

University of Dundee

Nonsyndromic cleft palate

Ishorst, Nina; Francheschelli, Paola; Böhmer, Anne C.; Khan, Mohammad Faisal J.; Heilmann-Heimbach, Stefanie; Fricker, Nadine

Published in:
Birth Defects Research

DOI:
[10.1002/bdr2.1213](https://doi.org/10.1002/bdr2.1213)

Publication date:
2018

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Ishorst, N., Francheschelli, P., Böhmer, A. C., Khan, M. F. J., Heilmann-Heimbach, S., Fricker, N., Little, J., Steegers-Theunissen, R. P. M., Peterlin, B., Nowak, S., Martini, M., Kruse, T., Dunsche, A., Kreuzsch, T., Gözl, L., Aldhore, K., Halboub, E., Reutter, H., Mossey, P., ... Mangold, E. (2018). Nonsyndromic cleft palate: An association study at GWAS candidate loci in a multiethnic sample. *Birth Defects Research*, 110(10), 871-882. <https://doi.org/10.1002/bdr2.1213>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

This is the pre-peer reviewed version of the following article Nonsyndromic cleft palate : An association study at GWAS candidate loci in a multiethnic sample. / Ishorst, Nina et. al., Birth defects research. which has been published in final form at DOI: 10.1002/bdr2.1213 . This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving

Nonsyndromic cleft palate: An association study at GWAS candidate loci in a multi-ethnic sample

Journal:	<i>Birth Defects Research</i>
Manuscript ID	BDR-17-0234.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	19-Jan-2018
Complete List of Authors:	<p>Ishorst, Nina; University of Bonn, Institute of Human Genetics; Life and Brain Center, Department of Genomics Francheschelli, Paola; University of Ferrara, Department of Biomedical and Specialty Surgical Sciences, Section of Medical Biochemistry, Molecular Biology and Genetics Böhmer, Anne; University of Bonn, Institute of Human Genetics; Department of Genomics, Life and Brain Center, University of Bonn Khan, Mohammad Faisal J.; University of Ferrara, Department of Biomedical and Specialty Surgical Sciences, Section of Medical Biochemistry, Molecular Biology and Genetics Heilmann-Heimbach, Stefanie; University of Bonn, Institute of Human Genetics; Department of Genomics, Life and Brain Center, University of Bonn Fricker, Nadine; University of Bonn, Institute of Human Genetics; Department of Genomics, Life and Brain Center, University of Bonn Little, Julian; Univ Ottawa, Epidemiol & Commun Hlth Steegers-Theunissen, Régine; Erasmus MC, University Medical Center, Obstetrics and Gynecology Peterlin, Borut; University Medical Center Ljubljana , Clinical Institute of Medical Genetics, Department of Obstetrics & Gynecology Nowak, Stefanie; University of Bonn, Institute of Human Genetics Martini, Markus; Rheinische Friedrich Wilhelms Universität Bonn, Department of Oral and Maxillo-Facial-Plastic Surgery Kruse, Teresa; University of Cologne, Department of Orthodontics Dunsche, Anton; Clinics Karlsruhe, Department of Oral and Maxillo-Facial Surgery Kreuzsch, Thomas; Asklepios Klinik Nord, Heidberg, Department of Oral and Maxillofacial Surgery, Head and Neck Centre Gölz, Lina; University of Bonn, Department of Orthodontics Aldharae, Khalid Ahmed; Tamar University, Orthodontic Department, College of Dentistry Halboub, Esam; Jazan University, College of Dentistry, Division of Oral Medicine and Diagnostic Sciences, Department of Maxillofacial Surgery and Diagnostic Sciences Reutter, Heiko; University of Bonn, Institute of Human Genetics; University of Bonn, Neonatology Mossey, Peter; University of Dundee, Dental School Nöthen, Markus; University of Bonn, Institute of Human Genetics; Life and Brain GmbH, Department of Genomics Rubini, Michele; University of Ferrara, Department of Biomedical and Specialty Surgical Sciences Ludwig, Kerstin; Institute of Human Genetics, Department of Genomics</p>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Knapp, Michael; University of Bonn, Institute of Medical Biometry, Informatics, and Epidemiology Mangold, Elisabeth; University of Bonn, Institute of Human Genetics
Key Words:	congenital malformation, common variants, imputed genome-wide association study, nonsyndromic cleft palate only, candidate loci, association study

SCHOLARONE™
Manuscripts

For Peer Review

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 **1 Nonsyndromic cleft palate: An association study at GWAS candidate loci in a multi-ethnic sample**

4
5 2

6
7 **3 Authors:**

8 4 Nina Ishorst ^{1,2*}, Paola Francheselli ^{3*}, Anne C. Böhmer ^{1,2}, Mohammad Faisal J. Khan ³, Stefanie
9 5 Heilmann-Heimbach ^{1,2}, Nadine Fricker ^{1,2}, Julian Little ⁴, Regine P.M. Steegers-Theunissen ⁵, Borut
10 6 Peterlin⁶, Stefanie Nowak ², Markus Martini ⁷, Teresa Kruse ⁸, Anton Dunsche ⁹, Thomas Kreuzsch ¹⁰,
11 7 Lina Gözl ¹¹, Khalid Aldhoreae ¹², Esam Halboub ¹³, Heiko Reutter ^{2,14}, Peter Mossey ¹⁵, Markus M.
12 8 Nöthen ^{1,2}, Michele Rubini ³, Kerstin U. Ludwig ^{1,2}, Michael Knapp ¹⁶, Elisabeth Mangold ²

13
14
15
16 9

17
18 **10 Affiliations:**

19 11 ¹ Department of Genomics, Life&Brain Center, University of Bonn, Bonn, Germany

20 12 ² Institute of Human Genetics, University of Bonn, Bonn, Germany

21 13 ³ Department of Biomedical and Specialty Surgical Sciences, Section of Medical Biochemistry,
22 14 Molecular Biology and Genetics, University of Ferrara, Ferrara, Italy

23 15 ⁴ School of Epidemiology, Public Health and Preventive Medicine, University of Ottawa, Ottawa,
24 16 Ontario, Canada

25 17 ⁵ Department of Obstetrics and Gynaecology, Department of Pediatrics, Division Neonatology
26 18 Erasmus MC, University Medical Center, Rotterdam, The Netherlands

27 19 ⁶ Clinical Institute of Medical Genetics, Department of Obstetrics & Gynecology, University Medical
28 20 Center Ljubljana, Ljubljana, Slovenia

29 21 ⁷ Department of Oral and Maxillo-Facial-Plastic Surgery, University of Bonn, Bonn, Germany

30 22 ⁸ Department of Orthodontics, University of Cologne, Cologne, Germany

31 23 ⁹ Clinics Karlsruhe, Department of Oral and Maxillo-Facial Surgery, Karlsruhe, Germany

32 24 ¹⁰ Department of Oral and Maxillofacial Surgery, Head and Neck Centre, Asklepios Klinik Nord-
33 25 Heidberg, Hamburg , Germany

34 26 ¹¹ Department of Orthodontics, University of Bonn, Bonn, Germany

35 27 ¹² Orthodontic Department, College of Dentistry, Tamar University, Tamar, Yemen

36 28 ¹³ Division of Oral Medicine and Diagnostic Sciences, Department of Maxillofacial Surgery and
37 29 Diagnostic Sciences, College of Dentistry, Jazan University, Jazan, Saudi Arabia

38 30 ¹⁴ Department of Neonatology, Children's Hospital, University of Bonn, Bonn, Germany

39 31 ¹⁵ Dental Hospital, University of Dundee, Dundee, UK

40 32 ¹⁶ Institute of Medical Biometry, Informatics and Epidemiology, University of Bonn, Bonn, Germany

41 33 * contributed equally

42
43
44
45
46
47
48
49
50
51
52
53
54
55
56 34

57 **35 Grant information:** The study was supported by the German Research Foundation (grant FOR 423

Ishorst et al., Candidate loci association study in nsCPO

1 and individual grants MA 2546/3-1, KR 1912/7-1, NO 246/6-1, and WI 1555/5-1, LU-1944/3-1), and
2 by the European Science Foundation (grant 09-RNP-023 (EMRC), ESF Network for Orofacial Clefts
3 Research, Prevention and Treatment: EUROCleftNet)

4
5 **Running title:** Candidate loci association study in nsCPO

6
7 **Correspondence to:**

8 Dr. Elisabeth Mangold, Institute of Human Genetics, University of Bonn, Sigmund-Freud-Straße 25,
9 53127 Bonn, Germany. Phone: +49 (0)228 287 51008; Fax: +49 (0) 228 287 51011 Email:
10 e.mangold@uni-bonn.de

11
12 **Abstract (250 words):**

13 **Background:** Nonsyndromic cleft palate only (nsCPO) is a common and multifactorial form of
14 orofacial clefting. In contrast to successes achieved for the other common form of orofacial clefting,
15 i.e., nonsyndromic cleft lip with/without cleft palate (nsCL/P), genome wide association studies
16 (GWAS) of nsCPO have identified only one genome wide significant locus. Aim of the present study
17 was to investigate whether common variants contribute to nsCPO and, if so, to identify novel risk
18 loci.

19 **Methods:** We genotyped 33 SNPs at 27 candidate loci from 2 previously published nsCPO GWAS in an
20 independent multi-ethnic sample. It included: (i) a family-based sample of European ancestry
21 (n=212); and (ii) two case/control samples of Central European (n=94/339) and Arabian ancestry
22 (n=38/231), respectively. A separate association analysis was performed for each genotyped dataset,
23 and meta-analyses were performed.

24 **Results:** After association analysis and meta-analyses, none of the 33 SNPs showed genome-wide
25 significance. Two variants showed nominally significant association in the imputed GWAS dataset and
26 exhibited a further decrease in P-value in a European and an overall meta-analysis including imputed
27 GWAS data, respectively (rs395572: $P_{\text{MetaEU}}=3.16 \times 10^{-4}$; rs6809420: $P_{\text{MetaAll}}=2.80 \times 10^{-4}$).

28 **Conclusion:** Our findings suggest that there is a limited contribution of common variants to nsCPO.
29 However, the individual effect sizes might be too small for detection of further associations in the
30 present sample sizes. Rare variants may play a more substantial role in nsCPO than in nsCL/P, for
31 which GWAS of smaller sample sizes have identified genome-wide significant loci. Whole-
32 exome/genome sequencing studies of nsCPO are now warranted.

33

Ishorst et al., Candidate loci association study in nsCPO

Keywords: congenital malformation, nonsyndromic cleft palate only, common variants, imputed genome-wide association study, candidate loci, association study

Introduction:

Orofacial clefting is one of the most common forms of congenital malformation, and displays substantial phenotypic variation. The two most common forms of orofacial clefting are cleft lip with/without cleft palate (CL/P) and cleft palate only (CPO) (Mangold et al., 2011). These two types are distinct in terms of both epidemiological findings and developmental origin, and can occur in a syndromic or nonsyndromic form. In the latter, no additional physical or cognitive disabilities are present (Mangold et al., 2011). In Europe, the estimated prevalence of true nonsyndromic clefts is 1 in 1,000 for nonsyndromic CL/P (nsCL/P) and 1 in 2,400 for nonsyndromic CPO (nsCPO), suggesting that nsCL/P is the more common form of nonsyndromic clefting in this population (Mangold et al., 2011). Research suggests that nsCL/P and nsCPO are multifactorial disorders, with both genetic and environmental factors contributing to their development. Heritability estimates of approximately 90% have been reported for both phenotypes (Grosen et al., 2012).

To date, research has identified a total of 39 genome-wide significant risk loci for nsCL/P across different populations. Of these, 37 loci were identified via genome-wide association studies (GWAS), meta-analyses of GWAS data, or follow-up studies (Rahimov et al., 2008; Birnbaum et al., 2009; Moreno et al., 2009; Beaty et al., 2010; Mangold et al., 2010; Ludwig et al., 2012; Beaty et al., 2013; Sun et al., 2015; Leslie et al., 2016b; Ludwig et al., 2016; Yu et al., 2017; Leslie et al., 2017; Ludwig et al., 2017).

In contrast, only two independent GWAS (Beaty et al., 2011; Leslie et al., 2016a), and one meta-analysis of both studies (Leslie et al., 2017), have been published for nsCPO. One of these studies (Leslie et al., 2016a) identified the first - and as yet only - genome-wide significant risk locus for nsCPO at chromosome 1p36, a finding that was then independently detected in a sequencing study (Mangold et al., 2016). Notably, this missense variant in the *grainyhead-like 3* gene (*GRHL3*) showed no association with nsCL/P (Mangold et al., 2016).

The issue of whether nsCPO and nsCL/P share a common genetic etiology is a matter of ongoing debate. However, available epidemiological- and molecular data suggest that the shared genetic etiology of these two frequent forms of clefting is limited. Ludwig et al. (2017) demonstrated that with the exception of the *FOXE1* locus (9q22), none of the nsCL/P risk loci known at the time of the analysis exhibited a significant association in an imputed nsCPO GWAS dataset (Ludwig et al., 2017). The finding for *FOXE1* was unsurprising, since previous research has established that this locus is associated with both traits (Moreno et al., 2009; Ludwig et al., 2014; Leslie et al., 2017). A polygenic score approach in the Ludwig et al. (2017) study also showed that nsCL/P loci in sum do not

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 contribute to nsCPO etiology (Ludwig et al., 2017). These findings are consistent with the results of a
4 2 current study of Moreno-Urbe et al. (2017).
5
6 3 Aim of the present study was to investigate whether common variants contribute to nsCPO and, if so,
7 4 to identify novel risk loci. To that end we investigated genetic association of candidate loci derived
8 5 from the already mentioned imputed nsCPO GWAS dataset (Ludwig et al., 2017). Additionally, we
9 6 included candidate single nucleotide polymorphisms (SNPs) described in another recent imputed
10 7 nsCPO GWAS (Leslie et al., 2016a). Association analyses were performed in a European trio sample,
11 8 and two independent case/control samples from Central Europe and Yemen, respectively.
12 9 Additionally, we investigated a possible overlap in signals between nsCL/P in nsCPO by checking 15
13 10 novel nsCL/P risk loci (Yu et al., 2017; Leslie et al., 2017) for association in the imputed nsCPO dataset
14 11 (Ludwig et al., 2017). With this investigation we completed the analysis performed by Ludwig et al.
15 12 (2017) for the 24 risk loci known at that time.
16 13
17 14

14 **Methods:**

15 **Imputed GWAS Dataset (Ludwig et al., 2017)**

16 16 The imputed nsCPO GWAS data that were followed up in the present study were based on a GWAS
17 17 that comprised approximately 498,000 genotypes of each member of 550 nuclear nsCPO trios (Beaty
18 18 et al., 2011). Ludwig et al. (2017) obtained dbGaP data of this nsCPO GWAS and performed genome-
19 19 wide imputation for 446 trios of European (n=212) and Asian (n=234) ancestry using IMPUTE2 on the
20 20 basis of 1000 genomes haplotypes (phase 3) (The 1000 Genomes Project Consortium 2015).
21 21 Statistical analyses were performed for around 8.38 million variants with an info-score of >0.4 and a
22 22 minor allele frequency (MAF) of > 1% using the FBAT dosage approach (Cobat et al., 2014). In
23 23 addition to the main imputed dataset of genotypes from 446 trios (EU+Asia_{Beaty}), a sub-dataset was
24 24 generated comprising the genotypes of 212 European trios only (EU_{Beaty}).
25 25

26 **Sample**

27 27 The genotyped sample comprised two independent case/control samples and one trio sample
28 28 recruited in the context of different studies. Further information on gender distribution and origin is
29 29 given in Table S1.
30 30 The first case/control sample comprised 94 nsCPO index patients and 339 population-matched
31 31 controls (EU_{Mangold}). The cases were drawn from a large sample of nonsyndromic cleft families of
32 32 Central European ancestry (recruitment method described in Mangold et al., 2010). The population-
33 33 matched controls were volunteer blood donors also of Central European ancestry (Mangold et al.,
34 34 2010).
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Ishorst et al., Candidate loci association study in nsCPO

1 The second case/control sample (Arab_{Aldhorae}) was recruited within the context of surgical outreach
2 programs, as described elsewhere (Aldhorae et al., 2014). The Arab_{Aldhorae} sample comprised 38
3 nsCPO cases and 231 blood donors of Arab ethnicity from Yemen.
4 The family-based sample comprised 212 nuclear trios with an nsCPO index patient (EU_{Mossey}). These
5 families were recruited from different areas of Europe (Great Britain, Italy, The Netherlands,
6 Slovenia, Slovakia, Hungary, Estonia, and Bulgaria) in the context of the EUROCRAN/ITALCLEFT study
7 (Mossey et al., 2017). Nonsyndromic status of affecteds was confirmed clinically at the point of
8 recruitment to the various studies.
9 The present study was approved by the ethics committees of the participating institutions, and
10 informed consent was obtained from all participants, or if not adult by parents or legal guardians,
11 prior to inclusion.

12 13 Genotyping

14 DNA from the EU_{Mangold} sample (patients and controls) had been isolated using standard salting out
15 procedures (Mangold et al., 2010). Isolation of DNA from Arab_{Aldhorae} blood samples (patients and
16 controls) had been performed using the Chemagic Magnetic Separation Module I (Perkin Elmer
17 chemagen Technologies GmbH, Baesweiler, Germany), in accordance with the manufacturer's
18 instructions (Aldhorae et al., 2014). Genomic DNA from the EU_{Mossey} sample had been extracted from
19 peripheral venous blood samples using the QIAamp Blood DNA Mini kit (QIAGEN, Germantown, MD,
20 US) (Mossey et al., 2017).
21 DNA was quantified using NanoDrop (PqLab Thermo Scientific, Wilmington, DE, US). Genotyping of
22 selected SNPs was performed with the multiplex MassARRAY system (Agena Bioscience, Hamburg,
23 Germany), as based on end-point PCR coupled with MALDI-ToF. Two plexes were designed using the
24 Assay Design Suite v2.0, and data analysis was performed using the Typer Analysis software (both:
25 Agena, Hamburg, Germany). Primers were synthesized by Metabion (Planegg, Germany). Primer
26 sequences are available upon request. One SNP was genotyped in both plexes as an internal control.

27 28 Selection of SNPs for Genotyping

29 SNPs with P-values of $< 1 \times 10^{-5}$ in at least one of the GWAS datasets (EU+Asia_{Beaty}, EU_{Beaty}) were
30 selected from the imputed GWAS data (Table S2). Of these, SNPs with an info-score < 0.8 were
31 excluded from further analysis.
32 For loci in which more than 1 SNP fulfilled the selection criteria we selected suitable backup SNPs.
33 SNPs that failed assay design were replaced by the second best hit, in terms of p-value and info-
34 score, at the respective locus. If only one SNP fulfilling the criteria was available at a given locus and
35 failed assay design, a backup SNP in high linkage disequilibrium (LD) ($r^2 > 0.99$) was selected from the

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 imputed dataset. In this special case also variants with a P-value slightly above the set threshold of
4 2 1×10^{-5} were considered. The selection procedure resulted in 25 SNPs at 19 loci (Table S2; Table S3).
5 3 Initially, 13 candidate SNPs (P-value $< 1 \times 10^{-5}$) derived from Leslie et al. (2016a) were to be forwarded
6 4 for genotyping in the present three independent samples, without considering their association in
7 5 the imputed datasets (EU_{Beaty} and EU+Asia_{Beaty}). However, only eight SNPs at eight loci could be
8 6 included into the genotyping assay (Table S4). Thus, a total of 33 SNPs at 27 loci were selected for
9 7 genotyping (Table S5).

10 Statistical Analysis

11 For the case/control samples, the standard chi-square test was used to test for deviation from Hardy-
12 Weinberg Equilibrium (HWE). The threshold for significant deviation from HWE was set at $P < 1 \times 10^{-4}$
13 in controls and $P < 1 \times 10^{-6}$ in cases.

14 Using the two-sided Armitage trend test, each SNP was tested for a difference in genotype
15 distribution between cases and controls. In the EU_{Mossey} trios, single-marker association analysis was
16 performed using the transmission disequilibrium test (TDT), as implemented in the FAMHAP
17 software package (Becker and Knapp, 2004).

18 Data evaluation comprised: (i) a separate analysis of each genotyped sample (EU_{Mangold}, Arab_{Aldhorae},
19 EU_{Mossey}); (ii) a z-score based meta-analysis and a fixed-effects meta-analysis of all genotyped
20 samples (Willer et al., 2010) (Meta_{geno}); (iii) a z-score based meta-analysis (Willer et al., 2010) of all
21 European samples (EU_{Mangold}, EU_{Mossey}, EU_{Beaty}; named Meta_{EU}); and (iv) z-score based meta-analysis
22 (Willer et al., 2010) of all samples (EU_{Mangold}, Arab_{Aldhorae}, EU_{Mossey}, EU+Asia_{Beaty}; termed Meta_{All}).

23 Relative risks were calculated for each of the genotyped samples and their meta-analysis based on
24 the fixed effects approach (Table S6). EU_{Beaty} and EU+Asia_{Beaty} (Ludwig et al., 2017) were imputed
25 using the FBAT dosage approach (Cobat et al., 2014). This method does not allow for calculation of
26 relative risks, thus this data is not available for Meta_{EU} and Meta_{All}.

27 A workflow of the present study is shown in Figure 1.

28 P-values of < 0.05 were considered nominally significant.

29 P-values given as $P_{corrected}$ were Bonferroni corrected for multiple testing. In the EU_{Mangold} and EU_{Mossey}
30 data the number of 33 genotyped SNPs was considered for Bonferroni correction (Table S5). In the
31 Arab_{Aldhorae} dataset the number of 32 SNPs was considered (Table S5). For the analysis of 15 nsCL/P
32 loci top SNPs from Yu et al. (2017) and Leslie et al. (2017) in the imputed nsCPO dataset we
33 considered the number of 15 markers for Bonferroni correction (Table S7). P-values of < 0.05 were
34 considered significant after correction for multiple testing.

35 LD values were determined using the web-based SNAP tool (Broad Institute).

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 Quality Control (QC) in the Genotyping Step

4 2 For the internal control SNP used in both plexes, genotypes were compared and checked for
5 3 agreement. Duplicates were identified by comparing individuals with shared genotypes, and
6 4 removed from the genotyping dataset. In addition, samples with >2 missing genotypes
7 5 (corresponding to an average call rate of 94%) and/or >2 Mendelian inconsistencies in the trio
8 6 sample, were excluded.
9 7

10 8 Analysis of nsCL/P loci top SNPs and associated regions of Yu et al. (2017) and Leslie et al. (2017) in
11 9 EU_{Beaty} and EU+Asia_{Beaty}

12 10 To complete the analysis of potential signal-overlap for nsCL/P and nsCPO performed in Ludwig et al.
13 11 (2017), the imputed nsCPO dataset was inspected on both a locus-wide and lead SNP level for the
14 12 novel loci identified in 2017 by Yu et al. (n= 14) and Leslie et al. (n= 1).
15 13

16 14 **Results:**

17 15 Imputation resulted in 8,384,634 and 8,384,636 markers (info-score ≥ 0.4 ; MAF > 1%) in EU_{Beaty} and
18 16 EU+Asia_{Beaty}, respectively (Ludwig et al., 2017). No genome-wide significant association for nsCPO
19 17 was reported (Ludwig et al., 2017). However, 83 SNPs at 26 loci showed P-values of $< 10^{-5}$ in at least
20 18 one analysis. The following had info-scores of < 0.8 and were thus excluded: (i) 6 SNPs at 6 loci with
21 19 only one associated SNP; and (ii) one SNP at a locus with multiple SNPs. This resulted in 76 SNPs at 20
22 20 loci. Of these 20 loci, one locus could not be included in the genotyping step due to technical issues.
23 21 For the remaining 19 loci, the selection procedure resulted in 25 SNPs for genotyping (Table S2). The
24 22 risk alleles of SNPs selected from the imputed data were identical in both datasets (EU_{Beaty} and
25 23 EU+Asia_{Beaty}). None of the eight SNPs at eight loci selected from Leslie et al. (2016a) showed
26 24 nominally significant association in either of the imputed datasets: The lowest P-values were
27 25 obtained for rs3740617 in EU_{Beaty} (P = 0.417) and rs595533 in EU+Asia_{Beaty} (P = 0.087) (Table S4).
28 26 Genotyping was successful for all 33 markers (representing 27 different loci). In the course of QC,
29 27 several samples were excluded due to missing genotype calls for more than two SNPs. These
30 28 comprised the samples of: 7 individuals from EU_{Mangold} (2 patients and 5 controls); 2 individuals from
31 29 Arab_{Aldhorae} (1 patient, 1 control); and 34 individuals from 27 EU_{Mossey} trios. Furthermore, the analyses
32 30 identified six EU_{Mossey} trios with more than two Mendelian inconsistencies, and two duplicates (1
33 31 control from Arab_{Aldhorae} and 1 EU_{Mossey} trio). These samples were excluded from further analysis. One
34 32 SNP (rs76706433) was monomorphic in the Yemeni population (Arab_{Aldhorae}). This variant was
35 33 therefore excluded from the analyses of Arab_{Aldhorae}, resulting in a set of 32 SNPs for this particular
36 34 sample. None of the analyzed SNPs showed a deviation from HWE at the selected thresholds.
37 35

Ishorst et al., Candidate loci association study in nsCPO

1 Genotyped sample and Meta_{geno}

2 Separate analysis of each genotyped sample (EU_{Mangold}, Arab_{Aldhorae}, EU_{Mossey}) (Figure 1) revealed four
3 SNPs (rs6747560, rs6809420, rs6805813, rs395572) with nominally significant P-values ($P < 0.05$).
4 These comprised one SNP in the EU_{Mangold} sample and three SNPs in the Arab_{Aldhorae} sample (Table 1).
5 At three of these loci, the risk allele was not identical to the risk allele of the imputed GWAS dataset.
6 In the Arab_{Aldhorae} sample, rs6809420 was the only SNP to show both nominally significant association
7 and a consistent effect direction ($P = 0.024$). This result did not withstand correction for multiple
8 testing of the 32 SNPs ($P_{corrected} = 0.752$). The fixed effects-based analysis revealed a relative risk (RR)
9 of 2.41(95% confidence interval (CI), 1.12-5.20) for this SNP (Table S6).
10 The z-score-based meta-analysis of all three genotyped samples (Meta_{geno}) revealed two nominally
11 significant SNPs (rs6747560, rs1283107). Both SNPs showed the opposite effect direction to that
12 observed in the initial datasets (EU_{Beaty} and EU+Asia_{Beaty}) (Table 1, Table S8). Notably, in Meta_{geno}
13 rs6809420 showed a tendency towards a nominally significant association ($P_{Meta\text{geno}} = 0.066$) and the
14 fixed-effects-based meta-analysis revealed an RR of 1.24 (95% CI, 0.97-1.60) (Table S6).

16 Meta_{EU}

17 In the meta-analysis of all samples with European ancestry (EU_{Mangold}, EU_{Mossey}, EU_{Beaty}) (Figure 1), a
18 decrease in the association P-value in comparison to EU_{Beaty} alone was observed for the SNPs
19 rs395572, rs7732608, and rs3740617 (Table 2). Of these, only rs395572 at 3q29 showed a nominally
20 significant association ($P_{EU\text{Beaty}} = 0.0017$; $P_{Meta\text{EU}} = 3.16 \times 10^{-4}$) (Figure 2). None of the SNPs from Leslie
21 et al. (2016a) showed decreased association P-values in the meta-analysis of all European samples.

23 Meta_{All}

24 In the z-score-based meta-analysis of all trios from the imputed EU+Asia_{Beaty} dataset and all
25 genotyped samples (EU_{Mangold}, Arab_{Aldhorae}, EU_{Mossey}) (Figure 1), lower P-values compared to
26 EU+Asia_{Beaty} alone were observed for the SNPs rs6809420, rs7732608, rs9347594, rs3740617, and
27 rs595533 (Table 3). Of these, only two SNPs showed nominally significant association. Rs6809420
28 was nominally significant in the initial imputed GWAS data (EU+Asia_{Beaty}), and the association signal
29 got stronger after combination with the genotyped samples ($P_{EU+Asia\text{Beaty}} = 0.0011$, $P_{Meta\text{All}} = 2.80 \times 10^{-4}$)
30 (Figure 3). Here, Arab_{Aldhorae} (Table 3) made the main contribution in terms of lowest P-value from the
31 genotyped samples as already reported in section 'Genotyped Sample and Meta_{geno}'.
32 The association P-value for rs595533 from Leslie et al. (2016a) became nominally significant following
33 data combination ($P_{EU+Asia\text{Beaty}} = 0.087$, $P_{Meta\text{All}} = 0.0412$) (Table 3).

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 In comparison to $Meta_{EU}$, the association P-value of rs395577 showed no further decrease in $Meta_{All}$.
4 2 This is probably due to an inconsistent effect direction for this SNP in the Yemen sample (Table S9,
5 3 Table S10).
6 4

7
8
9 5 Analysis of nsCL/P loci top SNPs and associated regions of Yu et al. (2017) and Leslie et al. (2017) in
10 6 EU_{Beaty} and $EU+Asia_{Beaty}$

11 7 Analysis of these 15 novel loci revealed no genome wide significant association at either the whole
12 8 locus or lead SNP level approach. In the locus-wide approach, the lowest P values were obtained for
13 9 rs73288895 and rs8009664 from Yu et al. (2017) in EU_{Beaty} and $EU+Asia_{Beaty}$, respectively ($P_{EU_{Beaty}} =$
14 10 2.87×10^{-4} and $P_{EU+Asia_{Beaty}} = 3.88 \times 10^{-5}$) (Table S11; Table S12). Both variants are located at 14q22.1.
15 11 According to a SNAP proxy search (Johnson et al., 2008), neither variant is in LD with the lead SNP
16 12 rs7148069 at the 14q22.1 locus (Yu et al., 2017) in European or Asian reference populations. In the
17 13 lead SNP approach, rs705704 at 12q13.2 showed a nominally significant P-value in $EU+Asia_{Beaty}$
18 14 ($P_{EU+Asia_{Beaty}} = 0.010$) (Table S7). However, this did not withstand Bonferroni correction for multiple
19 15 testing of the 15 tested lead SNPs ($P_{corrected} = 0.155$).
20 16

21 17 **Discussion:**

22 18 To date, research has identified only one genome-wide significant nsCPO risk variant, namely
23 19 rs41268753 at the *GRHL3* locus (1p36) (Leslie et al., 2016a; Mangold et al., 2016). The aim of the
24 20 present study was to further elucidate the genetic background of nsCPO, in particular whether
25 21 common variants do contribute to the phenotype.

26 22 We tried to detect such common nsCPO risk loci by combining imputed nsCPO GWAS data (Ludwig et
27 23 al., 2017) with genotyping data from independent samples in two different ways:

28 24
29 25 In the first approach data of all individuals under study were combined, regardless of ethnicity. This
30 26 strategy aimed at compensating for small sample sizes and is expected to identify variants that
31 27 contribute to nsCPO risk in diverse populations, albeit with low effect sizes. This led to the
32 28 identification of a lead variant at 3q13 (rs6809420) with a frequency of 81.9% in the general
33 29 population (Non-Finnish Europeans; gnomAD) (Exome Aggregation Consortium, 2016), which is
34 30 located in an intergenic region. The genes *TUSC1* and *IGSF11* are located 1 Mb upstream and 1.2 Mb
35 31 downstream of rs6809420, respectively, which suggests that this variant may have a regulatory
36 32 function (Figure 3). Further support for the association with rs6809420 is provided by the variants
37 33 rs9829549 ($P_{EU+Asia_{Beaty}} = 0.0013$; $r^2 = 1$) and rs4602427 ($P_{EU+Asia_{Beaty}} = 7.50 \times 10^{-4}$, $r^2 = 0.94$) which lie in
38 34 perfect/high LD with the lead variant (red dots, Figure 3).
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 Evidence for a potential regulatory function of genetic regions can be provided by publicly available
4 2 databases like the Roadmap Epigenomics Consortium (2015). These data show that rs6809420 is
5 3 located in an enhancer mark (H3K4me1) of CD56+ ectoderm cultured cells derived from human
6 4 embryonic stem cells, although it should be noted that these cells are derived from ectoderm, and
7 5 are not identical to the palate-forming neural crest cells. Nevertheless, due to the early
8 6 embryological time-point of facial development, access to appropriate human tissue samples for the
9 7 epigenetic or functional follow-up of non-coding regions is problematic (Thieme and Ludwig, 2017).
10 8 Furthermore, ENCODE transcription factor binding experiments have predicted that rs6809420 leads
11 9 to the alteration of a p300 binding site (Kheradpour and Kellis, 2014). Together with other
12 10 translational coactivators, p300 has been shown to exhibit distinct expression patterns during murine
13 11 craniofacial development (Bhattacharjee et al., 2009). An altered p300 binding site may thus lead to
14 12 the dysregulation of genes essential to this critical developmental process. Through an interaction
15 13 with the effects of other variants and/or environmental factors, this may lead to the development of
16 14 a cleft palate. Altogether, rs6809420 at 3q13 is an interesting finding. Another variant identified by
17 15 this multi-ethnic approach was rs595533 at 18q22.1. This variant was reported by Leslie et al.
18 16 (2016a). Only a limited decrease in the P-value was observed following data combination. We do not
19 17 consider rs595533 an interesting suggestive finding of the present study.

20 18
21 19 The second approach was done in view of the predominantly European nature of the samples under
22 20 study. In this 'European only' approach a meta-analysis of all European samples ($Meta_{EU}$) was
23 21 performed. This ethnicity-specific meta-analysis identified a second suggestive variant (rs395572)
24 22 with a frequency of 67.4% in the general population (Non-Finnish Europeans; gnomAD) (Exome
25 23 Aggregation Consortium, 2016). Rs395572 is located upstream of the microRNA gene *MIR4797* and of
26 24 the gene *DLG1* (Figure 2). Recent studies in humans and animal models showed that microRNAs play
27 25 a role in embryonic development including craniofacial development and might contribute to the risk
28 26 of orofacial clefting when disrupted in their function (Wang et al., 2013; Ding et al., 2016; Li et al.,
29 27 2016). Rs395572 may alter expression of *MIR4797*. Alternatively, rs395572 might be in LD with a true
30 28 causal variant located within the microRNA gene, altering binding properties of miR-4797. *DLG1*
31 29 encodes the *human discs large protein 1*, which belongs to the family of molecular scaffolding
32 30 proteins, and is implicated in biological processes such as the regulation of epithelial cell polarity and
33 31 cell migration (Roberts et al., 2012; Marziali et al., 2015). Moreover, a study in mice demonstrated
34 32 that murine *Dlg1* is required for craniofacial development, and leads to craniofacial abnormalities
35 33 including cleft palate when disrupted (Caruana and Bernstein 2001). The haplotype spans nearly the
36 34 whole *DLG1* gene providing additional evidence for *DLG1* as candidate gene (Figure 2). Of note,
37 35 strong support for our finding comes from a recent GWAS replication study for the other frequent

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 cleft phenotype, nsCL/P (Mostowska et al., 2017), which identified a genome-wide significant signal
4 2 at the *DLG1* locus. In their study the strongest associated SNP ($P = 1.77 \times 10^{-8}$) at the locus is rs338222,
5 3 which is in LD with our candidate SNP rs395572 ($r^2 > 0.8$). Our candidate SNP rs395572 reached a P-
6 4 value of 4.54×10^{-7} in their study. However, the risk alleles at rs338222 and rs395572 reported by
7 5 Mostowska et al. (2017) differ from the risk alleles in the present nsCPO data, indicating that one
8 6 allele confers risk for nsCL/P while the other allele confers risk for nsCPO. A similar finding has been
9 7 recently reported in a population-based study of orofacial clefting (Moreno Uribe et al., 2017).
10 8 Moreno Uribe et al. (2017) found that the variant rs227731 at the *NOG1* locus showed an association
11 9 with the opposite risk allele for nsCL/P compared to nsCPO. The functional significance of this
12 10 observation is unclear. One plausible hypothesis is that the binding of transcription factors to these
13 11 regions is allele- and context-specific.

14
15
16 13 Our results support the concept of differences in the genetic background of nsCPO and nsCL/P. First,
17 14 the 15 novel risk loci for nsCL/P (Yu et al., 2017; Leslie et al., 2017) showed no genome-wide
18 15 association in the present nsCPO GWAS data. Second, our present study identified only suggestive
19 16 loci, and nsCPO GWASs have identified only one genome-wide significant finding, namely rs41268753
20 17 at the *GRHL3* locus (Beaty et al., 2011; Leslie et al., 2016a). In contrast, GWASs of nsCL/P even when
21 18 investigating smaller sample sizes identified five genome-wide significant risk loci (Birnbaum et al.,
22 19 2009; Beaty et al., 2010; Mangold et al., 2010), and many loci approaching genome-wide significance
23 20 which were confirmed in further studies (Ludwig et al., 2012; Beaty et al., 2013; Sun et al., 2015;
24 21 Leslie et al., 2016b). Sufficient evidence is therefore now available to consider nsCL/P a complex trait.
25 22 In contrast, the precise etiological nature of nsCPO remains unclear.

26
27
28 24 The development of nsCPO may involve fewer genetic factors and more exogenic factors than is the
29 25 case for nsCL/P. Nonetheless, heritability estimates of approximately 90% and a 50-60 fold increase
30 26 in the recurrence risk for siblings of nsCPO patients compared to the population - data which strongly
31 27 resemble estimates for nsCL/P - indicate that genetic factors make a strong contribution to the
32 28 nsCPO phenotype.

33
34
35 30 One could raise the hypothesis that the genetic variants involved in nsCPO and nsCL/P differ in terms
36 31 of their frequency spectrum, and that nsCPO risk variants have lower frequencies. Taking this into
37 32 consideration common nsCPO risk variants may have been missed with the previous GWASs and also
38 33 with our approach simply due to sample size, i.e. power issues. Of note, the proven risk variant
39 34 rs41268753 at the *GRHL3* locus has a frequency of 3.2% in the general population (Non-Finnish

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 Europeans; gnomAD) (Exome Aggregation Consortium, 2016), and thus is among the low-frequency
4 2 variants in the European population.
5
6 3

7 4 Another hypothesis is that in general, rare variants might play a primary role in nsCPO etiology.
8 5 These rare variants may be located in genes associated with syndromic forms of orofacial clefting, as
9 6 shown previously by the present authors for *GRHL3* and the Van-der-Woude syndrome (Mangold et
10 7 al., 2016). Van der Woude syndrome is characterized by orofacial clefting and lip pits. In the
11 8 mentioned sequencing study by our group members of the Mangold_{EU} nsCPO sample were found to
12 9 have Van der Woude syndrome without lip pits. Furthermore, an exome sequencing study of nsCPO
13 10 identified putative causal variants in *GRHL3* and *CREBBP* (Hoebel et al., 2017). *CREBBP* is associated
14 11 with Rubinstein-Taybi syndrome, which is characterized among others by CPO (Hennekam et al.,
15 12 1993), and which might represent a further hypomorphic syndrome with reduced penetrance of
16 13 additional anomalies. The present and other nsCPO samples may include similar cases. To generate
17 14 more evidence for this hypothesis, further whole-exome or whole-genome sequencing would be a
18 15 useful tool to detect rare variants in nsCPO patients.
19 16

20 17 All of our findings require follow-up in order to assess their true contribution to nsCPO risk. This
21 18 might involve meta-analyses of the present data and other available nsCPO GWAS datasets, or the
22 19 use of larger samples in order to maximize statistical power. However, recruitment is hampered by
23 20 the low prevalence of nsCPO (around 1 in 2,400). An alternative strategy would be to maximize
24 21 statistical power by increasing marker density and data quality. As the newest generation of
25 22 haplotype reference panels for imputation – such as the Haplotype Reference Consortium (HRC)
26 23 (McCarthy et al., 2016) – combine sequencing data from multiple samples, genotype imputation is
27 24 becoming more accurate for the entire MAF spectrum. In the present study, imputation was
28 25 performed using the 1000 genomes haplotypes (phase 3). Thus imputation accuracy (and marker
29 26 density) in general could be increased by imputing the dataset using HRC as a reference panel.
30 27 Furthermore, variants with MAF < 1% could be considered. Nevertheless, the sample size of the
31 28 imputed GWAS sample (EU+Asia_{BeaTy}) was limited, and may have been insufficient in terms of the
32 29 discovery of associations with low frequency variants.
33 30

34 31 There are several limitations to the study.

35 32 First, in our study we may have overlooked true associated variants in the imputed datasets (EU_{BeaTy}
36 33 and EU+Asia_{BeaTy}) (Ludwig et al., 2017) due to the stringent selection criteria applied (only SNPs with P
37 34 < 1x10⁻⁵ were considered for genotyping). An example for this is again rs41268753 at the *GRHL3*
38 35 locus, the only proven nsCPO risk variant known to date (Leslie et al., 2016a; Mangold et al., 2016).

Ishorst et al., Candidate loci association study in nsCPO

1 We checked the P-value of this variant in both imputed datasets. Of note, the P-value of rs4126875
2 is higher than 1×10^{-5} ($P_{\text{EUBeaty}} = 1.18 \times 10^{-4}$; $P_{\text{EU+AsiaBeaty}} = 1.42 \times 10^{-4}$) (data not shown). This is above our
3 set P-value threshold in the SNP selection procedure and demonstrates that our approach can miss a
4 true finding.

5
6 Second, although CPO often occurs as part of a syndrome, it may be misdiagnosed as nonsyndromic
7 when reduced penetrance of additional symptoms occurs. Therefore, the present GWAS and
8 genotyped samples may have included syndromic cases with a monogenetic background. If so, this
9 would have led to a loss of statistical power. However, Moreno Uribe et al. compared locus effects in
10 nonsyndromic (isolated) and syndromic cases, and showed that for most loci, the associations were
11 similar in terms of direction and magnitude (Moreno Uribe et al., 2017). Notably, the Moreno Uribe
12 analysis was performed predominantly in nsCL/P, since only one nsCPO risk locus was analyzed.

13
14 Third, the study is based on imputed GWAS data of trios. Therefore not all of the meta-analyses
15 provided effect sizes, which hampers the interpretation of results.

16
17 A fourth limitation are the sizes and ethnicities of the samples used in this study. Sample sizes might
18 be too small to detect true associations. Asian trios were part of the dataset that led to candidate
19 SNPs, but none of the genotyped samples was of Asian ethnicity. Variants with a population-specific
20 effect restricted to Asians may have been missed by our approach.

21
22 In conclusion, our findings suggest that there is a limited contribution of common variants to nsCPO.
23 Rare variants may play a more substantial role in nsCPO etiology than in nsCL/P. Whole-
24 exome/genome sequencing studies are necessary to identify such rare causative variants.
25 However, our study identified two interesting suggestive nsCPO risk variants. The first finding
26 rs395572 is located at a recently identified risk locus for nsCL/P at 3q29 and lies in high LD with the
27 genome-wide significant top SNP at the locus (Mostowska et al., 2017). Although it has to be noted
28 that the SNP shows the opposite risk allele in our study. The other finding rs6809420 at 3q13 is
29 located in a potential p300 binding site, a coactivator that is known to be involved in craniofacial
30 development. Follow-up studies with increased sample sizes could further elucidate whether these
31 variants are truly associated with nsCPO.

32 **Acknowledgements:**

33
34 The authors thank the patients and their families, and all control individuals, for their participation in
35 the study. We acknowledge the valuable assistance of all clinical and laboratory staff. We thank the

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 German support group for individuals with cleft lip and/or palate (Wolfgang Rosenthal Gesellschaft),
4 2 and the Network for Orofacial Clefts Research, Prevention, and Treatment (EUROcleftNet) for
5 3 assistance with recruitment. Furthermore, the authors would like to thank the Genome Aggregation
6 4 Database (gnomAD) and the groups that provided exome and genome variant data to this resource.
7 5 A full list of contributing groups can be found at <http://gnomad.broadinstitute.org/about>. This work
8 6 was supported by the European Science Foundation.
9
10
11
12
13
14

14 8 **Conflict of interest:**

15 9 The authors have no conflict of interest.
16
17
18

18 11 **References:**

- 19
20 12 Aldhorae KA, Böhmer AC, Ludwig KU, Esmail AHA, Al-Hebshi NN, Lippke B, Gözl L, Nöthen MM,
21 13 Daratsianos N, Knapp M, Jäger A, Mangold E. 2014. Nonsyndromic cleft lip with or without cleft
22 14 palate in Arab populations: Genetic analysis of 15 risk loci in a novel case-control sample recruited in
23 15 Yemen. *Birth Defects Res Part A - Clin Mol Teratol* 100:307–313.
- 24
25 16 Beaty TH, Murray JC, Marazita ML, Munger RG, Ruczinski I, Hetmanski JB, Liang KY, Wu T, Murray T,
26 17 Fallin MD, Redett RA, Raymond G, Schwender H, Jin S, Cooper ME, Dunnwald M, Mansilla MA, Leslie
27 18 E, Bullard S, Lidral AC, Moreno LM, Menezes R, Vieira AR, Petrin A, Wilcox AJ, Lie RT, Jabs EW, Wu-
28 19 Chou YH, Chen PK, Wang H, Ye X, Huang S, Yeow V, Chong SS, Jee SH, Shi B, Christensen K, Melbye M,
29 20 Doheny KF, Pugh EW, Ling H, Castilla EE, Czeizel AE, Ma L, Field LL, Brody L, Pangilinan F, Mills JL,
30 21 Molloy AM, Kirke PN, Scott JM, Arcos-Burgos M, Scott AF. 2010. A genome-wide association study of
31 22 cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. *Nat Genet*
32 23 42:525–9.
- 33
34 24 Beaty TH, Ruczinski I, Murray JC, Marazita ML, Munger RG, Hetmanski JB, Murray T, Redett RJ, Fallin
35 25 MD, Liang KY, Wu T, Patel PJ, Jin S, Zhang TX, Schwender H, Wu-Chou YH, Chen PK, Chong SS, Cheah
36 26 F, Yeow V, Ye X, Wang H, Huang S, Jabs EW, Shi B, Wilcox AJ, Lie RT, Jee SH, Christensen K, Doheny
37 27 KF, Pugh EW, Ling H, Scott AF. 2011. Evidence for gene-environment interaction in a genome wide
38 28 study of nonsyndromic cleft palate. *Genet Epidemiol* 35:469–478.
- 39
40 29 Beaty TH, Taub MA, Scott AF, Murray JC, Marazita ML, Schwender H, Parker MM, Hetmanski JB,
41 30 Balakrishnan P, Mansilla MA, Mangold E, Ludwig KU, Nöthen MM, Rubini M, Elcioglu N, Ruczinski I.
42 31 2013. Confirming genes influencing risk to cleft lip with/without cleft palate in a case-parent trio
43 32 study. *Hum Genet* 132:771-781.
- 44
45 33 Becker T, Knapp M. 2004. A powerful strategy to account for multiple testing in the context of
46 34 haplotype analysis. *Am J Hum Genet* 75:561–570.
- 47
48 35 Bhattacharjee V, Horn KH, Singh S, Webb CL, Pisano MM, Greene RM. 2009. CBP/p300 and
49 36 associated transcriptional co-activators exhibit distinct expression patterns during murine
50 37 craniofacial and neural tube development. *Int J Dev Biol* 53:1097–1104.
- 51
52 38 Birnbaum S, Ludwig KU, Reutter H, Herms S, Steffens M, Rubini M, Baluardo C, Ferrian M, Almeida de
53 39 Assis N, Alblas MA, Barth S, Freudenberg J, Lauster C, Schmidt G, Scheer M, Braumann B, Berge SJ,
54 40 Reich RH, Schiefke F, Hemprich A, Potzsch S, Steegers-Theunissen RP, Potzsch B, Moebus S,
55 41 Horsthemke B, Kramer FJ, Wienker TF, Mossey PA, Propping P, Cichon S, Hoffmann P, Knapp M,
56 42 Nöthen MM, Mangold E. 2009. Key susceptibility locus for nonsyndromic cleft lip with or without
57 43 cleft palate on chromosome 8q24. *Nat Genet* 41:473-477.

Ishorst et al., Candidate loci association study in nsCPO

- 1
2
3 1 Cobat, A, Abel L., Alcaïa A, Schurr E. 2014. A general efficient and flexible approach for genome-wide
4 2 association analyses of imputed genotypes in family-based designs. *Genet Epidemiol* 38:560-571.
- 5 3 Ding H-L, Hooper JE, Batzel P, Eames BF, Postlethwait JH, Artinger KB, Clouthier DE. 2016. MicroRNA
6 4 Profiling during Craniofacial Development: Potential Roles for Mir23b and Mir133b. *Front Physiol*
7 5 7:281.
- 8
9 6 Exome Aggregation Consortium. 2016. *Nature* 536:285-291.
- 10
11 7 Grosen D, Bille C, Petersen I, Skytthe A, Bornemann J Von, Pedersen JK, Murray JC, Christensen K.
12 8 2011. Risk of Oral Clefts in twins. *Epidemiology* 22:313–319.
- 13
14 9 Hennekam RC, Tilanus M, Hamel BC, Voshart-van Heeren H, Mariman EC, van Beersum SE, van den
15 10 Boogaard MJ, Breuning MH. 1993. Deletion at chromosome 16p13.3 as a cause of Rubinstein-Taybi
16 11 syndrome: clinical aspects. *Am J Hum Genet* 52:255–262.
- 17
18 12 Hoebel AK, Drichel D, van de Vorst M, Böhmer AC, Sivalingam S, Ishorst N, Klamt J, Gözl L, Alblas M,
19 13 Maaser A, Keppler K, Zink AM, Dixon MJ, Dixon M, Hemprich A, Kruse T, Graf I, Dunsche A, Schmidt
20 14 G, Daratsianos N, Nowak S, Aldhorae KA, Nöthen MM, Knapp M, Thiele H, Gilissen C, Reutter H,
21 15 Hoischen A, Mangold E, Ludwig KU. 2017. Candidate Genes for Nonsyndromic Cleft Palate Detected
22 16 by Exome Sequencing. *J Dent Res* 96:1314-1321.
- 23
24 17 Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. 2008. SNAP: a web-
25 18 based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 24:2938–
26 19 2939.
- 27
28 20 Kheradpour P, Kellis M. 2013. Systematic discovery and characterization of regulatory motifs in
29 21 ENCODE TF binding experiments. *Nucleic Acids Res* 42:2976–2987.
- 30
31 22 Leslie EJ, Liu H, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF,
32 23 McHenry T, Resick J, Sanchez C, Jacobs J, Emanuele B, Vieira AR, Neiswanger K, Standley J, Czeizel AE,
33 24 Deleyiannis F, Christensen K, Munger RG, Lie RT, Wilcox A, Romitti PA, Field LL, Padilla CD, Cutiongco-
34 25 de la Paz EM, Lidral AC, Valencia-Ramirez LC, Lopez-Palacio AM, Valencia DR, Arcos- Burgos M,
35 26 Castilla EE, Mereb JC, Poleta FA, Orioli IM, Carvalho FM, Hecht JT, Blanton SH, Buxó CJ, Butali A,
36 27 Mossey PA, Adeyemo WL, James O, Braimah RO, Aregbesola BS, Eshete MA, Deribew M, Koruyucu
37 28 M, Seymen F, Ma L, de Salamanca JE, Weinberg SM, Moreno L, Cornell RA, Murray JC, Marazita ML.
38 29 2016a. A Genome-wide Association Study of Nonsyndromic Cleft Palate Identifies an Etiologic
39 30 Missense Variant in GRHL3. *Am J Hum Genet* 98:744–754.
- 40
41 31 Leslie EJ, Carlson JC, Shaffer JR, Feingold E, Wehby G, Laurie CA, Jain D, Laurie CC, Doheny KF,
42 32 McHenry T, Resick J, Sanchez C, Jacobs J, Emanuele B, Vieira AR, Neiswanger K, Lidral AC, Valencia-
43 33 Ramirez LC, Lopez-Palacio AM, Valencia DR, Arcos-Burgos M, Czeizel AE, Field LL, Padilla CD,
44 34 Cutiongco-de la Paz EM, Deleyiannis F, Christensen K, Munger RG, Lie RT, Wilcox A, Romitti PA,
45 35 Castilla EE, Mereb JC, Poletta FA, Orioli IM, Carvalho FM, Hecht JT, Blanton SH, Buxo CJ, Butali A,
46 36 Mossey PA, Adeyemo WL, James O, Braimah RO, Aregbesola BS, Eshete MA, Abate F, Koruyucu M,
47 37 Seymen F, Ma L, de Salamanca JE, Weinberg SM, Moreno L, Murray JC, Marazita ML. 2016b. A multi-
48 38 ethnic genome-wide association study identifies novel loci for non-syndromic cleft lip with or without
49 39 cleft palate on 2p24.2, 17q23 and 19q13. *Hum Mol Genet* 25:2862-2872.
- 50
51 40 Leslie EJ, Carlson JC, Shaffer JR, Butali A, Buxo CJ, Castilla EE, Christensen K, Deleyiannis FW, Leigh
52 41 Field L, Hecht JT, Moreno L, Orioli IM, Padilla C, Vieira AR, Wehby GL, Feingold E, Weinberg SM,
53 42 Murray JC, Beaty TH, Marazita ML. 2017. Genome-wide meta-analyses of nonsyndromic orofacial
54 43 clefts identify novel associations between FOXE1 and all orofacial clefts, and TP63 and cleft lip with
55 44 or without cleft palate. *Hum Genet* 136:275-286.

Ishorst et al., Candidate loci association study in nsCPO

- 1
2
3 1 Li D, Zhang H, Ma L, Han Y, Xu M, Wang Z, Jiang H, Zhang W, Wang L, Pan Y. 2016. Associations
4 2 between microRNA binding site SNPs in FGFs and FGFRs and the risk of non-syndromic orofacial cleft.
5 3 Sci Rep 6:31054.
- 6 4 Ludwig KU, Ahmed ST, Böhmer AC, Sangani NB, Varghese S, Klamt J, Schuenke H, Gultepe P,
7 5 Hofmann A, Rubini M, Aldhorae KA, Steegers-Theunissen RP, Rojas-Martinez A, Reiter R, Borck G,
8 6 Knapp M, Nakatomi M, Graf D, Mangold E, Peters H. 2016. Meta-analysis Reveals Genome-Wide
9 7 Significance at 15q13 for Nonsyndromic Clefting of Both the Lip and the Palate, and Functional
10 8 Analyses Implicate GREM1 As a Plausible Causative Gene. PLoS genetics 12:e1005914.
- 11 9 Ludwig KU, Böhmer AC, Rubini M, Mossey PA, Herms S, Nowak S, Reutter H, Alblas MA, Lippke B,
12 10 Barth S, Paredes-Zenteno M, Munoz-Jimenez SG, Ortiz-Lopez R, Kreuzsch T, Hemprich A, Martini M,
13 11 Braumann B, Jager A, Potzsch B, Molloy A, Peterlin B, Hoffmann P, Nöthen MM, Rojas-Martinez A,
14 12 Knapp M, Steegers-Theunissen RP, Mangold E. 2014. Strong association of variants around FOXE1
15 13 and orofacial clefting. J Dent Res 93:376-381.
- 16 14 Ludwig KU, Böhmer AC, Bowes J, Nikolic M, Ishorst N, Wyatt N, Hammond NL, Gölz L, Thieme F, Barth
17 15 S, Schuenke H, Klamt J, Spielmann M, Aldhorae K, Rojas-Martinez A, Nöthen MM, Rada-Iglesias A,
18 16 Dixon MJ, Knapp M, Mangold E. 2017. Imputation of Orofacial Clefting Data Identifies Novel Risk Loci
19 17 and Sheds Light on the Genetic Background of Cleft Lip +/- Cleft Palate and Cleft Palate Only. Hum
20 18 Mol Genet 26:829-842.
- 21 19 Ludwig KU, Mangold E, Herms S, Nowak S, Reutter H, Paul A, Becker J, Herberz R, AlChawa T, Nasser
22 20 E, Böhmer AC, Mattheisen M, Alblas MA, Barth S, Kluck N, Lauster C, Braumann B, Reich RH,
23 21 Hemprich A, Potzsch S, Blaumeiser B, Daratsianos N, Kreuzsch T, Murray JC, Marazita ML, Ruczinski I,
24 22 Scott AF, Beaty TH, Kramer FJ, Wienker TF, Steegers-Theunissen RP, Rubini M, Mossey PA, Hoffmann
25 23 P, Lange C, Cichon S, Propping P, Knapp M, Nöthen MM. 2012. Genome-wide meta-analyses of
26 24 nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet 44:968-971.
- 27 25 Mangold E, Böhmer AC, Ishorst N, Hoebel AK, Gültepe P, Schuenke H, Klamt J, Hofmann A, Gölz L,
28 26 Raff R, Tessmann P, Nowak S, Reutter H, Hemprich A, Kreuzsch T, Kramer FJ, Braumann B, Reich R,
29 27 Schmidt G, Jäger A, Reiter R, Brosch S, Stavakis J, Ishida M, Seselgyte R, Moore GE, Nöthen MM,
30 28 Borck G, Aldhorae KA, Lace B, Stanier P, Knapp M, Ludwig KU. 2016. Sequencing the GRHL3 Coding
31 29 Region Reveals Rare Truncating Mutations and a Common Susceptibility Variant for Nonsyndromic
32 30 Cleft Palate. Am J Hum Genet 98:755–762.
- 33 31 Mangold E, Ludwig KU, Birnbaum S, Baluardo C, Ferrian M, Herms S, Reutter H, de Assis NA, Chawa
34 32 TA, Mattheisen M, Steffens M, Barth S, Kluck N, Paul A, Becker J, Lauster C, Schmidt G, Braumann B,
35 33 Scheer M, Reich RH, Hemprich A, Pöttsch S, Blaumeiser B, Moebus S, Krawczak M, Schreiber S,
36 34 Meitinger T, Wichmann HE, Steegers-Theunissen RP, Kramer FJ, Cichon S, Propping P, Wienker TF,
37 35 Knapp M, Rubini M, Mossey PA, Hoffmann P, Nöthen MM. 2010. Genome-wide association study
38 36 identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat. Genet.
39 37 42:24–26.
- 40 38 Mangold E, Ludwig KU, Nöthen MM. 2011. Breakthroughs in the genetics of orofacial clefting. Trends
41 39 Mol Med 17:725-733.
- 42 40 McCarthy S, Das S, Kretschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C,
43 41 Danecek P, Sharp K, Luo Y, Sidore C, Kwong A, Timpson N, Koskinen S, Vrieze S, Scott LJ, Zhang H,
44 42 Mahajan A, Veldink J, Peters U, Pato C, van Duijn CM, Gillies CE, Gandin I, Mezzavilla M, Gilly A, Cocca
45 43 M, Traglia M, Angius A, Barrett JC, Boomsma D, Branham K, Breen G, Brummett CM, Busonero F,
46 44 Campbell H, Chan A, Chen S, Chew E, Collins FS, Corbin LJ, Smith GD, Dedoussis G, Dorr M, Farmaki
47 45 AE, Ferrucci L, Forer L, Fraser RM, Gabriel S, Levy S, Groop L, Harrison T, Hattersley A, Holmen OL,
48 46 Hveem K, Kretzler M, Lee JC, McGue M, Meitinger T, Melzer D, Min JL, Mohlke KL, Vincent JB, Nauck
49 47 M, Nickerson D, Palotie A, Pato M, Pirastu N, McInnis M, Richards JB, Sala C, Salomaa V, Schlessinger
50 48 D, Schoenherr S, Slagboom PE, Small K, Spector T, Stambolian D, Tuke M, Tuomilehto J, Van den Berg

Ishorst et al., Candidate loci association study in nsCPO

- 1
2
3 1 LH, Van Rheenen W, Volker U, Wijmenga C, Toniolo D, Zeggini E, Gasparini P, Sampson MG, Wilson
4 2 JF, Frayling T, de Bakker PI, Swertz MA, McCarroll S, Kooperberg C, Dekker A, Altshuler D, Willer C,
5 3 lacono W, Ripatti S, Soranzo N, Walter K, Swaroop A, Cucca F, Anderson CA, Myers RM, Boehnke M,
6 4 McCarthy MI, Durbin R; Haplotype Reference Consortium. 2016. A reference panel of 64,976
7 5 haplotypes for genotype imputation. *Nat Genet* 48:1279–1283.
- 8
9 6 Moreno LM, Mansilla MA, Bullard SA, Cooper ME, Busch TD, Machida J, Johnson MK, Brauer D, Krahn
10 7 K, Daack-Hirsch S, L'Heureux J, Valencia-Ramirez C, Rivera D, Lopez AM, Moreno MA, Hing A, Lammer
11 8 EJ, Jones M, Christensen K, Lie RT, Jugessur A, Wilcox AJ, Chines P, Pugh E, Doheny K, Arcos-Burgos
12 9 M, Marazita ML, Murray JC, Lidral AC. 2009. FOXE1 association with both isolated cleft lip with or
13 10 without cleft palate, and isolated cleft palate. *Hum Mol Genet* 18:4879-4896.
- 14
15 11 Moreno Uribe LM, Fomina T, Munger RG, Romitti PA, Jenkins MM, Gjessing HK, Gjerdevik M,
16 12 Christensen K, Wilcox AJ, Murray JC, Lie RT, Wehby GL. 2017. A Population-Based Study of Effects of
17 13 Genetic Loci on Orofacial Clefts. *J Dent Res* 96:1322-1329.
- 18
19 14 Mossey PA, Little J, Steegers-Theunissen R, Molloy A, Peterlin B, Shaw WC, Johnson C, FitzPatrick DR,
20 15 Franceschelli P, Rubini M. 2017. Genetic Interactions in Nonsyndromic Orofacial Clefts in Europe—
21 16 EUROCRAN Study. *Cleft Palate Craniofac J* 54:623-630.
- 22
23 17 Mostowska A, Gaczkowska A, Żukowski K, Ludwig KU, Hozyasze KK, Wójcickif P, Mangold E, Böhmer
24 18 AC, Heilmann-Heimbach S, Knapp M, Zadurskai M, Biedziak B, Budner M, Lasota A, Daktera-Micker
25 19 A, Jagodziński PP. 2017. Common variants in DLG1 locus are associated with non-syndromic cleft lip
26 20 with or without cleft palate. *Clin Genet* doi: 10.1111/cge.13141.
- 27
28 21 Rahimov F, Marazita ML, Visel A, Cooper ME, Hitchler MJ, Rubini M, Domann FE, Govil M,
29 22 Christensen K, Bille C, Melbye M, Jugessur A, Lie RT, Wilcox AJ, Fitzpatrick DR, Green ED, Mossey PA,
30 23 Little J, Steegers-Theunissen RP, Pennacchio LA, Schutte BC, Murray JC. 2008. Disruption of an AP-
31 24 2alpha binding site in an IRF6 enhancer is associated with cleft lip. *Nat Genet* 40:1341-1347.
- 32
33 25 Sun Y, Huang Y, Yin A, Pan Y, Wang Y, Wang C, Du Y, Wang M, Lan F, Hu Z, Wang G, Jiang M, Ma J,
34 26 Zhang X, Ma H, Ma J, Zhang W, Huang Q, Zhou Z, Ma L, Li Y, Jiang H, Xie L, Jiang Y, Shi B, Cheng J,
35 27 Shen H, Wang L, Yang Y. 2015. Genome-wide association study identifies a new susceptibility locus
36 28 for cleft lip with or without a cleft palate. *Nat Commun* 6:6414.
- 37
38 29 The 1000 genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature*
39 30 526:68-74.
- 40
41 31 Thieme F, Ludwig KU. 2017. The Role of Noncoding Genetic Variation in Isolated Orofacial Clefts. *J*
42 32 *Dent Res* 96:1238-1247.
- 43
44 33 Wang J, Bai Y, Li H, Greene SB, Klysis E, Yu W, Schwartz RJ, Williams TJ, Martin JF. 2013. MicroRNA-
45 34 17-92, a direct Ap-2α transcriptional target, modulates T-box Factor activity in orofacial clefting. *PLoS*
46 35 *Genet* 9:e1003785.
- 47
48 36 Willer CJ, Li Y, Abecasis GR. 2010. METAL: fast and efficient meta-analysis of genomewide association
49 37 scans. *Bioinformatics*. 26:2190-1.
- 50
51 38 Yu Y, Zuo X, He M, Gao J, Fu Y, Qin C, Meng L, Wang W, Song Y, Cheng Y, Zhou F, Chen G, Zheng X,
52 39 Wang X, Liang B, Zhu Z, Fu X, Sheng Y, Hao J, Liu Z, Yan H, Mangold E, Ruczinski I, Liu J, Marazita ML,
53 40 Ludwig KU, Beaty TH, Zhang X, Sun L, Bian Z. 2017. Genome-wide analyses of non-syndromic cleft lip
54 41 with palate identify 14 novel loci and genetic heterogeneity. *Nat Commun* 8:14364.

55 43 **Table and Figure legends:**

56
57 44 **Figure 1:** Workflow of the present study.

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 Figure 2: Regional association plot of rs395572 locus +/- 500kb. $P_{EU_{Beaty}}$ = P-value from imputed GWAS
4 2 of European trios. P_{MetaEU} = P-value from imputed GWAS of European trios and all genotyped samples
5 3 with European ancestry.

6
7
8 4 Figure 3: Regional association plot of rs6809420 locus +/- 500kb. $P_{EU+Asia_{Beaty}}$ = P-value from imputed
9 5 GWAS of all trios. $P_{MetaAll}$ = P-value from imputed GWAS of all trios and all genotyped samples. Genes
10 6 *TUSC1* and *IGSF11* are located 1 Mb upstream and 1.2 Mb downstream of rs6809420, respectively.
11 7 Rs7633157 belongs to another locus indicated by LD (represented by dashed line) and failed
12 8 genotyping.

13
14
15
16 9 Table 1: Association results for SNPs with nominally significant associations in at least one of three
17 10 genotyped samples and/or their meta-analysis ($Meta_{geno}$).

18
19
20 11 Table 2: Results of European meta-analysis (SNPs with a decreased P-value in meta-analysis as
21 12 compared to initial imputed GWAS dataset (EU_{Beaty}) are shown.

22
23
24 13 Table 3: Results of meta-analysis of all genotyped samples and $EU+Asia_{Beaty}$ (SNPs with a decreased P-
25 14 value after meta-analysis in comparison to the initial imputed GWAS dataset ($EU+Asia_{Beaty}$) are shown
26 15 here.)

27
28
29 16 Supplementary Figure S1: Regional association plot of rs595533 locus +/- 500kb. $P_{EU+Asia_{Beaty}}$ = P-value
30 17 from imputed GWAS of all trios. $P_{MetaAll}$ = P-value from imputed GWAS of all trios and all genotyped
31 18 samples.

32
33
34
35 19 Supplementary Table S1: Gender distribution and origin of the three genotyped samples

36
37 20 Supplementary Table S2: 83 SNPs with $P < 1 \times 10^{-5}$ in imputed datasets EU_{Beaty} and $EU+Asia_{Beaty}$

38
39 21 Supplementary Table S3: Genotyped SNPs from imputed GWAS (EU_{Beaty} , $EU+Asia_{Beaty}$)

40
41 22 Supplementary Table S4: SNPs from Leslie et al. (2016a)

42
43 23 Supplementary Table S5: SNPs finally selected for genotyping

44
45
46 24 Supplementary Table S6: Relative risks of genotyped samples and meta-analysis of all genotyped
47 25 samples ($Meta_{geno}$) derived from a fixed-effects based meta-analysis.

48
49
50 26 Supplementary Table S7: Association P-values of EU_{Beaty} and $EU+Asia_{Beaty}$ at the 14 novel nsCL/P loci
51 27 from Yu et al. (2017) and one novel locus from Leslie et al. (2017) at lead SNP level

52
53
54 28 Supplementary Table S8: Association results for two European ($EU_{Mangold}$ and EU_{Mossey}) and one
55 29 Arabian ($Arab_{Aldhorae}$) nsCPO sample and the corresponding meta-analysis

Ishorst et al., Candidate loci association study in nsCPO

1
2
3 1 Supplementary Table S9: Association results for the European nsCPO samples (EU_{Mangold} and EU_{Mossey}),
4 2 imputation of the European nsCPO trios (EU_{Beaty}) and the corresponding meta-analysis (Meta_{EU})

5
6 3 Supplementary Table S10: Association results for the European and the Arabian nsCPO samples
7 4 (EU_{Mangold}, EU_{Mossey} and Arab_{Aldhorae}), imputation of the nsCPO trios (EU+Asia_{Beaty}) and the
8 5 corresponding meta-analysis (Meta_{All}).

9
10 6 Supplementary Table S11: Association P-values of EU_{Beaty} at 14 novel nsCL/P loci from Yu et al. (2017)
11 7 and one novel locus from Leslie et al. (2017). For each locus all SNPs with P-value $< 1 \times 10^{-3}$ (info-score
12 8 > 0.8) or if none present the SNPs with the two lowest P-values (info-score > 0.8) are shown.

13
14
15
16
17 9 Supplementary Table S12: Association P-values of EU+Asia_{Beaty} at 14 novel nsCL/P loci from Yu et al.
18 10 (2017) and one novel locus from Leslie et al. (2017). For each locus all SNPs with P-value $< 1 \times 10^{-3}$
19 11 (info-score > 0.8) or if none present the SNPs with the two lowest P-values (info-score > 0.8) are
20 12 shown.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

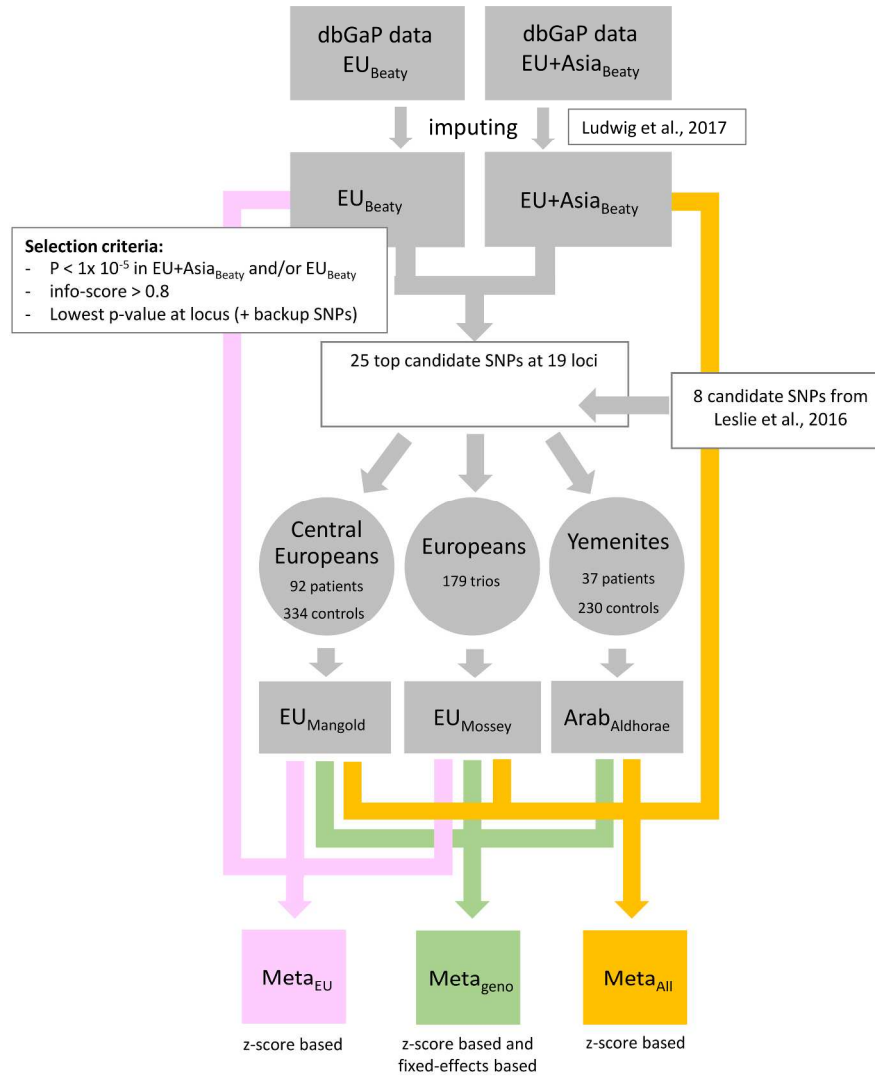


Figure 1: Workflow of the present study.

233x318mm (300 x 300 DPI)

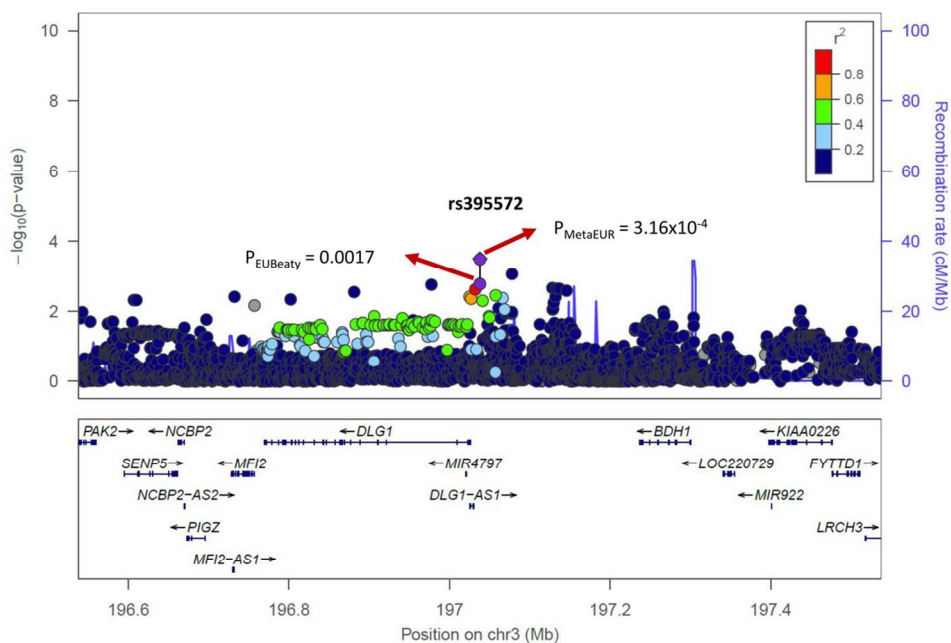


Figure 2: Regional association plot of rs395572 locus +/- 500kb. P_{EUBeaty} = P-value from imputed GWAS of European trios. P_{MetaEU} = P-value from imputed GWAS of European trios and all genotyped samples with European ancestry.

109x74mm (300 x 300 DPI)

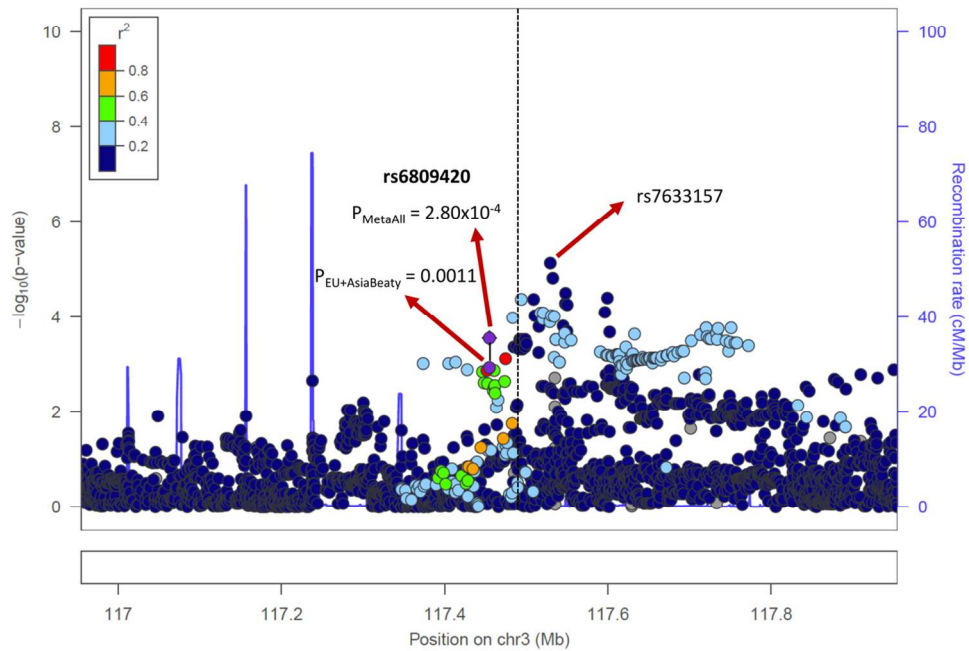


Figure 3: Regional association plot of rs6809420 locus +/- 500kb. $P_{\text{EU+AsiaBeaty}}$ = P-value from imputed GWAS of all trios. P_{MetaAll} = P-value from imputed GWAS of all trios and all genotyped samples. Genes *TUSC1* and *IGSF11* are located 1 Mb upstream and 1.2 Mb downstream of rs6809420, respectively. Rs7633157 belongs to another locus indicated by LD (represented by dashed line) and failed genotyping.

119x91mm (300 x 300 DPI)

Table 1 Association results for SNPs with nominally significant associations in at least one of three genotyped samples and/or their meta-analysis (Meta_{geno})

SNP information				Association P-values ^c			
SNP	Chr	Pos (hg19)	Risk/other allele ^a	EU _{Mangold} P _{Trend}	Arab _{Aldhorae} P _{Trend}	EU _{Mossey} P _{TDT}	Meta _{geno}
rs6747560	2	32784766	G/T ^b	0.007^d	<i>0.414</i>	0.320	0.046^d
rs1283107	3	106961320	G/C ^b	0.383	0.281	0.142	0.049^d
rs6809420	3	117454822	T/G	0.320	0.024	0.647	0.066
rs6805813	3	187595294	G/A ^b	0.733	0.019^d	<i>0.535</i>	0.456
rs395572	3	197037952	A/G	0.309	0.042^e	0.058	0.278

SNP = single nucleotide polymorphism; Chr = Chromosome; Pos (hg19) = position in human reference genome version 19

Nominally significant P-values are given in bold; P-values given in italics indicate an opposite effect direction as in the meta-analysis (Meta_{geno})

^a risk allele according to Meta_{geno}

^b different risk allele as compared to imputed nsCPO datasets (EU_{Beaty} and EU+Asia_{Beaty})

^c Association P-values for EU_{Mangold} and Arab_{Aldhorae} cohorts were calculated using the Armitage trend test (P_{trend}); Association P-values for EU_{Mossey} were calculated using the transmission disequilibrium test (P_{TDT})

^d SNP shows nominally significant association but in the opposite effect direction to that observed in the initial GWAS data

^e SNP shows nominally significant association but in an opposite effect direction as the initial GWAS data and Meta_{geno}

Table 2 Results of European meta-analysis (SNPs with a decreased P-value in meta-analysis as compared to initial imputed GWAS dataset (EU_{Beaty}) are shown.)

SNP information				Association P-values ^c			
SNP	Chr	Pos (hg19)	Risk/other allele ^b	EU _{Mangold} P _{Trend}	EU _{Mossey} P _{TDT}	EU _{Beaty}	Meta _{EU}
rs395572	3	197037952	A/G	0.309	0.058	0.002	3.16x10⁻⁴
rs7732608 ^a	5	155799466	A/G	0.588	0.117	0.496	0.106
rs3740617 ^a	11	33881016	C/T	0.062	0.651	0.417	0.081

SNP = single nucleotide polymorphism; Chr = Chromosome; Pos (hg19) = position in human reference genome version 19

Nominally significant P-values of the meta-analysis are given in bold

^a SNP derived from Leslie et al. 2016

^b risk allele according to Meta_{EU}

^c Association P-values for EU_{Mangold} and EU_{Beaty} were calculated using the Armitage trend test (P_{trend}); Association p-values for EU_{Rubini&Mossey} were calculated using the transmission disequilibrium test (P_{TDT})

Table 3 Results of meta-analysis of all genotyped samples and EU+Asia_{Beaty} (SNPs with a decreased P-value after meta-analysis in comparison to the initial imputed GWAS dataset (EU+Asia_{Beaty}) are shown here.

SNP information				Association P-values ^c				
SNP	Chr	Pos (hg19)	Risk/other allele ^b	EU _{Mangold} P _{Trend}	Arab _{Aldhorae} P _{Trend}	EU _{Mossey} P _{TDT}	EU+Asia _{Beaty}	Meta _{All}
rs6809420	3	117454822	T/G	0.320	0.024	0.647	0.001	2.80x10⁻⁴
rs7732608 ^a	5	155799466	A/G	0.588	<i>0.434</i>	0.117	0.498	0.218
rs9347594 ^a	6	162587892	C/T	0.084	0.789	0.824	0.728	0.249
rs3740617 ^a	11	33881016	C/T	0.062	<i>0.626</i>	0.651	0.640	0.232
rs595533 ^a	18	66441589	T/C	0.147	0.961	0.712	0.087	0.041

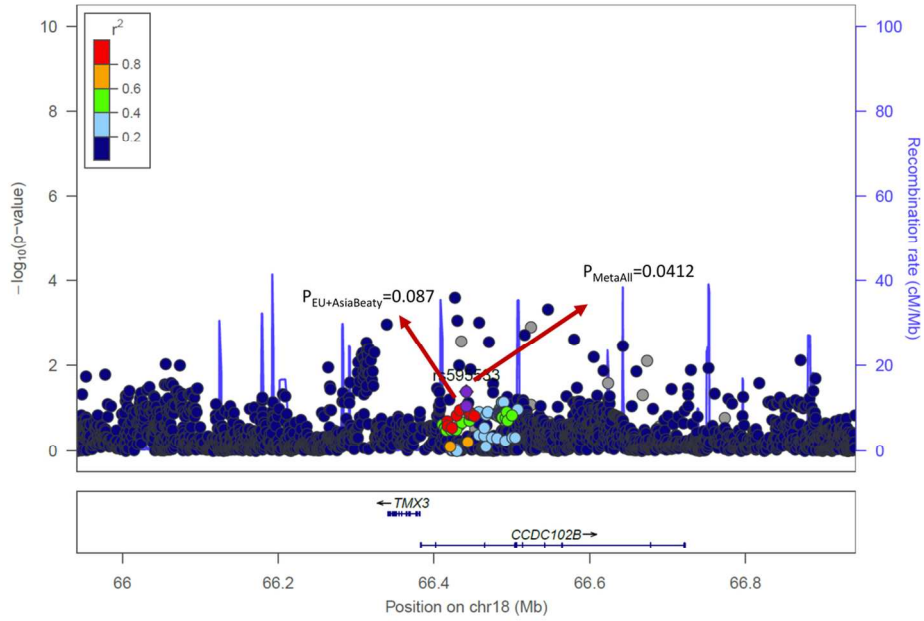
SNP = single nucleotide polymorphism; Chr = Chromosome; Pos (hg19) = position in human reference genome version 19

Nominally significant P-values of the meta-analysis are given in bold; P-values given in italics indicate an opposite effect direction as in the meta-analysis (Meta_{All})

^a SNP derived from Leslie et al. 2016

^b risk allele according to Meta_{All}

^c Association P-values for EU_{Mangold}, Arab_{Aldhorae} and EU_{Beaty} were calculated using the Armitage trend test (P_{Trend}); Association P-values for EU_{Mossey} were calculated using the transmission disequilibrium test (P_{TDT})



Supplementary Figure S1: Regional association plot of rs595533 locus +/- 500kb. $P_{EU+AsiaBeaty}$ = P-value from imputed GWAS of all trios.

119x85mm (300 x 300 DPI)