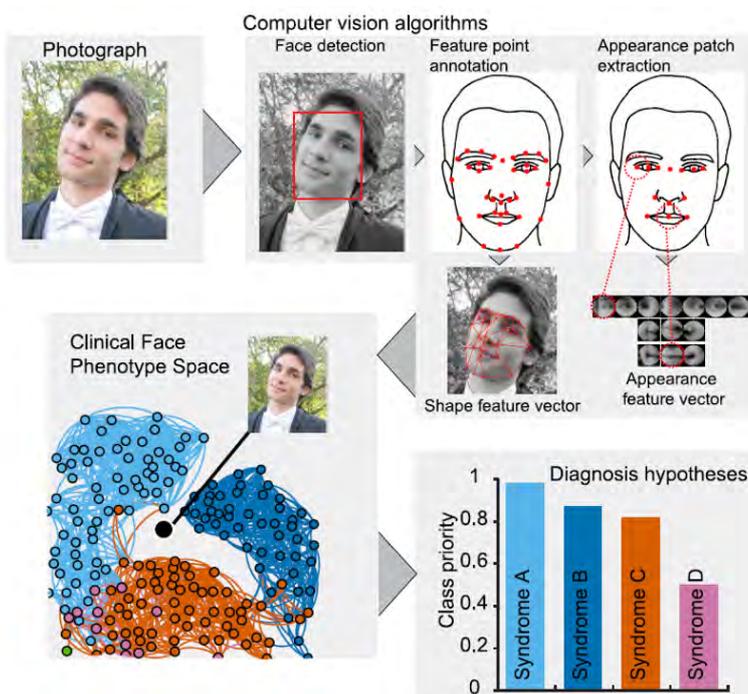


Figure 1: Computational pipeline – from photograph to diagnosis

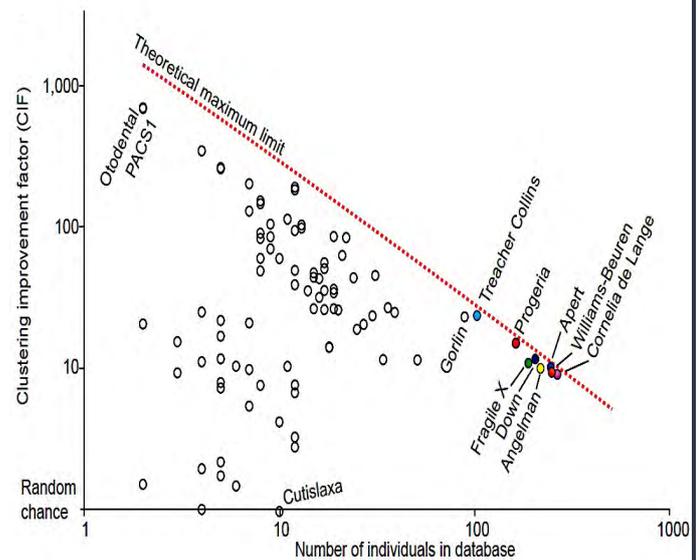


Figure 2: Narrowing of the search space to identify rare disease.

Reference

Ferry Q, Steinberg J, Webber C, FitzPatrick DR, Ponting CP, Zisserman A, Nellåker C. Diagnostically relevant facial gestalt information from ordinary photos. *Elife*. 2014 Jun 24;3:e02020.

An Automatic 3D Facial Landmarking Algorithm Using 2D Gabor Wavelets

Markus A. de Jong

Landmark registration of the human face is an important pre-requisite to determine facial traits for heritability, genetic associations, and medical conditions with distinctive facial characteristics.

Introduction

The approach is to work with 2D-projections of 3D-surface data and to employ 2D algorithms on that transformed data. In this process the original detailed surface data is retained throughout.

The process

1. Map projection:
a region of interest is defined from the frontal face. A two-stage ellipsoid process is used to register the face (Figure 1).

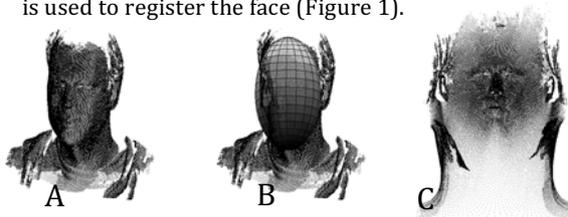


Figure 1: The ellipsoid fitting and map projection process of the human face. A) original point cloud, B) ellipsoid fitting, C) 3D map projection based on an ellipsoid

2. Image feature layer generation:
 - the texture of the 3D model is rendered orthographically using the 3D editing software Blender under full brightness conditions, i.e. without artificial shadows or specular reflections.
 - the relief map (heightmap) is constructed.
 - 3 feature layers are derived: x-axis, y-axis and the Laplacian of Gaussian (Figure 2).

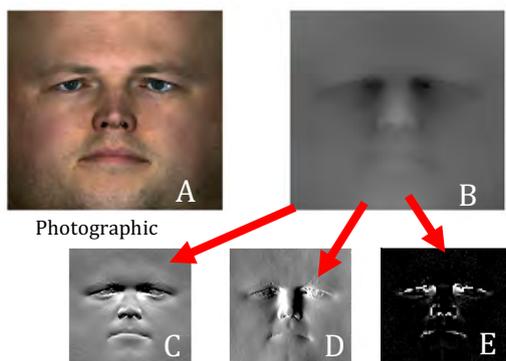


Figure 2: The 5 feature layers generated from a human face image based on the map projection. A) photographic, B) heightmap, C) derivative with respect to X-axis, (D) Y-axis, (E) heightmap Laplacian of Gaussian

3. Training and landmarking:

Elastic Bunch Graph Matching method is employed (EBGM). A maximum correlation template search is performed between a set of example images and the image to be landmarked.

The features used are Gabor-Wavelet transformations of the 21 facial landmarks (Figure 3). The search is performed per landmark with pre-defined constraints. Finally, the registered 2D landmarks are mapped back into 3D, inverting the projection.

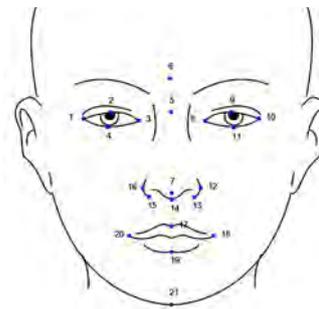


Figure 3: 21 facial landmarks

EBGM was run on all feature layers simultaneously, i.e. Gabor-Wavelet coefficients were extracted from all layers using the same Gabor-Wavelets and resulting coefficients were combined into a single vector. The EBGM algorithm is trained using 30 images with the 21 manually placed landmarks.

Finally, the 2D landmark pixel coordinates were mapped back to the nearest points in the relief map model and subsequently to the 3D landmark coordinates in the 3D cloud. This is achieved by using the inverse of the mapping performed or, more efficiently, using vertex indexing.

Conclusions

The automatic landmarking of facial 3D surface data requires little investment in the training phase in a single iteration (30 training samples taking 1 minute for 21 landmarks). Errors of 1-2mm are achieved for most of the facial landmarks with facial images of differing quality, gender and ethnicity. The algorithm can be easily and quickly re-trained when a different set of landmarks is required.

Reference

de Jong M, Boehringer S, Kayser M, Wollstein A, Ruff C, Dunaway D, Hysi P, Spector T, Liu F, Niessen W, Koudstaal M, Wolvius E. An automatic 3D facial landmarking algorithm using 2D Gabor Wavelets. IEEE Transactions On Image Processing, Vol. 25, No. 2, 580-588, 2016.

Markus A. de Jong
Department of Oral
& Maxillofacial
Surgery, Special
Dental Care, and
Orthodontics,
Erasmus MC
University Medical
Center Rotterdam,
Rotterdam, The
Netherlands

Facial Growth And Facial Genetics

Stephen Richmond, Caryl Wilson-Nagrani, Alexei Zhurov, Jelena Djordjevic, Nizar Mhani,

Perti Perttiniemi

Sequential studies undertaken in Cardiff collaborating with the Universities of Bristol, Oulu and Pittsburgh, also with Visigen have highlighted interesting aspects of facial heritability and facial variation.

The studies

Six studies i) normal variation in a population of 15 year old children (2233 Males and 2514 females - ALSPAC), ii) Fathers and offspring (995 fathers; 465 sons- 530 daughters - ALSPAC), iii) UK twins (263 MZ and 341 DZ -Visigen) iv) Welsh (n=98) and Finnish (n=56) facial cohorts, v) Lip traits and cleft genes (ALSPAC 15 year olds), vi) Cleft and non-cleft trios (n=597) were reviewed.

ALSPAC 15 year old cohort¹

The first 7 principal components of facial variation highlight the most important facial features in this population.

- i) facial size and height
- ii) width of the face
- iii) nose prominence
- iv) lip and chin protrusion
- v) eye depth
- vi) vertical nose height
- vii) mouth width and depth

Fathers and offspring

The narrow sense heritability ($h^2 > 0.60$) was calculated for the frequency of ratios that involve a landmark (represented by the cubic root of the radius - Figure 1). Progressing through the levels of h^2 it was noticed that asymmetry was less and that the heritability of landmarks associated with the lips and nose were much higher in the daughters compared to sons. The frequency of ratios involving the nasion and mid-endocanthion (n-men) were high for sons and daughters ($h^2 > 0.73$) which would explain the discovery of the *Pax3* gene associated with these landmarks.²

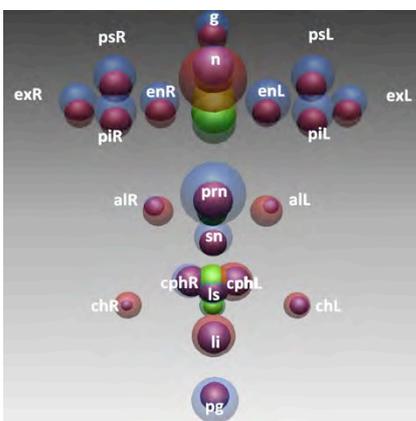


Figure 1 Radius of spheres = Cubic root of the frequency of landmarks of 71,253 possible ratios for 28 landmarks (blue males - red females), including mid points (yellow M and green F) $h^2 = 0.6$.

Twins UK (Visigen)

Looking at the heritability of MZ and DZ twins ($h^2 = 2(r_{MZ} - r_{DZ})$), the principal components were similar for both scaled and unscaled 3D facial shells. These PC's also compared favourably with the 15 year olds. Face size and height were important in the scaled and lip length in both scaled and unscaled.

Facial growth cohorts (Welsh Finnish)

A UK growth cohort verifies the close relationship of hard and soft tissue landmarks. The Welsh and Finnish cohorts highlight the difference in facial proportions in boys and girls from 11 to 17 years of age which may explain dominance of h^2 of daughters over sons in the fathers and offspring study. The lip shape changes very little in males and females from 11 to 17 years of age and growth will not influence the dominance of the fathers features in their daughters.

Lip traits and gene associations

The Wilson-Richmond lip trait scale³ was used to phenotype the 15 year old ALSPAC children (Figure 2). To illustrate this both Mick Jagger and her daughter Georgia May have a groove/notch in the upper lip and a double lower lip border with very similar lip shape. 14 candidate SNP's were associated with V-shaped cupid's bow, narrow philtrum, upturned commissures and deep groove. An automated lip/nose shape recognition is currently being trialled

Cleft and non-cleft trios

The prevalence of lip traits among parents of cleft and non-cleft trios highlighted the predictive value of the lip features. 9 lip traits were statistically significant between case and controls.

Conclusions

- 2½ datasets (fathers/offspring and female twins) confirm relative importance of similar facial features.
- Different patterns of heritability of facial landmarks in male and female offspring – particularly ch, chr & alae (F) and prn, pg (M)
- Male and female facial proportions will change between 12-18 years of age.
- Lips have distinctive features most present at birth can be identified easily on individuals after a period of training.
- These lip traits may be useful in understanding facial development, inheritance and CLP.
- Parental lip trait and gene interaction may increase the risk of CLP
- Automated systems will help in some areas of facial trait recognition.

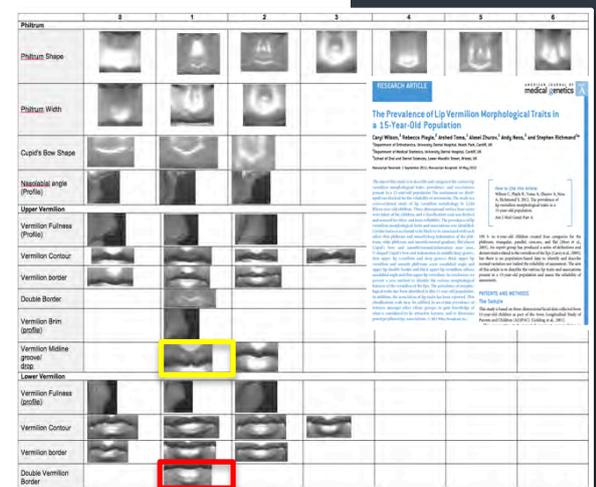


Figure 2 Wilson-Richmond lip scale highlighting notch in upper lip (yellow) and double lower border (red).

References

1. Toma AM, Zhurov AI, Playle R, Marshall D, Rosin PL, Richmond S. The assessment of facial variation in 4747 British school children. *Eur J Orthod.* 2012 Dec;34(6):655-64
2. Paternoster L, Zhurov AI, Toma AM, Kemp JP, St Pourcain B, Timpson NJ, McMahon G, McArdle W, Ring SM, Smith GD, Richmond S, Evans DM. Genome-wide association study of three-dimensional facial morphology identifies a variant in *PAX3* associated with nasion position. *Am J Hum Genet.* 2012 Mar 9;90(3):478-85.
3. Wilson C, Playle R, Toma A, Zhurov A, Ness A, Richmond S. The prevalence of lip vermilion morphological traits in a 15-year-old population. *Am J Med Genet A.* 2013 Jan;161A(1):4-12.

Abbreviations: ALSPAC: Avon Longitudinal Study of Parents and Children. Visigen: consortium dedicated to uncover the genetic foundations of visible traits

Editors:

Stephen Richmond and Peter Claes

Copyright of Cardiff University and KU Leuven

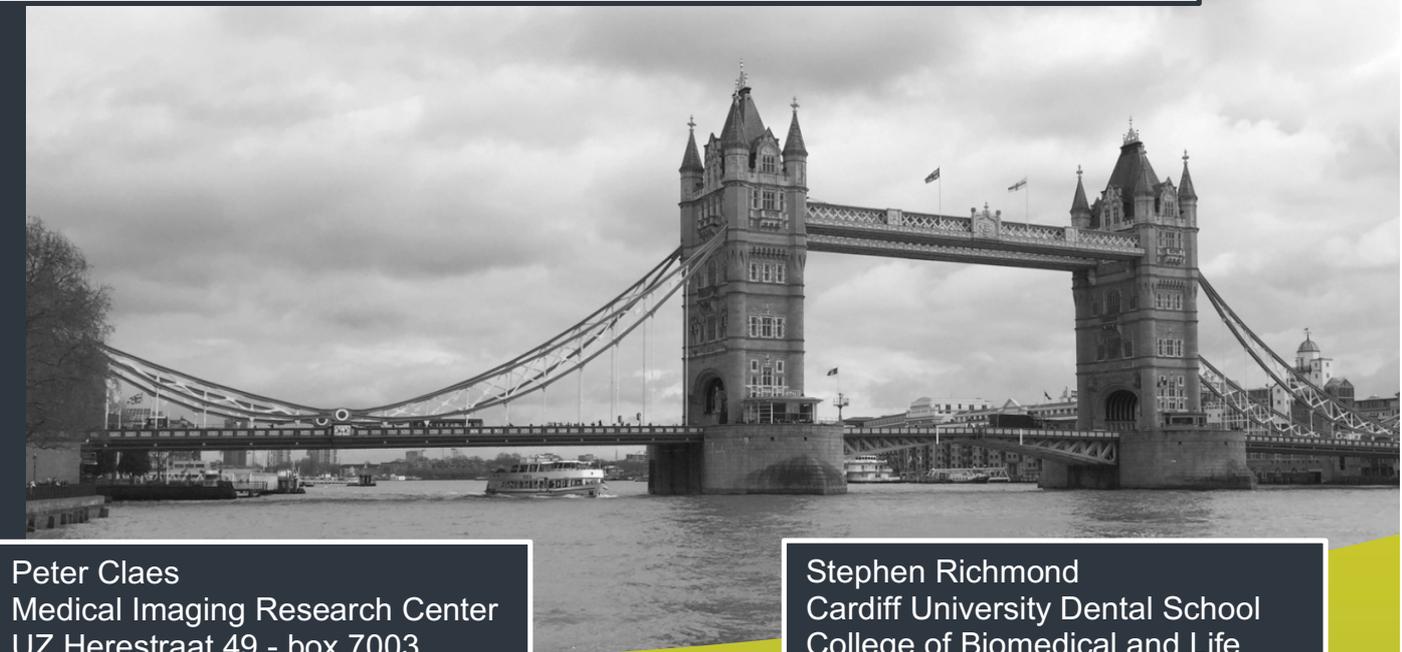
Any material taken from this document must be fully referenced.

“We all expect to meet up again and hopefully those who could not take part this year will be able to make the next meeting”.

Acknowledgements:

The hosts would like to thank everyone for their enthusiastic contributions that made this meeting a success.

We would also like to thank Cardiff University and KU Leuven for their financial support that made the First Facial Genetics Workshop happen.



Peter Claes
Medical Imaging Research Center
UZ Herestraat 49 - box 7003
3000 Leuven
Belgium

peter.claes@kuleuven.be

KU LEUVEN

Stephen Richmond
Cardiff University Dental School
College of Biomedical and Life
Sciences
Heath Park
Cardiff
CF14 4XY
United Kingdom

richmonds@cardiff.ac.uk

CARDIFF
UNIVERSITY
DENTAL SCHOOL