Human Molecular Genetics, 2026, 35(12), ddag052
https://doi.org/10.1093/hmg/ddag052
Published: 21 June 2026
Original Article

OXFORD

Evidence that disruption of Discoidin domain receptor 2 contributes to palate malformations through effects on the extracellular matrix

Julia A. Capecki[1,2], Helena Shkuro[2], Öznur Yilmaz[1], Lisa Schmitt[1], Khadija Channab[3], Tobias T. Lindenberg[1], Teresa Kruse[4], Sarah Achterrath[4], Berta Crespo[5], Anna Siewert[2], Mostafa Bakhshi[3], Leandra Pantel[2], Kerstin U. Ludwig[2], Matthias Geyer[6], Elisabeth Mangold[2,†], Benjamin Odermatt[1,3,†], Nina Ishorst[1,2,†,*]

[1]Institute of Neuroanatomy, University of Bonn, School of Medicine & University Hospital Bonn, Nussallee 10, 53115 Bonn, Germany
[2]Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany
[3]Institute of Anatomy and Cell Biology, University of Bonn, School of Medicine & University Hospital Bonn, Nussallee 10, 53115 Bonn, Germany
[4]Department of Orthodontics, University of Cologne, Faculty of Medicine and University Hospital Cologne, Kerpener Str. 32, 50931 Cologne, Germany
[5]Human Developmental Biology Resource (HDBR), Great Ormond Street Institute of Child Health, University College London, 30 Guilford Street, WC1N 1EH London, United Kingdom
[6]Institute of Structural Biology, University of Bonn, School of Medicine & University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

*Corresponding author. Institute of Neuroanatomy, University of Bonn, Nussallee 10, 53115 Bonn. E-mail: Nina.Ishorst@uni-bonn.de
†Elisabeth Mangold, Benjamin Odermatt and Nina Ishorst Contributed equally

Abstract

Orofacial clefting (OFC) is among the most common birth defects and can occur either as part of a syndrome or in isolation (nonsyndromic, ns). Cleft palate only (CPO) is an OFC subtype. Here, we searched for novel nsCPO risk genes carrying homozygous and compound-heterozygous variants, by analyzing exome data from six sibling pairs with nsCPO born to unaffected parents. After stringent quality control and filtering, we identified 6 homozygous variants and 32 compound-heterozygous variants in 5 and 16 candidate genes, respectively. We prioritized *DDR2*, a collagen-activated receptor-tyrosine-kinase influencing extracellular matrix composition, as our top candidate for functional follow-up, since variants in this gene can cause Warburg-Cinotti syndrome, the phenotypic spectrum of which includes palatal abnormalities. Knock-down and knock-out of *DDR2*-orthologs in zebrafish caused craniofacial abnormalities resembling CPO in humans. Zebrafish immunostaining indicated that *DDR2*-orthologs were expressed in mature head muscle cells, while murine single-cell RNA-Sequencing data detected *Ddr2* expression only in head muscle progenitor cells; the latter finding was confirmed in human embryo sections stained for *DDR2*. *DDR2*-expressing head muscle progenitor cells may influence extracellular matrix composition through DDR2-mediated signaling, thereby affecting outgrowth, elevation, and fusion of the palatal shelves, a previously postulated mechanism involved in palatogenesis. Most established OFC genes (e.g. *CDH1*, *CTNND1*, *IRF6*, and *GRHL3*) act via mechanisms related to epithelial integrity and periderm differentiation, whereas our data provide evidence supporting *DDR2* as a risk gene for nsOFC that functions by influencing extracellular matrix composition.

Significance Statement

Nonsyndromic orofacial clefting (OFC) is a common birth defect with a significant impact on the lives of affected individuals and broader society. The etiology of OFC is multifactorial, involving environmental and genetic factors, which both remain poorly understood. We investigated a promising risk gene, encoding the receptor-tyrosine-kinase, DDR2, in zebrafish, mouse, and human, and found evidence for its role in OFC. Interestingly, established OFC genes act via mechanisms related to epithelial integrity and periderm differentiation, while DDR2 appears to act via extracellular matrix interactions mediated by muscle progenitors. Our findings contribute to deeper understanding of DDR2's function in craniofacial development and mechanisms underlying OFC. This knowledge could inform risk prediction and genetic counseling for affected families.

Keywords craniofacial development, palatogenesis, DDR2, cleft palate, zebrafish, bulk RNA-Seq

Received: November 20, 2025. **Revised:** April 23, 2026. **Accepted:** June 5, 2026

Introduction

Orofacial clefting (OFC) is a common congenital anomaly, occurring in approximately 1 in every 700 newborns worldwide [1, 2]. Cleft palate only (CPO; MIM 119540) is among the most frequent forms of OFC and is defined by a gap between either the two palatal shelves that form the hard palate, or the muscular structures of the soft palate [3]. Newborns with this condition often have feeding difficulties and require a combination of corrective surgeries, therapies, and speech training, which profoundly impact the quality of life of affected individuals and their families, as well as broader society [4, 5].

In humans, the hard palate emerges from paired maxillary processes; its development begins around the fifth week post conception and is concluded by the twelfth week [6, 7]. Outgrowths form vertically alongside the tongue before reorienting horizontally above it, in a process known as palatal shelf elevation. Subsequently, the palatal shelves converge and fuse at the midline [8]. This complex palate fusion process requires a delicate equilibrium of cell communication, migration, proliferation, and apoptosis [1], any disruptions of which can result in orofacial malformations, such as clefting of the hard and/or soft palate [8].

CPO occurs in either a nonsyndromic form, without any further anomalies, or in the context of syndromes; approximately 55% of patients with CPO have the nonsyndromic form (nsCPO) [9], which has a multifactorial etiology, with a high genetic component indicated by heritability estimates of more than 90% [10]. Genetic and epidemiological data indicate that variants in 'major' genes act on a multifactorial background [11–13]. A few major genes have been identified by studies sequencing candidate genes selected based on evidence from syndromic forms of OFC or via exome sequencing (ES) in multiple affected families [11, 14]. Notably, those studies focused primarily on genes with autosomal-dominant inheritance.

In the present study, we aimed to identify recessively inherited nsCPO candidate genes by reanalyzing a previously published ES dataset from six sibling pairs (sib-pairs) with nsCPO and their unaffected parents, which was originally investigated to identify dominant heterozygous variants [15]. Further, we conducted functional analysis of the selected candidate gene, Discoidin Domain Receptor 2 (*DDR2*), using zebrafish and mouse models, as well as determining *DDR2* expression patterns during human embryonic development. Our findings provide a foundation for improved diagnostic and predictive genetic counseling.

Results

Exome sequencing analysis, filtering, and prioritization

Reanalysis of an existing exome sequencing (ES) dataset from six affected sib-pairs [15] identified recessive (compound-heterozygous and homozygous) variants in 79 genes (Fig. 1A). A detailed description of the filtering strategy and prioritization based on pathogenicity scores, such as SIFT (Sorting Intolerant From Tolerant), PolyPhen-2 (Polymorphism Phenotyping), and Z-score (GnomAD v4.1.0, https://gnomad.broadinstitute.org), as well as expression analyses of RNASeq data from murine palates at E13.5 [16, 17] is provided in the appendix (*SI Appendix*, Fig. S1). After filtering and prioritization, 6 homozygous variants in 5 candidate genes and 32 compound-heterozygous variants in 16 candidate genes remained (*SI Appendix*, Table S1). Among these 21 candidate genes, one carried a homozygous splice site variant, one had a compound-heterozygous deletion and a synonymous variant, and the remainder carried missense variants.

Validation of *DDR2* variants and additional findings

In one family (BN00293), two compound-heterozygous missense variants were detected in *DDR2* (Fig. 1B and C), one of which (NM_006182.4:c.1352C > T; rs150920581; CADD score 24.5) was among missense variants classified as high priority, while the other (NM_006182.4:c.518C > G; rs202091759; CADD score 25.3) was classified as intermediate-priority. Further analyses revealed *DDR2* as one of the two candidate genes most intolerant towards missense mutations (Z-score = 2.29) and as having the highest expression level (RPKM = 48.02). In addition, family BN00293 carried two compound-heterozygous variants in *GRHL3*: a previously identified low-frequency variant (NM_198173.3:c.1361C > T; minor allele frequency in Non-Finnish Europeans = 0.0355 (GnomAD v4.1.0)) [11] and a novel in-frame deletion (NM_198173.3:c.1722_1724del). All four variants were validated as compound-heterozygous based on parent sequence data.

Prioritization of *DDR2* as a candidate gene

Literature-based research using Pubmed (https://pubmed.ncbi.nlm.nih.gov/) and GeneCards (https://genecards.org/) (both accessed in June 2021) revealed that *DDR2* is a collagen-activated receptor-tyrosine-kinase that plays a crucial role in various developmental processes, including cell adhesion, migration, and proliferation [18]. *DDR2* is highly expressed during human and murine embryonic development, at time-points and in tissues relevant to craniofacial development, particularly in mesenchymal cells [19]. Through its interactions with the extracellular matrix (ECM), DDR2 has been suggested to regulate the expression and activity of ECM components, such as matrix metalloproteinases [20]. The critical role of *DDR2* in development is also highlighted by its association with several developmental disorders, including spondylo-meta-epiphyseal dysplasia (SMED; MIM 271665), a type of autosomal-recessive dwarfism characterized by vertebral, epiphyseal, and metaphyseal abnormalities, often accompanied by cleft palate [21, 22]. Additionally, *DDR2* mutations have been linked to Warburg-Cinotti syndrome, which features skeletal and orofacial malformations, including high-arched palate [23]. Studies in *Ddr2*-deficient mice have demonstrated impaired anterior–posterior skull growth and frontal suture and bone formation, highlighting the critical role of *DDR2* in craniofacial development [19]. Therefore, we prioritized *DDR2* for further investigation. Additionally, protein–protein-interaction analysis using STRING (v12.0, https://string-db.org/) revealed interactions of DDR2 with proteins, including CDH1, CDH2 and COL11A2 (*SI Appendix*, Fig. S2), which are associated with OFC [24–26].

DDR2 structure and variant localization

Analysis of *DDR2* gene and protein structure (Fig. 1G–I) showed that one variant in the index family, c.518C > G (S173C), is located within the region encoding the discoidin domain of the extracellular domain. By contrast, the other variant, c.1352C > T (R451H), is situated in the cytoplasmic domain. Previously reported activating variants associated with Warburg-Cinotti syndrome are all located within the protein kinase domain of the cytoplasmic domain [23]; however, variants

Figures and Tables

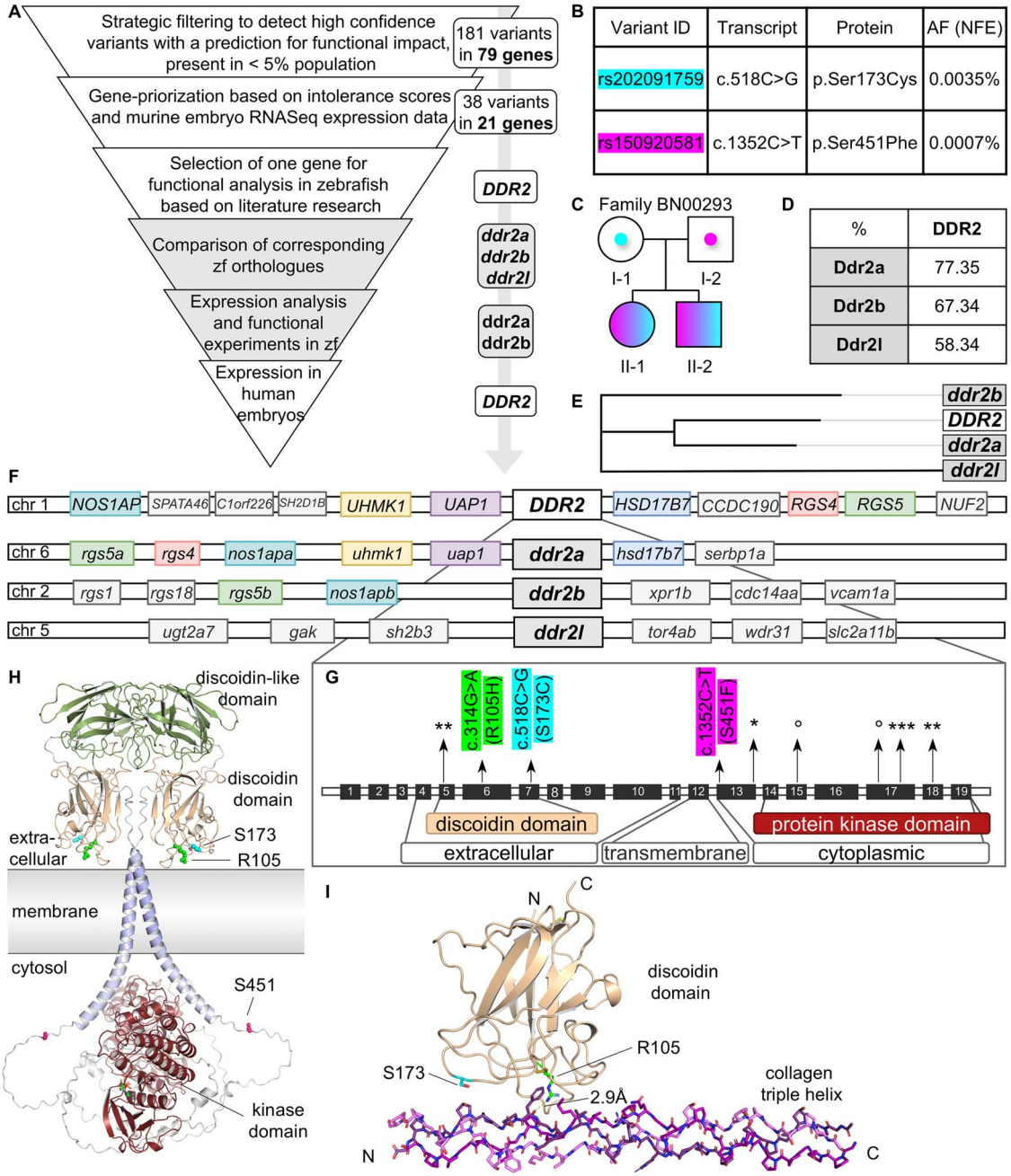

Figure 1 Selection of a suitable nsCPO candidate gene for analysis in zebrafish models. (A) workflow from exome sequencing (ES) reanalysis of data from nsCPO sib-pairs (white) to functional experiments in the zebrafish model (grey). (B) *DDR2* variants in our index family (BN00293). Each parent carried a different variant in *DDR2*. Allele frequencies (AF) were obtained from GnomAD v.4.1.0. (C) index family pedigree. The parents were unaffected, while both children have nsCPO and are compound-heterozygotes for the parental variants. (D–F) comparative analysis of human *DDR2* and zebrafish *ddr2a*, *ddr2b*, and *ddr2l*. Comparison of alignment of Ddr2a, Ddr2b, and Ddr2l with the human DDR2 amino acid sequence (D) and phylogenetic relationships (E). Comparison of human *DDR2* and zebrafish *ddr2a*, *ddr2b*, and *ddr2l* genetic loci: Corresponding genes are presented in the same color (F). (G) human *DDR2* gene structure: Exons are shown approximately to scale, relative to the entire gene. Protein domains are marked according to coding exons. Previously reported variants are indicated with arrows in their corresponding exons: * and °, variants associated with SMED and Warburg-Cinotti syndrome, respectively. The two variants identified in our study are shown (c.518C>G, c.1352C>T), as well as one further variant potentially causing a cleft palate from the DECIPHER database (c.314G>A). For further information refer to the text and the appendix (*SI Appendix*, Table S2). (H) localization of three potential causative variants in the DDR2 protein. AlphaFold3 model of a human DDR2 dimer, based on individual modeling of the N-terminal discoidin (29–185), discoidin-like (190–370), transmembrane helix (400–420), and C-terminal tyrosine-kinase (563–849) domains. Positions of the three variant sites are labeled (R105, S173, S451). (I) variants R105H and S173C in the discoidin domain are in the collagen triple-helix binding interface. R105 forms a hydrogen-bond with a hydroxyproline, which cannot be substituted by the shorter histidine sidechain. S173 is close to two prolines of the collagen-helix and to a disulfide bond (C177–C73), which stabilizes the extracellular domain. Hydrogen residues are red, nitrogen blue. N, N-term; C, C-term.

suggested to cause SMED have been described in the discoidin, protein kinase, and intracellular domains [27].

Visualization of the variants detected in this study using a DDR2 protein AlphaFold3 model (Fig. 1H) revealed that the S173C variant is in direct proximity to the collagen-binding interface, while the intracellular variant, R451H, is close to the transmembrane helix of the receptor. Research in the DECIPHER database (https://www.deciphergenomics.org/) revealed another patient with a cleft palate, carrying a dominant missense variant in *DDR2*, c.314G > A (R105H), which directly binds the collagen triple-helix (Fig. 1I). In an additional DDR2-specific ES analysis of 90 nsCPO families from the Bonn cohort (n = 314 individuals; mean target coverageindex = 121×; mean target coveragerelatives = 70×), including 73 trios and 17 multiplex families (8 quattros and 9 families with 5–8 sequenced members), we have not identified any further recessive variants in DDR2. For more detailed listing of variant numbers in the different filtering steps, refer to the supplementary information (SI Appendix, Fig. S3).

Zebrafish Orthologs of human *DDR2*

The zebrafish is a well-established model of craniofacial development [28, 29] and we identified three different zebrafish *DDR2*-orthologs: *ddr2a*, *ddr2b*, and *ddr2-like* (*ddr2l*). To determine the most relevant zebrafish orthologs to human *DDR2*, we conducted comprehensive analyses of phylogenetic relationships (Fig. 1E), percentage of identity in amino acid alignment (Fig. 1D), and genomic context (Fig. 1F). Phylogenetic analysis of the *DDR2* gene family revealed a complex evolutionary history, with *ddr2a* and *ddr2b* identified as paralogs that arose during the third teleost-specific whole-genome duplication. By contrast, *ddr2l* originated from a distinct duplication event, likely an earlier round of vertebrate whole-genome duplication, and was subsequently lost in mammals [30]. Analysis using the Uniprot align tool (https://uniprot.org/) showed that Ddr2a has the highest degree of similarity to DDR2 at the amino acid level, followed by Ddr2b. Examining the genomic contexts of *DDR2* and its zebrafish orthologs revealed that the *ddr2a* locus exhibits a high level of conservation with the human *DDR2* gene; 6 of the 10 genes closest to *DDR2* are also located near *ddr2a*. The *ddr2b* locus also showed a certain degree of similarity, with 2 of the closest genes to *DDR2* being present. However, there was no synteny between *DDR2* and *ddr2l* (Fig. 1F). Together, these findings led us to conclude that *ddr2a* and *ddr2b* (subsequently referred to as *DDR2*-orthologs) are the most significant orthologs, while *ddr2l* was excluded from further consideration.

Expression of *ddr2a* and *ddr2b* in zebrafish larvae

In situ hybridization

To investigate *ddr2a* and *ddr2b* mRNA expression patterns in zebrafish larvae (zfl), we conducted whole-mount in situ hybridization at various time-points between 1 and 4 days post fertilization (dpf). Throughout this developmental period, expression of both *ddr2a* and *ddr2b* was detected exclusively in the head region. At 3 dpf, the expression patterns of both orthologs were similar (Fig. 2A–C). In the lateral view, intense expression was observed projecting into the hyosympletic, Meckel's, and palatoquadrate cartilage regions. Expression posterior to the eyes became compact and localized to the hyosympletic cartilage. For *ddr2b*, intense staining was detected in the anterior portions of the fused ethmoid plate and ceratobranchial cartilage regions. For other time-points and more details, refer to the appendix (*SI Appendix*, Fig. S4).

Immunofluorescence

Further investigation of expression on protein level at larval stage 4 dpf using antibody co-staining of DDR2-orthologs and Collagen type II (indicating cartilage), revealed a stranded pattern of *DDR2*-ortholog expression in the lower jaw, corresponding to muscular structures (Fig. 2D and G). A similar stranded staining pattern of DDR2-orthologs was observed in the upper jaw region and ethmoid plate (Fig. 2E and G). Magnification of the oral side of the viscero-cranium revealed muscle strands expressing *DDR2* in direct connection with the ethmoid plate cartilage (Fig. 2F).

Single-cell RNA-sequencing

Single-Cell RNA-Sequencing (scRNASeq) expression data (Zebrahub, https://zebrahub.sf.czbiohub.org/) revealed that *ddr2a* and *ddr2b* were expressed in cell clusters important for ethmoid plate and jaw development at relevant time-points (3 and 5 dpf); *ddr2a* presented a broader expression pattern than *ddr2b* (*SI Appendix*, Figs S5 and S6). At 3 dpf, *ddr2a* was mainly expressed in neural crest, head mesenchyme, pharyngeal arch, pectoral fin, pectoral fin cartilage, and mesenchymal cells, while *ddr2b* expression was limited to neural crest, head mesenchyme, and pharyngeal arch cell clusters (*SI Appendix*, Fig. S5). The same pattern was detected at 5 dpf. While *ddr2b* was expressed in clusters highly specific to ethmoid plate and jaw development (cartilage element, cranial cartilage, pharyngeal arch), *ddr2a* was additionally expressed in mesoderm, myotome, and myoblast (*SI Appendix*, Fig. S6).

Loss-of-function of Ddr2a and Ddr2b in Zfl

To investigate the functional consequences of Ddr2a loss in zfl, we employed both Morpholino (MO)-mediated knock-down (KD) of Ddr2a and CRISPR/Cas9 (CRISPR)-mediated F0-knock-out (KO) of *ddr2a*, which both resulted in significant craniofacial abnormalities (Fig. 3A and B). CRISPR/Cas9-mediated *ddr2a* KO in *ddr2b* germline-KO (*ddr2a/b* double-KO) zfl was also performed (Fig. 3C). KD and KO zfl were always evaluated in comparison with control-MO (ctrl-MO) injected and scrambled-control (scr-ctrl) injected littermates, as well as uninjected littermates. To further examine the phenotypic effects of Ddr2a and Ddr2b loss, we inspected Alcian blue cartilage-stained zfl at 4 dpf and assigned them to two groups: one presenting with anomalies of the ethmoid plate or trabeculae, the other presenting no such anomalies. This analysis revealed a highly significant increase in craniofacial abnormalities in Ddr2a-deficient zfl (Fig. 3D, left). Specifically, 45.4% of Ddr2a MO-KD zfl exhibited an abnormal craniofacial phenotype ($P = 0.0067$). Similarly, 32.6% of *ddr2a* CRISPR F0-KOs showed an abnormal craniofacial phenotype ($P = 0.0044$). In *ddr2b* germline-KO (*ddr2b* KO) zfl, 16.3% of craniofacial anomalies were observed, relative to 2.9% in wildtype littermates ($P = 0.0442$). However, *ddr2a/b* double-KOs had a similar number of craniofacial anomalies to Ddr2a MO-KD zfl, at 45.3% relative to 17.4% in scr-ctrl injected *ddr2b* KOs (not significant: $P = 0.0528$; not shown in figure) and 16.3% in uninjected *ddr2b* KOs ($P = 0.0408$). No further obvious developmental anomalies were detected. Slightly higher mortality was observed in both control-injected and Ddr2a KD/KO zfl (Fig. 3D, right). The highest mortality rates were observed in *ddr2b* KOs and *ddr2a/b* double-KOs, with < 50% survival after 1 dpf. For CRISPR F0-KO verification, we performed PCR of targeted *ddr2a* and *ddr2b* exons 2–4 using 4 dpf zfl DNA samples from each group (Fig. 3E); CRISPR-gene-KO was successful in 92.9% of tested zfl (overall n = 100).

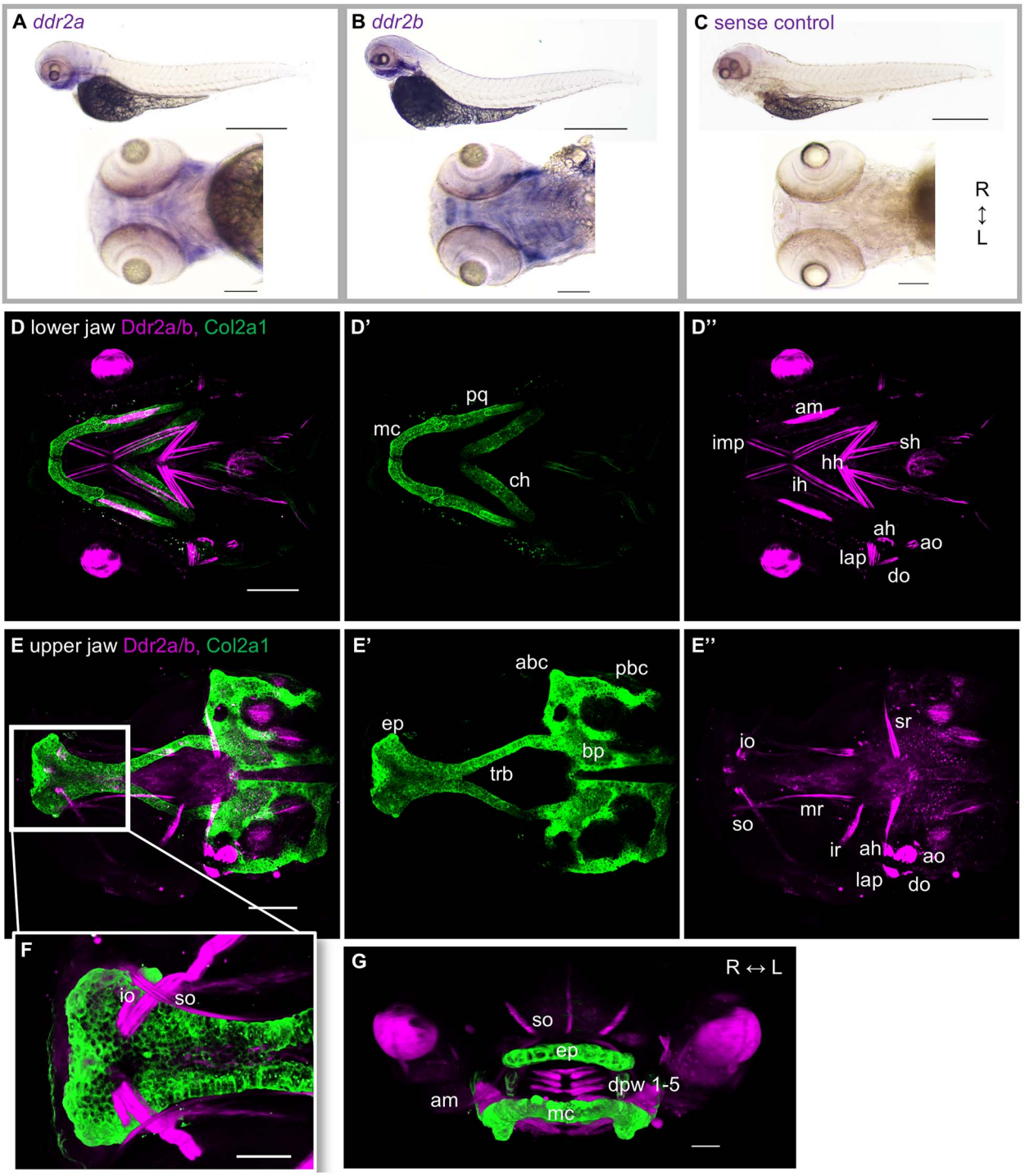

Figure 2 Analysis of *ddr2a* and *ddr2b* expression in wildtype zebrafish larvae (zfl). (A–C) RNA in situ hybridization at 3 days post fertilization (dpf). Lateral view of whole-mount zfl and dorsal view of the head (R, right; L, left). N = 3 independent experiments per group. Scale bars: Whole-mount zfl, 500 μm; head images, 100 μm. (C) sense control: *ddr2a* negative control. No expression was detected in negative control samples. (D–G) Immunofluorescence analysis of 4 dpf zfl with anti-DDR2 (expected to target all DDR2-orthologs) and anti-Collagen2a1. Visible cartilages are: Mc, Meckel's; pq, palatoquadrate; ch, ceratohyal; ep, ethmoid plate; trb, trabeculae; abc/pbc, anterior and posterior basicranial commissures; bp, basal plate. Visible muscles are: Imp, intermandibularis posterior; am, adductor mandibulae; ih, interhyals; hh, hyohyals; sh, sternohyoideus; lap, levator arcus palatini; ah, adductor hyoideus; ao, adductor operculi; do, dilator operculi; io/so, inferior/superior oblique; mr/ir/sr, medial/inferior/superior rectus; dpw, dorsal pharyngeal wall. (D) dorsal view of the zfl lower jaw. A stranded pattern of *DDR2*-ortholog expression is visible, corresponding to the developing muscle structures of the lower jaw (D''). Scale bar, 100 μm. (E) dorsal view of the zfl upper jaw, including the ethmoid plate. In this plain, stranded staining of DDR2-orthologs is also visible (E''). Scale bar, 100 μm. (F) high-magnification oral view of the ethmoid plate region; muscle strands close to the ethmoid plate, corresponding to the inferior and superior oblique muscles, express DDR2-orthologs. A distinct signal in the posterior portion of the ethmoid plate is also visible. Scale bar, 50 μm. (G) frontal view of the zfl upper and lower jaw (R, right; L, left). Scale bar, 100 μm.

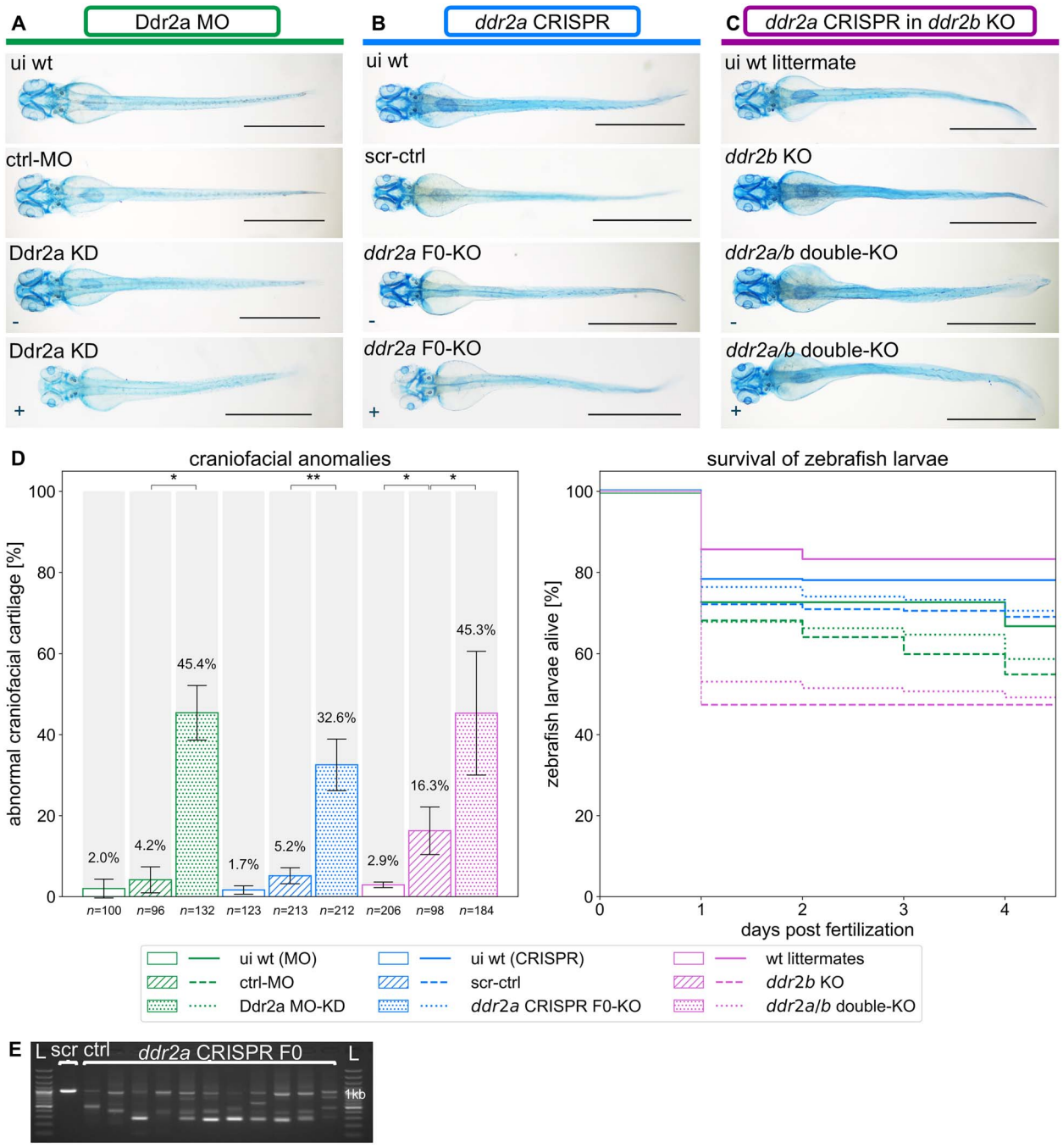

Figure 3 Phenotype classification of Alcian blue-stained zebrafish larvae (zfl). N = 3 independent experiments per group. (A–C) dorsal view of whole-mount zfl at 4 dpf. Uninjected and control-injected littermates are presented alongside the KD and KO for each experimental group: Ddr2a Morpholino-knock-down (MO-KD) and controls (ctrl-MO); *ddr2a* CRISPR F0-knock-out (CRISPR F0-KO) and controls (scr-ctrl); *ddr2a/b* double-KO and uninjected *ddr2b* KOs, alongside with their wildtype littermates. Zfl were separated into two categories: Those showing abnormal structures of the ethmoid plate and/or trabeculae (+) and those not showing any clear craniofacial alterations (−). Scale bars, 1000 μm. (D) left: Bar-graph showing the statistical occurrence of an abnormal craniofacial cartilage in different KD and KO zfl compared with uninjected and control-injected littermates. Experimental groups, as described in (A-C). The percentage of detected anomalies is given above the corresponding bar. $^{*}P < 0.05$, $^{**}P < 0.005$, $^{***}P < 0.0005$. Error bars represent the standard deviation. n, number of zfl. Right, Kaplan-Meyer graph displaying zfl survival curves from fertilization to 4 dpf. (E) PCR analysis of genomic DNA from *ddr2a* CRISPR and scr-ctrl injected zfl at 4 dpf to validate the CRISPR F0-KO. The expected PCR product in an undisrupted gene is approximately 1 kb; in case of successful CRISPR F0-KO, bands should be of unspecified different mobilities. L, ladder.

We observed severely altered craniofacial structures in KD and KO zfl, specifically concerning the ethmoid plate, which corresponds to the human hard palate, and the connected paired trabeculae. The cartilage of the ethmoid plate was often shorter and the trabeculae appeared warped, relative to zfl from the control injection groups (Fig. 4C, E, G).

To quantify the effects of Ddr2a MO-KD, *ddr2a* CRISPR F0-KO, *ddr2b* KO, and *ddr2a/b* double-KO on larval morphology, we performed

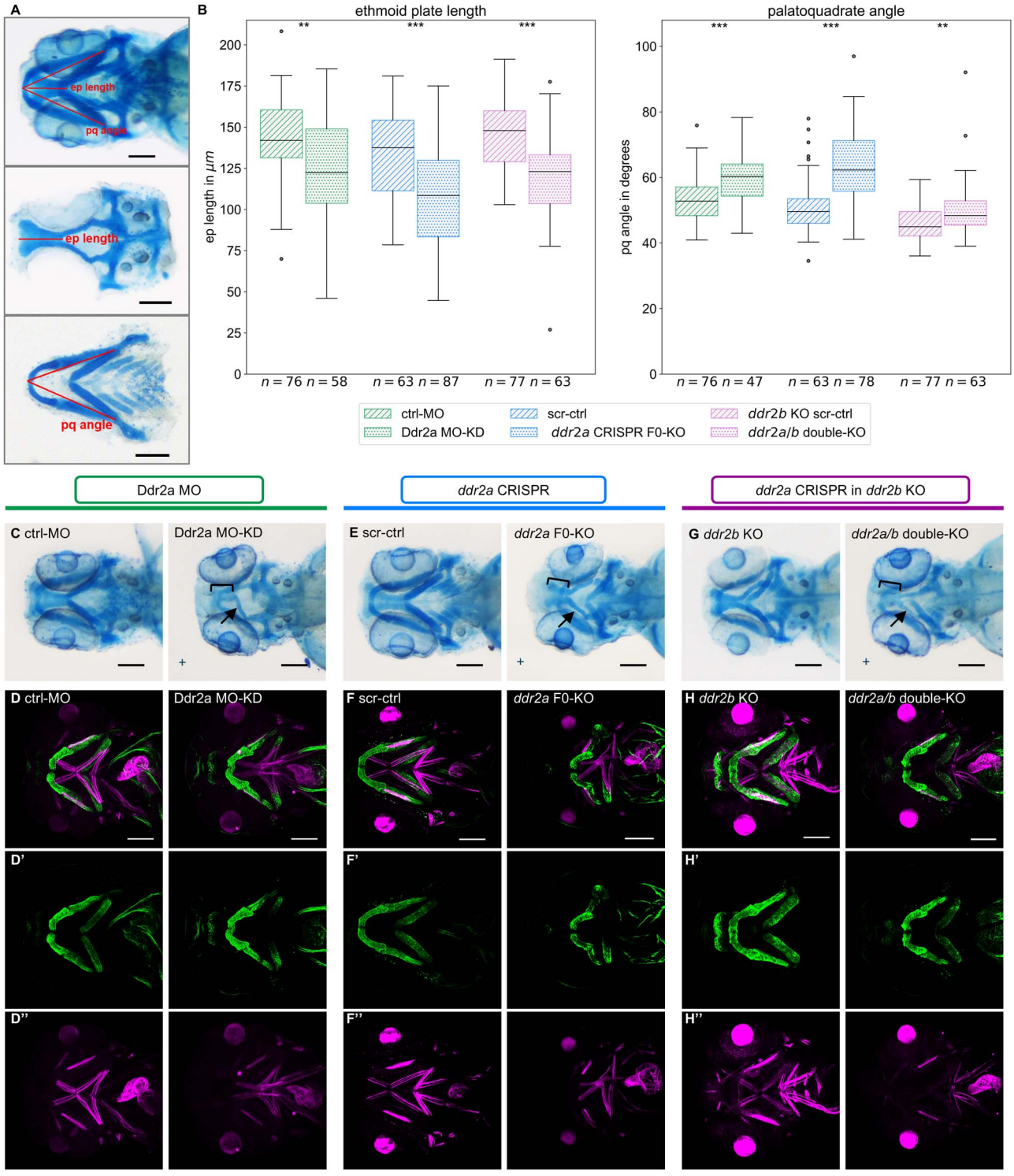

Figure 4 Phenotype evaluation of Alcian blue and immunofluorescence (IF) staining of the neuro- and viscero-cranium of various zebrafish larvae (zfl) groups. All zfl were at 4 dpf. N = 3 independent experiments per group. Scale bars, 100 μm. (A) overview of morphometric measurements of the upper and lower jaw structures of 4 dpf zfl: Ethmoid plate length (ep length) and palatoquadrate angle (pq angle) measurement lines presented in red in both whole-mount zfl head and separated viscero- and neuro-cranium. (B) box-plots of ep length and pq angle morphometric measurements. $^{*}P < 0.05$, $^{**}P < 0.005$ $^{***}P < 0.0005$. n, number of zfl. (C–H) representative examples of observed phenotypes, stained with Alcian blue (whole-mount heads) and IF (lower jaw structures): anti-DDR2; cartilage (anti-Col2a1). (C, D): Ddr2a Morpholino-knock-down (Ddr2a MO-KD) zfl, and controls (ctrl-MO); (E, F): *ddr2a* CRISPR knock-out (*ddr2a* CRISPR F0-KO) zfl and controls (scr-ctrl); (G, H): *ddr2b* KO zfl injected with *ddr2a* CRISPR, resulting in a *ddr2a/ddr2b* double-KO, and controls (scr-ctrl). Arrows indicate warped trabeculae; black brackets indicate a shortened ethmoid plate.

morphometric measurements of ethmoid plate length and the angle between the two palatoquadrates (Fig. 4A). Our analysis revealed significant differences in both parameters between Ddr2a MO-KD, *ddr2a* CRISPR F0-KO, and *ddr2a/b* double-KOs and their corresponding controls. Ethmoid plate length was significantly decreased in Ddr2a MO-KD ($P = 0.0022$), *ddr2a* CRISPR F0-KO zfl ($P = 1.01e-6$), and *ddr2a/b* double-KOs ($P = 1.03e-9$), relative to their respective controls (Fig. 4B, left). Additional quantitative morphometric analyses of the overall body length or head height revealed no significant differences in between Ddr2a MO-KD larvae and their controls. In contrast, *ddr2a* CRISPR F0-KOs and *ddr2a/b* double-KOs exhibited significantly reduced body length and head height, indicating a generally smaller embryonic size. To account for potential effects of generally delayed development and smaller body size on the craniofacial skeletons in the loss-of-function groups, we normalized ethmoid plate length to total zfl body length; p-values remained < 0.015 for all groups (*SI Appendix*, Fig. S7). The palatoquadrate angle was significantly increased in Ddr2a MO-KD ($P = 7.48e-5$), *ddr2a* CRISPR F0-KO ($P = 7.01e-13$), and *ddr2a/b* double-KO ($P = 0.0023$) zfl (Fig. 4B, right).

Furthermore, we performed immunofluorescence analysis of Ddr2 and Col2a, corresponding to the expression analysis presented in Fig. 2, for all KO and KD zfl, to evaluate development of the lower jaw and its adjacent muscular structures in Ddr2a and Ddr2b loss-of-function zfl (Fig. 4D, F, H). Severe alteration of muscle strand patterning was observed in various muscles, including the interhyal, hyohyal, and sternohyoideus muscles, in both Ddr2a MO-KD and *ddr2a* CRISPR F0-KO relative to controls (Fig. 4D" and E"). Zfl with KO of both *ddr2a* and *ddr2b* presented with much weaker overall staining for DDR2-orthologs on immunofluorescence analysis (Fig. 4H"), and the interhyals, hyohyals, and sternohyoideus muscles were undetectable relative to controls.

Craniofacial *DDR2* expression in human and mouse embryos

Next, we analyzed *DDR2* expression patterns using human embryo sections, assessing crucial points in human palate development, including: outgrowth of the palatal processes (Carnegie-Stage (CS) 18), elevation of palatal shelves (CS21), and fusion of bilateral palatal shelves (9 weeks post conception (PCW)). *DDR2* expression was detected at all stages. At CS18, *DDR2* expression was widespread, including in crucial orofacial structures, such as the emerging palatal shelves, Meckel's cartilage, and the tongue (Fig. 5A). Published murine expression data [19] and scRNASeq data from E11.5 mouse embryo lambdoid junctions [31] revealed that *DDR2* is expressed in head embryonic mesenchyme, while our analysis of scRNASeq data from human embryo facial tissue at CS12–16 [31, 32] further specified the expression to the frontonasal mesenchyme, as well as the anterior and posterior presomitic mesoderm (Fig. 5D). Interestingly, in the human embryo sections, *DDR2* was not detected in epithelial cells of the oral cavity at any stage examined. This finding was replicated by analysis of our scRNASeq data from both human and mouse embryos. As development proceeded, *DDR2* expression patterns in human CS21 embryo sections became more concentrated in the subepithelial tissue of the oral cavity and very strongly focused in the tissue around the emerging parotid gland (Fig. 5B). In the tongue, *DDR2* was detected in small cells with round nuclei, presumably myoblasts evolving from mesenchyme, but not in cells elongating into muscle fibers, with transverse striation (Fig. 5B"'). Our human embryo scRNASeq data confirmed *DDR2* expression in head muscle progenitor cells. After fusion of the palatal shelves at 9 PCW, *DDR2* was clearly localized in developing bone structures, such as the mandibula, as well as subepithelial palatal structures (Fig. 5C). Further, *DDR2* was still visible in some progenitor cells in tongue muscle, but not at all in muscle fibers, which were now clearly distinguishable by their transverse striation (Fig. 5C"').

Bulk RNASeq of ddr2a CRISPR F0-KO zfl

In our bulk RNASeq of 4 dpf zfl heads of *ddr2a* F0-KO zfl, scr-ctrl and uninjected zfl to assess the impact of *DDR2* loss-of-function we obtained an average of 18.1 M reads per sample (> 90% uniquely mapped). In the Principal Component Analysis, two outliers could be detected and were excluded from further analysis. The remaining samples showed clear separation between different experimental groups (*SI Appendix*, Fig. S8). The analysis revealed differential expression of 18 665 genes, 2193 of which were significant (padj < 0.05). To assess ECM genes disproportionately affected by *DDR2* loss-of-function, absolute Gene Set Enrichment Analysis (GSEA) was performed using the zebrafish Matrisome dataset [33] composing of 1015 ECM-related genes, of which 693 were detected in our dataset. We obtained an overall enrichment score (ES) of 0.7755 for the *ddr2a* F0-KO group, compared to scr-ctrl ($p = 0.049$). In comparison, scr-ctrl versus uninjected zfl showed no significant effect ($p = 0.876$), with an ES of 0.5968. From the Matrisome gene list, 85 genes showed differential expression, 71 of those with a significant effect (padj < 0.05; SI Appendix, Table S4). Among the 71 significantly dysregulated leading-edge genes (Fig. 6), 62 were down-regulated, including ADAM metallopeptidase with thrombospondin type 1, motif 13 (*adamts13*), ECM protein *ecm1a*, and proteoglycan *prg4b*, a component localized to articular cartilage. In contrast, 9 up-regulated genes included Serpin Family H Member 1 (*serpinh1b*), which is involved in collagen biosynthesis, and Cathepsin L (*ctslb*), associated with collagen and elastin degradation.

Discussion

The reanalysis of ES data from six sib-pairs affected with nsCPO revealed interesting recessive variants in genes potentially implicated in craniofacial development or related developmental pathways (e.g. *DDR2*, *PRDM1*, *PDLIM5*, *ACACB*; *SI Appendix*, Table S1). In this study, we prioritized *DDR2*, however, the other genes also represent promising candidates for future genetic and experimental analyses.

In the ES data, two siblings (family BN00293) were of particular note, as both had not only compound-heterozygous variants in the known clefting gene, *GRHL3*, but also in *DDR2*. Truncating dominant *GRHL3* variants with incomplete penetrance that lead to nsCPO have been described [11]. Since the novel in-frame deletion in *GRHL3* (c.1722_1724del) is not truncating, it appears unlikely that it has such an effect. The other *GRHL3* variant (c.1361C > T) is a common risk variant (rs41268753) in the coding region with a minor allele frequency of 3.5% in the European population and an effect size (odds ratio) of 2.46 only [11, 34], and thus may contribute to the polygenic background. Previously reported variants in *DDR2* have been established as causal for autosomal-recessive SMED [27], while gain-of-function *DDR2* variants cause the rare autosomal-dominant inherited Warburg-Cinotti syndrome. In addition to various other phenotypes, individuals with Warburg-Cinotti syndrome can exhibit a high-arched

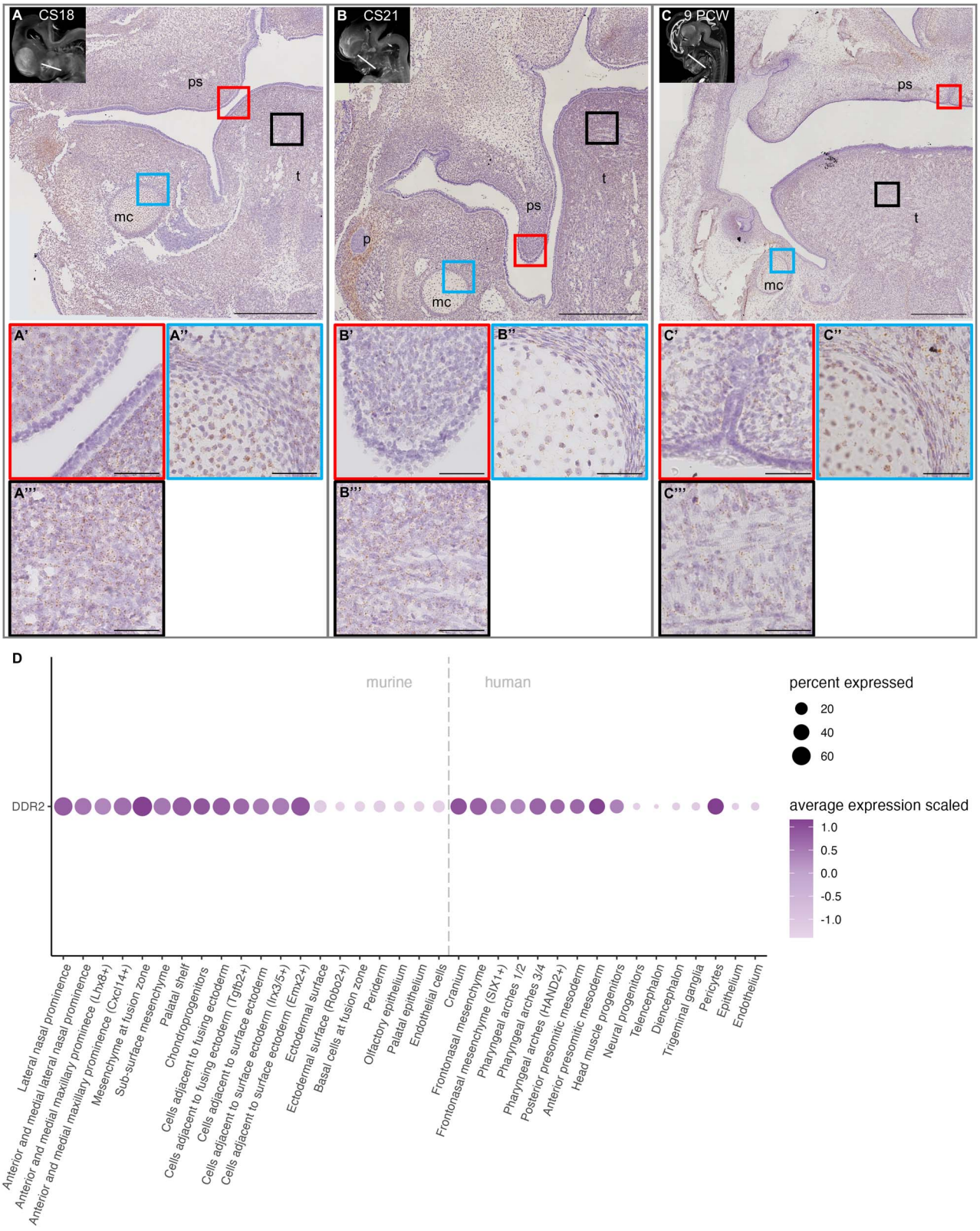

Figure 5 *DDR2* expression analysis in human and murine embryos. (A–C) RNA-scope-in-situ (brown staining) of cranial human embryo sections at Carnegie-stage (CS) 18 and 21, as well as 9 weeks post conception (PCW). The approximate plain of sections is marked with a white line in the upper left corner. For each image, magnified palatal shelf (ps; A′, B′, C′), Meckel's (mc; A″, B″, C″), and tongue (t; A‴, B‴, C‴) areas are shown. Scale bars: Overviews, 500 μm; magnified regions, 50 μm. (D) Dotplot of *DDR2* gene expression in the lambdoid junction of embryonic day 11.5 murine embryos and in the head of CS12–16 human embryos. The size of the dots corresponds to the percentage of cells expressing the gene in the respective cell type. Dot color intensity represents the average scaled expression level.

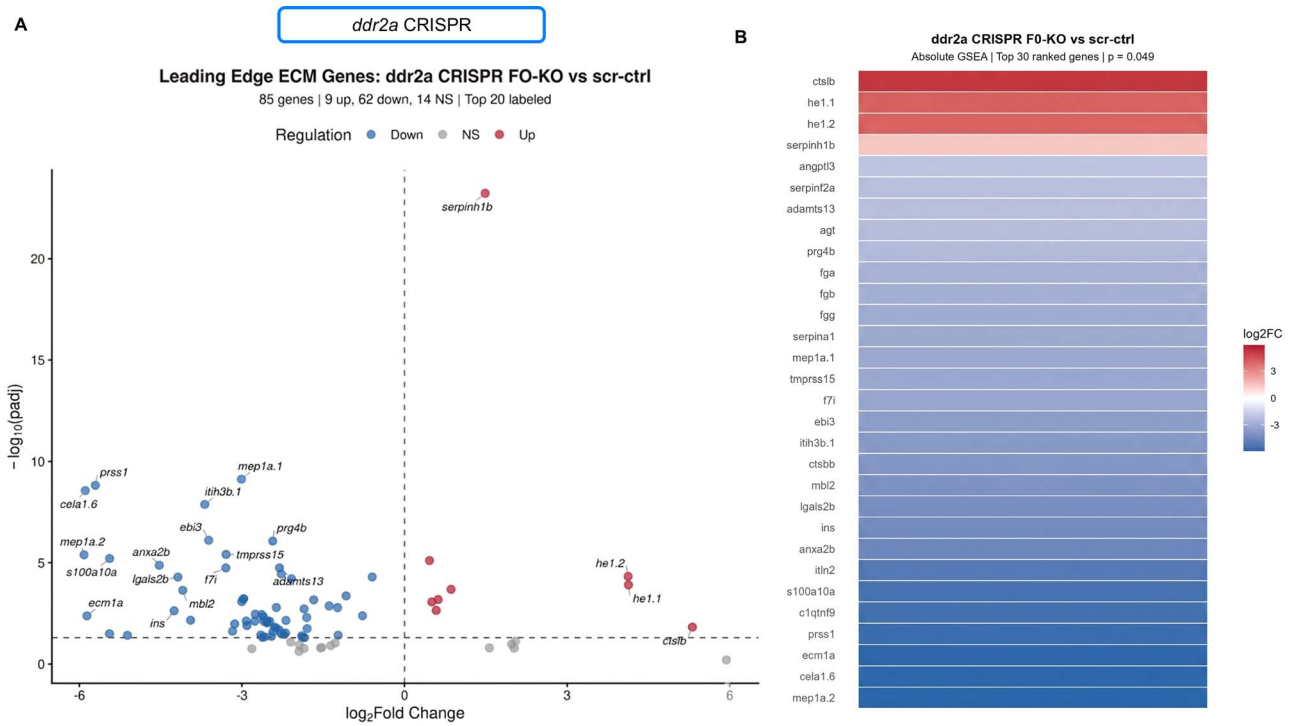

Figure 6 Differential expression of leading-edge extracellular matrix (ECM) genes in 4 dpf *ddr2a* CRISPR-F0 zebrafish larvae (zfl) heads compared with scrambled controls (scr-ctrl). (A) volcano plot showing $\log_2$ fold change versus $-\log_{10}$ adjusted p-value for ECM-related leading-edge genes identified by absolute gene-set enrichment analysis (GSEA). Each dot represents one gene. The dashed vertical line marks no change in expression, and the horizontal dashed line indicates the significance cutoff ($p_{adj} < 0.05$). In total, 62 of 71 significantly dysregulated genes were down-regulated in ddr2a F0-KOs, 9 were up-regulated. (B) Heatmap of $\log_2$ fold changes for the top 30 leading-edge genes identified by absolute GSEA in the comparison of *ddr2a* F0-KO versus scrambled control ($P = 0.049$). Gene expression changes were derived from DESeq2 differential expression analysis. Colors indicate the direction and magnitude of the expression.

palate, which is a known cleft palate microform [23]. Importantly, both siblings had CPO and did not exhibit any further abnormalities suggestive of an underlying syndrome. Given the multifactorial etiology of nsCPO, it is conceivable that the combination of *GRHL3* and *DDR2* variants contribute significantly to development of the condition in these two affected individuals. In an additional, unpublished nsCPO exome dataset, we conducted a targeted analysis for *DDR2* and could not identify any further recessive variants, emphasizing the rarity of pathogenic variants in *DDR2* as the cause of nsCPO. However, the previously described findings highlight *DDR2* as a highly interesting nsCPO candidate gene warranting further investigation.

From a developmental perspective, *DDR2* has previously been described as crucial for craniofacial development [19]. As a cell-surface receptor responding to binding of fibrillar collagen, DDR2 influences cell differentiation, proliferation, and migration, as well as ECM composition and regulation [35], where cell and ECM interactions are crucial in shaping cranial bones [36]. Thus, DDR2 impairment may contribute to altered structural development of craniofacial bones, such as the outgrowth, elevation, and fusion of palatal shelves, leading to cleft palate, as observed in the sib-pair included in this study.

DDR2 belongs to the transmembrane receptor tyrosine-kinase family and has various domains. One of the identified variants (NM_006182.4:c.518C > G) is located within the extracellular discoidin domain, in direct proximity to the collagen-binding interface, and is very likely to alter DDR2 binding-properties and consequent receptor-activation (Fig. 1G and H). Supporting this theory, we identified another individual with CPO (DECIPHER v11.31, patient 376 276) with a *de novo*-missense variant in *DDR2* (NM_006182.4:c.314G > A) within the collagen-binding interface, directly interacting with the collagen-helix [37]. In contrast to our sib-pair, patient 376 276 presented with a syndromic phenotype, including anemia, corneal erosion, epileptic encephalopathy, metopic synostosis, primary hyperparathyroidism, and short stature. We believe that those variations in phenotypic expression could be well explained through different consequences of gene variants on protein structure and activity, due to alterations in amino acid interactions and folding. The other variant we identified (NM_006182.4:c.1352C > T), is located in the intracellular portion of DDR2 (Fig. 1F, *SI Appendix*, Table S2), close to the transmembrane helix. The functional consequences of these variants at the protein level remain unclear; an influence on DDR2 dimerization or its kinase activity are plausible, as both would result in altered signal transduction, a mechanism seen in patients with variants causing Warburg-Cinotti syndrome [23].

To further validate and functionally characterize *DDR2* as a candidate gene for nsCPO, we employed zebrafish models. The specific expression patterns of *DDR2*-orthologs we observed in the zebrafish craniofacial region confirmed *DDR2* as a candidate gene warranting further analysis (Fig. 2A–C). Further, co-immunostaining of DDR2 and collagen II showed muscle fiber-specific localization of DDR2-orthologs (Fig. 2D–G). Publicly available scRNASeq data (Zebrahub) allowed us to distinguish between the *DDR2*-orthologs and revealed clear differences in local expression patterns, as well as significant changes in expression across developmental stages (*SI Appendix*,

Figs S5 and S6). However, the Zebrahub data showed strong general expression of both *ddr2a* and *ddr2b* in relevant palatal tissues and their progenitors, such as neural crest, head mesenchyme, and pharyngeal arches, at time-points relevant to ethmoid plate development, with *ddr2a* showing a broader expression pattern than *ddr2b* (*SI Appendix*, Figs S5 and S6). Discrepancies in expression patterns among different analysis methods may be attributable to technical variations and/or inconsistencies in developmental staging.

KD and KO of *DDR2*-orthologs in zebrafish resulted in malformations of regions equivalent to the palate and jaw, somewhat resembling the cleft palate phenotype in our patients (Figs 3 and 4). We note that the effects of zebrafish gene KD or KO are not directly corresponding to those of compound-heterozygous missense variants; however, incomplete KD or KO of only one *DDR2*-ortholog could simulate our *DDR2* variants with a low effect size, leading to a phenotype equivalent to nsCPO.

Up to half of our *DDR2*-ortholog KOs, as well as KD zebrafish morphants, presented with similar phenotypes: trabeculae and ethmoid plate (equivalents of the human palate) were severely malformed or zfl were even partly agnathic (Figs 3D and 4).

Histologically, we observed severe disorganization of DDR2-expressing muscle fibers in the jaw regions of the various *DDR2*-ortholog KD and KO zebrafish morphants generated in this study (Fig. 4D, F, H). Mechanical stretch, produced at musculo-cartilagenous attachment points drifting apart due to cartilage growth has been demonstrated to guide the orientation of adhering myocytes and induce cell fusion and muscle growth in a zebrafish model [38]. Interestingly, in humans, palate elevation and fusion occur at around the same time as first facial muscle contractions, leading to fetal swallowing and grimacing [39]. Furthermore, lack of tongue and masseter muscle movements can cause cleft palate [39, 40]. To investigate this hypothesis, we analyzed *DDR2* expression in human and murine embryonic orofacial muscle progenitor tissues using scRNASeq data (Fig. 5D). Further, we examined human embryonic cranial RNA-Scope-in-situ sections and observed *DDR2* expression at embryonic stages prior to palatal fusion in muscle progenitor tissue, but not in fully differentiated transverse striated muscle. After fusion, as myoblasts differentiate further into clearly detectable muscle fibers, *DDR2* expression was no longer visible (Fig. 5A–C). This finding is distinct to the immunofluorescence results in 4 dpf zfl, where DDR2-orthologs were clearly detected in jaw muscle fibers. Yet, this observation is not necessarily inconsistent with the observation in human data. Craniofacial muscles in zfl are morphologically differentiated and capable of supporting feeding behavior by approximately 4 dpf; however, they continue to undergo maturation, structural remodeling, and subdivision into adult muscle units beyond this stage [41]. Thus, muscles at 4 dpf can be considered differentiated but not yet fully mature. In this context, our observations may reflect expression during a transitional phase spanning late progenitor and early differentiated states. Additionally, observed differences in expression may be explained by dissimilarities in palatogenesis between human and zebrafish. While palatogenesis in humans involves elevation of paired palatal shelves that grow vertically alongside the tongue before reorienting horizontally above it and then fusing at the midline through an epithelial seam, in zebrafish chondrocytes organize into two parallel rows between the eyes, forming the trabeculae, which eventually fuse at their anterior ends to form the ethmoid plate [42], without undergoing the characteristic reorientation observed during human palatogenesis. Nevertheless, previous studies have demonstrated zebrafish as a suitable model for human OFC development [28, 29]. Further analyses would be required to definitively determine whether DDR2's function is restricted to muscle progenitors or persists in mature muscle.

We propose that *DDR2*-expressing muscle progenitor cells influence the outgrowth, elevation, and fusion of the palatal shelves during human palatogenesis. This effect may be mediated either through collagen-induced intracellular signaling pathways that modulate ECM constitution [35] or even through direct mechanical interactions between DDR2 and the ECM. The first hypothesis is supported by our bulk RNASeq absolute GSEA results, which revealed significant differential expression of several zebrafish Matrisome-associated genes (Fig. 6), including the matrix metalloproteinase *adamts13* and the ECM protein *ecm1a*, both of which are critical regulators of ECM composition, as well as *serpinh1b*, involved in collagen biosynthesis, and *ctslb*, which contributes to collagen and elastin degradation. Dysregulation of these critical genes, potentially resulting from *DDR2* loss-of-function, may substantially alter ECM composition and thereby disrupt palatogenesis. Future studies employing high-resolution temporal expression profiling or lineage-tracing approaches will be required to more precisely define the temporal dynamics of DDR2 activity during muscle and/or palate development.

However, zfl may not be a suitable model for further detailed analysis of these mechanisms, due to the partial differences in *DDR2*-ortholog expression patterns through development, as well as dissimilarities between palatal shelf development and ethmoid plate shaping.

While known OFC genes, such as *CDH1*, *CTNND1*, *IRF6*, and *GRHL3*, act mainly via mechanisms related to epithelial integrity and periderm differentiation [34, 43, 44], we show here that *DDR2* is not expressed in epithelial cells and may contribute to OFC development via hampered ECM interactions, a mechanism that has been postulated previously [45].

Limitations of this study include the small number of individuals harboring *DDR2* variants across both analyzed ES cohorts. However, given the presumed rarity of recessive causative *DDR2* variants in nsCPO, this finding is not unexpected.

In addition, functional characterization of other candidate genes identified in the initial ES dataset was not the focus of this work and remains an important area for future investigation.

Another limitation arises from potential off-target effects associated with MO and CRISPR approaches, including mosaicism in developing larvae. These effects were mitigated through dose titration, exclusion of embryos with gross developmental abnormalities, quantitative morphometric analyses, and the observation of consistent phenotypes across both independent loss-of-function strategies. Nevertheless, future studies employing germline KO models are needed to further validate these findings.

Finally, although RNASeq analyses showed an impact of *DDR2* loss-of-function on ECM-regulating genes, the precise molecular interactions remain unresolved and highlight the need for further investigations.

In conclusion, *DDR2* is a strong candidate for nsCPO, as we clearly demonstrate a correlation of zfl phenotypes in loss-of-function experiments with those of human patients with *DDR2* variants. In human embryos, *DDR2* is present in crucial structures of the developing palate and is likely to be involved in its formation. Possible mechanisms include disruption of ECM remodeling through altered collagen-response and/or intracellular signaling mediated by DDR2, which is emphasized by the dysregulation of several crucial ECM-related

genes in the absolute GSEA of our bulk RNASeq data in zfl. Although more patient genetic data is needed to further elucidate genotype–phenotype correlations, *DDR2* should be considered a causative factor, not only in patients with severe complex syndromes, but also in individuals with mild syndromic CPO and nsCPO.

Materials and methods

Ethical approval

Written informed consent was obtained from all participants, or their legal guardians, prior to inclusion. The study was approved by the ethics committee of the Medical Faculty of the University of Bonn (Ethics approval number, 295/14, last dated June 20th, 2022). Animal husbandry and experimental setups were in accordance with European Legislation for the Protection of Animals used for Scientific Purposes (Directive 2010/62/EU). National law exempts all zebrafish experiments conducted at larval stages up to 5 dpf before zfl begin to feed from ethical approval. Human embryonic samples were obtained from the HDBR with appropriate written informed consent from the donor and ethical approval from the Fulham Research Ethics Committee (23/LO/0312). The HDBR is regulated by the UK Human Tissue Authority (https://hta.gov.uk/) and operates in accordance with the relevant Human Tissue Authority Codes of Practice.

Database

ES data from six families with nsCPO from the Bonn cohort, previously described by Hoebel et al. [15], were reanalyzed. To identify novel nsCPO candidate genes that may be recessively inherited, families with two affected siblings and healthy parents were selected. ES data from 12 individuals (six sib-pairs) were analyzed using Varbank 1.0 (https://varbank.ccg.uni-koeln.de/). To identify potential causal variants, strategic pipeline-internal filters were applied to each sib-pair data using the following criteria: (I) high-confidence variants with read coverage $\geq 15\times$ and quality score ≥ 20; (II) variants with predicted functional impact, including single nucleotide variants, single nucleotide polymorphisms, and InDels that resulted in an alteration in primary protein structure (missense or nonsense variants) or had a strong or medium splice site effect; (III) variants with a population frequency $\leq 5\%$ according to the 1000Genomes database (phase3 release 20 130 502); and (IV) variants appearing ≤ 25 times in the Varbank in-house epilepsy database and ≤ 30 times in the Varbank in-house structural database, to minimize pipeline-specific artifacts. For variant detection, two separate lists were generated: one for homozygous variants with an allele read frequency 75%–100%, and another for compound-heterozygous variants with frequency 25%–75%.

Variant annotation

After filtering, the remaining variants were annotated. Genotype frequencies in Non-Finnish European population data were obtained using the Genome Aggregation Database (GnomAD v.2.1.1). Variants with frequency $> 5\%$ were excluded from further analysis, while those with unknown frequencies were retained for further consideration. Remaining variants were visually inspected in Varbank Browse Reads to rule out technical artifacts and to verify compound heterozygosity. Missense variants were divided into three sub-groups: high, medium, and low priority, based on SIFT and PolyPhen-2 prediction scores, both generated using the Ensembl Variant Effect Predictor tool (https://ensembl.org/, GRCh37/hg19 reference genome) [46], and CADD v.1.4 (https://cadd.gs.washington.edu/). For further details refer to the appendix (*SI Appendix* Fig. S1).

Prioritization of candidate genes

To prioritize genes harboring putative disease-causing recessive variants, two key factors were considered: gene expression in murine embryonic facial tissue and gene intolerance to specific types of mutations. Data from the 'RNASeq Analysis in E13.5 Mouse Palates' dataset on FaceBase [16, 17] were leveraged. Murine gene orthologs were assigned using NCBI (https://ncbi.nlm.nih.gov/). Gene intolerance scores for missense mutations were obtained from GnomAD v.2.1.1. Regions of interest in prioritized gene variants from both affected individuals and their parents were validated by Sanger sequencing.

Selection of candidate genes for functional analysis in zebrafish

To prioritize genes for a functional follow-up in zebrafish (*Danio rerio*), an extensive literature search of PubMed was conducted using each gene name in combination with the keywords 'palate', 'orofacial', or 'development' as search terms. Genes involved in cell–cell communication, regulation, and migration were prioritized, while those without any associations with embryonic development were excluded from further consideration. Candidate genes were analyzed for protein–protein-interaction using STRING v.12.0 and the DECIPHER database searched for further patients with a cleft palate and variants in *DDR2*.

DDR2-specific analysis of ES data

To search for further recessive variants (homozygous and compound heterozygous) in *DDR2*, a targeted analysis of ES data from individuals with nsCPO has been performed. The ES dataset comprised 90 families (n = 314 individuals) from the Bonn cohort, including 73 trios (affected child born to healthy parents), eight quattros (affected sib-pairs born to heathy parents) and nine multiplex families with five to eight sequenced members. Exome capture was performed using Twist Human Core Exome Kit and sequencing was carried out on Illumina Nova Seq (2×100 bp). Reads were aligned to the GRCh37 reference genome using BWA implemented in Parabricks (GPU-accelerated). Variant calling was performed on a per-sample basis using Haplotype Caller, followed by joint genotyping across all samples using GLnexus to generate a cohort-level BCF file. Individual and family-level VCFs were subsequently derived. For variant annotation, filtering strategy and prioritization please refer to the supplementary information (*SI Appendix*, Fig. S3).

Modeling of *DDR2* gene and protein structures

To visualize the structural implications of variants identified in human *DDR2* in our patients, the DECIPHER database, and patients with Warburg-Cinotti syndrome and SMED, *DDR2* gene structure was modeled using UCSC genome browser with the human reference genome (https://genome-euro.ucsc.edu/,GRCh37/hg19). Additionally, AlphaFold3 was used to model human DDR2 structure, based on crystal structure 2WUH [47].

Assigning zebrafish Orthologs to human *DDR2*

The ZFIN database (https://zfin.org/) was used to investigate zebrafish orthologs of prioritized candidate genes. Genes with amino acid identity ≥70% between human and zebrafish orthologs were selected for possible functional analysis. For genes with multiple orthologs, each ortholog's genomic conservation was compared with human, considering amino acid alignment and phylogenetic tree, obtained using the UniProt align tool (https://uniprot.org/). Genomic contexts were also compared using the UCSC genome browser (reference genomes GRCh37/hg19 and GRCz11/danRer11). Expression data of candidate genes available through the ZFIN database were analyzed and those not expressed in the craniofacial region during zebrafish development excluded. Since there were no available zebrafish expression data for the highly prioritized gene *DDR2*, whole-mount in situ hybridization was performed in wildtype zfl to assess its expression pattern.

Zebrafish husbandry and embryo maintenance

All zebrafish strains and larvae used were maintained in the zebrafish core facility in Bonn, Germany following national law and recommendations by Westerfield [48]. KO *ddr2b*$^{+/-}$ zfl (sanger embryos, sa19669, item #23663) were obtained from Karlsruhe Institute of Technology. Adult *ddr2b* KO fish were genotyped using a PCR digestion protocol, with DNA assessed through skin-swabbing. New generations of all strains were obtained by natural spawning in the morning, and raised at 28°C in Danieau (30%) medium on a 14-h light/10-h dark cycle. After fertilization, embryos were maintained in Danieau's solution. After 24 h, 1-phenyl-2-thiourea was added (final concentration 0.003%) to inhibit pigmentation. All experiments including zebrafish were conducted prior to 5 dpf, before the start of independent feeding.

Analysis of *DDR2*-Ortholog expression in zebrafish

In situ hybridization

Whole-mount in situ hybridization against both *ddr2a* and *ddr2b* was performed on 1–4 dpf zfl. Specific probes for both *DDR2*-orthologs were designed to include untranslated regions (UTRs) to ensure specific targeting (*SI Appendix*, Table S3). For probe generation, cDNA was synthesized from 3 dpf larval tissue mRNA using iScript Reverse Transcription (Biorad; Cat. No. 1708840) and PCR-cloned into a pBlue-Script II SK vector. Successful cloning was verified by test digestion and DNA sequencing (performed by EurofinsGenomics; Sanger Sequencing; TubeSeq Supreme). Probe synthesis and RNA labeling with digoxigenin-UTP by in vitro transcription was conducted using T3 and T7 RNA polymerase and a DIG RNA Labeling Kit (Roche Applied Sciences; Cat. No. 11175025910). The high-resolution in situ hybridization protocol from Thisse lab [49] was used for *ddr2a* and *ddr2b* mRNA-labeling. Stained zfl were imaged using a Nikon Eclipse Ni microscope.

Immunofluorescence

Whole-mount antibody labeling was performed on 4 dpf zfl using antibodies against DDR2 (ab76967, Abcam) and anti-Col2a1-ab (CHX-7005, Biozol). Collected zfl were fixed in 4% PFA for 2 h at room temperature, washed with phosphate-buffered saline (PBS, pH 7.3), dehydrated with methanol, stored overnight at −20°C, then rehydrated with PBS + 0.1% Tween20. Antigen retrieval was induced by incubating in sodium citrate buffer (pH 6) at 70°C for 20 min. Samples were permeabilized with Proteinase K (10 μg/μl) for 50 min at room temperature, then preincubated and blocked with freshly prepared 10% goat serum, 2% bovine serum albumin, 0.1 M phosphate buffer +0.8% TritonX100 for 4 h at 4°C. Primary antibodies (anti-DDR2 (1:300) and anti-collagen type II (1:500)) were added and incubated for 3 days at 4°C. Samples treated with anti-DDR2 and anti-Col2a1 were then incubated with the secondary antibodies Alexa Fluor 546 goat anti-rabbit (ThermoFisher Scientific; Invitrogen A11035) and Alexa Fluor 488 goat anti-mouse (ThermoFisher Scientific; Invitrogen A11001), respectively, for 2 days at 4°C, avoiding light. After final washing and refixation with 4% PFA, zfl were mounted in 1.25% ultra-low-melting agarose (Sigma-Aldrich; A2576-5G) for imaging using a Nikon Eclipse Ti2 fluorescent microscope.

ScRNA-sequencing data

Expression of *ddr2a* and *ddr2b* at 3 and 5 dpf, representing the zfl period of craniofacial development, was evaluated using a publicly available scRNASeq Atlas of zfl development (Zebrahub) [50].

Loss-of-function experiments

KD of *Ddr2a* mRNA was performed using MO oligonucleotides targeted at the translational start site (GeneTools, LLC, Philomath, OR, US) (*SI Appendix*, Table S3); injection solution contained CutSmart buffer, 10% Phenol red, water, and MO stock solution (final concentration, 1 mM) and was heated to 65°C for 5 min before injection, to reduce self-complementary binding. A standard ctrl-MO, with no specific mRNA target, was used as a negative control. MO concentrations were empirically titrated to the lowest dose that produced reproducible phenotypes while minimizing nonspecific toxicity.

A genomic level *ddr2a* F0-KO was generated using the CRISPR/Cas9 approach. Five specific crRNAs, targeting exons 2, 3, 4, and 5, were designed using CRISPRscan [51] (*SI Appendix*, Table S3). Stock containing equal amounts (100 mM) of all crRNAs was combined with tracrRNA (100 mM) and heated to 95°C for 5 min to form gRNA. For the injection mix, Alt-R S.p. HiFi Cas9 nuclease was diluted 1:10 in Cas9 working buffer (20 mM HEPES; 150 mM KCL, pH 7.5), combined with the gRNA and Phenol red, then heated to 37°C for 10 min. A standard scr-ctrl, with no genomic target, was used as a negative control. Cas9 nuclease, crRNAs, and tracrRNA were purchased from Integrated DNA Technologies (IDT, Belgium). CRISPR injections were performed using a guide RNA and Cas9 concentration substantially lower than those typically reported in zebrafish studies [52], in order to reduce potential toxicity; formal dosage titration was not performed.

Injections (volume, approximately 1.7 nl) were performed into the yolk of one-cell staged embryos within 20 min after fertilization using an air-pressure micro-injection setup. Uninjected embryos served as controls for all injections. Post-injection care and monitoring of zfl was equal for each experimental group. Dead zfl in each group were counted and removed daily to obtain survival statistics.

Genomic PCR of the targeted exons 2–4 was performed to evaluate the success of the CRISPR KO (Fig. 3E). DNA was extracted from 10 representative 4 dpf zfl from each group, as previously described [53].

Phenotype assessment and morphometric measurements

At 4 dpf, zfl were subjected to Alcian blue cartilage staining, as previously described [54] and imaged using a ZEISS Stemi 508 stereomicroscope equipped with a Nikon DSFi2 camera. Zfl subjected to MO-KD or CRISPR F0-KO as well as their respective controls were systematically

screened for general developmental abnormalities. Phenotype evaluation was performed with particular focus on the ethmoid plate, trabeculae, Meckel's, and palatoquadrate cartilages and classified as either normal (−) or abnormal (+) in each experimental group (Fig. 3A–C). NIS Elements software (Nikon) was used to conduct morphometric measurements to quantify: (I) total body-length (mean of lateral and dorsal measurement, to avoid errors due to bending of the tail); (II) ethmoid plate length; and (III) palatoquadrate angle (Fig. 4A).

DDR2 expression analysis in human embryos

RNA-Scope staining was performed in human embryo sections of the head at CS18 and 21, as well as 9 PCW, in cooperation with Human Developmental Biology Resource (HDBR, https://hdbr.org/). For each time-point, one section localized in the rostral part of the palate, representing the hard palate, and one section in a more caudal plain, representing the soft muscular palate, were examined. Staged human embryonic and fetal tissue was fixed in formalin overnight, processed through graded alcohols and xylene changes, and embedded in paraffin. Tissue sections (7 μm) were cut on a microtome and collected on SuperFrost microscope slides. Slides were baked for 1 h at 60°C before paraffin was removed in Histo-Clear (National Diagnostics). RNAscope-in-situ hybridization was used to analyze the expression profile of *DDR2* using a commercially available probe (Hs-*DDR2*, #551011), RNAscope 2.5HD assay kits (Advanced Cell Diagnostics), and BROWN detection reagent. A negative control probe (#320751) was used to confirm specificity. Manufacturer-recommended protocols were followed, with slight modifications: incubation times for Amp 5 hybridization and DAB signal detection were increased to 35 and 15 min, respectively; sections were counterstained with Mayer's hematoxylin (Sigma-Aldrich) for 1 min 30 s; and, after final dehydration, slides were cleared in Histo-Clear and mounted with VectaMount (Vector laboratories). Stained slides were analyzed with a Zeiss Axioplan imaging system. Minor adjustments of contrast and brightness were made to the entire image; no further alterations were introduced.

Human and murine tissue scRNASeq data

Expression analysis of both human *DDR2* and murine *Ddr2* in scRNASeq data was performed as previously described [31, 32].

Bulk RNASeq of ddr2a F0-KO zfl heads

Sample preparation and sequencing

Four dpf zfl were anesthetized using MS-222 (tricaine methanesulfonate) and subsequently decapitated. Dissected heads were washed with Danieau's solution and transferred to sterile microcentrifuge tubes on ice. For each group (*ddr2a* CRISPR F0-KO, scr-ctrl, and uninjected zfl), three pools of 20 zfl heads were suspended in 500 μl TRIzol reagent (Zymo Research, Irvine, CA, USA) (n = 3 pools per experimental group). For homogenization, samples were transferred into homogenizer-compatible tubes containing 2.0 mm lysis beads and processed using a high-speed homogenizer at 6500 rpm for 30 seconds, followed by a 20-second pause.

RNA isolation and purification were performed according to the manufacturer's protocol using the Direct-zol™ RNA Miniprep Plus Kit (Zymo Research, Irvine, CA, USA). RNA was eluted in 50 μl of nuclease-free water and stored at −80°C. RNA concentration was determined using the Qubit 4 Fluorometer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA). Sequencing was performed on NovaSeq X Plus with 2×150 bp read length. RNA enrichment was performed using Watchmaker mRNA capture kit (Watchmaker Genomics, Boulder, CO, USA) with 200 ng RNA input.

Bulk RNASeq data processing and differential expression analysis

RNASeq reads were processed using the nf-core/rnaseq pipeline (v3.x) with STAR for genome alignment and Salmon for transcript quantification against the zebrafish (*D. rerio*) reference genome (GRCz11). Gene-level count matrices were imported into R (v4.x). Low-expression genes were filtered by retaining only genes with $\geq$10 counts in at least 2 samples across all groups (18 665 genes retained). PCA on variance-stabilizing transformed (VST) counts was used to assess sample clustering; three outlier samples were identified and excluded. Differential expression analysis was performed using DESeq2 (v1.38+) with pairwise contrasts between experimental groups. Genes with Benjamini–Hochberg adjusted p-value < 0.05 were considered significantly differentially expressed. Gene annotations were obtained from Ensembl (BioMart).

Gene-set enrichment analysis (GSEA)

To test whether ECM genes are disproportionately affected by *DDR2* loss-of-function, enrichment analysis was performed using the zebrafish Matrisome gene list (1015 genes; [33]), of which 693 were detected in our dataset. Absolute GSEA, a modified GSEA to detect bidirectional dysregulation, where genes were ranked by abs(log2FoldChange) $\times$ −log10(p-value) and a weighted running enrichment score was computed, with significance assessed against 10 000 gene-label permutations. Leading edge genes—the Matrisome subset driving the enrichment signal—were extracted for interpretation. As negative controls, the same analyses were applied to injection-only comparisons (scr_control versus UI) to confirm specificity to DDR2 knockout. All analyses and visualizations were performed in R using DESeq2, ggplot2, fgsea, and dplyr.

Statistical analysis and data visualization

Survival analysis of zfl was performed using Kaplan–Meier curves. Statistical analysis of phenotypic data and morphometric measurements was conducted to compare the effects of Ddr2a MO-KD and ctrl-MO groups, as well as *ddr2a* CRISPR F0-KO and scr-ctrl groups. The significance of differences between groups was evaluated using Welch's *t*-test. All statistical analyses and plots were generated using Python (v3.11.4) with the libraries: Matplotlib, Pandas, SciPy, and NumPy.

Graphic design

Graphics for Fig. 1A–F were designed using Draw.io (https://app.diagrams.net/, free online version).

Acknowledgements

We thank all patients and relatives for their participation. We thank the German support group for individuals with cleft lip and/or palate (Selbsthilfevereinigung für Lippen-Gaumen-Fehlbildungen e.V., Wolfgang Rosenthal Gesellschaft) for assistance with recruitment. We

acknowledge the invaluable assistance of all clinical, laboratory and bioinformatics personnel. Technical support was obtained from the Cologne Center for Genomics, which provided the Varbank software. We are thankful for the support of the Zebrafish Core Facility (Medical Faculty, University of Bonn, Germany) for assistance with zebrafish work. This study was supported by a BONFOR grant of the Medical Faculty, University of Bonn (O-249.0244).

This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats/ and via email from contact@deciphergenomics.org. DECIPHER is hosted by EMBL-EBI and funding for the DECIPHER project was provided by the Wellcome Trust [grant number WT223718/Z/21/Z]. The human embryonic and fetal material was provided by the Joint MRC/Wellcome Trust (grant# MR/×008304/1 and 226202/Z/22/Z) Human Developmental Biology Resource (https://hdbr.org/).

We also would like to thank Prof. Andreas Miething (Institute of Neuroanatomy, University of Bonn, Germany) for insightful and productive discussion.

J. A. Capecki was supported by a scholarship from the Studienstiftung des Deutschen Volkes (German Academic Scholarship Foundation) during research semesters undertaken as part of her medical studies. K. U. Ludwig and M. Geyer are members of the ImmunoSensation^2 Cluster of Excellence, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2151—390873048). N. Ishorst was supported by a DFG-funded position (IS 427/2-1). We would like to thank the NGS Core Facility of the Medical Faculty at the University of Bonn/ West German Genome Center (WGGC) for providing support and instrumentation for the bulk RNASeq. We would like to thank the Core Unit for Bioinformatics Data Analysis (CUBA) of University Hospital Bonn for providing computing resources for the analysis of the additional exome sequencing dataset. This publication was supported by the Open Access Publication Fund of the University of Bonn.

Author contributions

Julia Anna Capecki (Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing—original draft, Writing—review & editing), Helena Shkuro (Data curation, Formal analysis, Investigation), Öznur Yilmaz (Data curation, Methodology), Lisa Schmitt (Data curation, Formal analysis, Methodology, Writing—review & editing), Khadija Channab (Data curation, Methodology), Tobias T Lindenberg (Data curation, Methodology), Teresa Kruse (Data curation), Sarah Achterrath (Data curation), Berta Crespo (Data curation, Investigation), Anna Siewert (Data curation, Formal analysis, Investigation), Mostafa Bakhshi (Data curation, Formal analysis, Methodology, Writing—review & editing), Leandra Pantel (Data curation, Formal analysis, Writing—review & editing), Kerstin Ludwig (Project administration, Supervision, Writing—review & editing), Matthias Geyer (Formal analysis, Resources, Software), Elisabeth Mangold (Conceptualization, Funding acquisition, Project administration, Supervision, Writing—review & editing), Benjamin Odermatt (Conceptualization, Formal analysis, Funding acquisition, Project administration, Supervision, Writing—review & editing), and Nina Ishorst (Conceptualization, Formal analysis, Funding acquisition, Project administration, Supervision, Writing—review & editing)

Supplementary material

Supplementary material is available at *Human Molecular Genetics* online.

Conflicts of interest

None of the authors has any competing interests to declare.

References

1. Dixon MJ, Marazita ML, Beaty TH *et al.* Cleft lip and palate: understanding genetic and environmental influences. *Nat Rev Genet* 2011;**12**:167–178. https://doi.org/10.1038/nrg2933
2. Mossey PA, Modell B. Epidemiology of oral clefts 2012: an international perspective. *Front Oral Biol* 2012;**16**:1–18. https://doi.org/10.1159/000337464
3. Mangold E, Ludwig KU, Nöthen MM. Breakthroughs in the genetics of orofacial clefting. *Trends Mol Med* 2011;**17**:725–733. https://doi.org/10.1016/j.molmed.2011.07.007
4. Kosowski TR, Weathers WM, Wolfswinkel EM *et al.* Cleft palate. *Semin Plast Surg* 2012;**26**:164–169. https://doi.org/10.1055/s-0033-1333883
5. Miller CK. Feeding issues and interventions in infants and children with clefts and craniofacial syndromes. *Semin Speech Lang* 2011;**32**:115–126. https://doi.org/10.1055/s-0031-1277714
6. Lan Y, Xu J, Jiang R. Cellular and molecular mechanisms of Palatogenesis. *Curr Top Dev Biol* 2015;**115**:59–84. https://doi.org/10.1016/bs.ctdb.2015.07.002
7. Levi B, Brugman S, Wong VW *et al.* Palatogenesis: engineering, pathways and pathologies. *Organogenesis* 2011;**7**:242–254. https://doi.org/10.4161/org.7.4.17926
8. Bush JO, Jiang R. Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. *Development* 2012;**139**:231–243. https://doi.org/10.1242/dev.067082
9. Calzolari E, Bianchi F, Rubini M *et al.* Epidemiology of cleft palate in Europe: implications for genetic research. *Cleft Palate Craniofac J* 2004;**41**:244–249. https://doi.org/10.1597/02-074.1
10. Grosen D, Bille C, Petersen I *et al.* Risk of oral clefts in twins. *Epidemiology* 2011;**22**:313–319. https://doi.org/10.1097/EDE.0b013e3182125f9c
11. Mangold E, Böhmer AC, Ishorst N *et al.* Sequencing the GRHL3 coding region reveals rare truncating mutations and a common susceptibility variant for nonsyndromic cleft palate. *Am J Hum Genet* 2016;**98**:755–762. https://doi.org/10.1016/j.ajhg.2016.02.013
12. Marazita ML. The evolution of human genetic studies of cleft lip and cleft palate. *Annu Rev Genomics Hum Genet* 2012;**13**:263–283. https://doi.org/10.1146/annurev-genom-090711-163729
13. Ranji P, Pairet E, Helaers R *et al.* Four putative pathogenic ARHGAP29 variants in patients with non-syndromic orofacial clefts (NsOFC). *Eur J Hum Genet* 2025;**33**:38–43. https://doi.org/10.1038/s41431-024-01727-3
14. Basha M, Demeer B, Revencu N *et al.* Whole exome sequencing identifies mutations in 10% of patients with familial non-syndromic cleft lip and/or palate in genes mutated in well-known syndromes. *J Med Genet* 2018;**55**:449–458. https://doi.org/10.1136/jmedgenet-2017-105110
15. Hoebel AK, Drichel D, van de Vorst M *et al.* Candidate genes for nonsyndromic cleft palate detected by exome sequencing. *J Dent Res* 2017;**96**:1314–1321. https://doi.org/10.1177/0022034517722761

16. Visel A, Attanasio C. *RNA-Seq analysis in E13.5 mouse palates. Face Base Consortium*, 2015. https://doi.org/10.25550/TN2
17. Samuels BD, Aho R, Brinkley JF *et al.* Face Base 3: analytical tools and FAIR resources for craniofacial and dental research. *Development* 2020;**147**:dev191213. https://doi.org/10.1242/dev.191213
18. Vogel WF, Abdulhussein R, Ford CE. Sensing extracellular matrix: an update on discoidin domain receptor function. *Cell Signal* 2006;**18**:1108–1116. https://doi.org/10.1016/j.cellsig.2006.02.012
19. Mohamed FF, Ge C, Hallett SA *et al.* Control of craniofacial development by the collagen receptor, discoidin domain receptor 2. *eLife* 2023;**12**:e77257. https://doi.org/10.7554/eLife.77257
20. Leitinger B, Kwan APL. The discoidin domain receptor DDR2 is a receptor for type X collagen. *Matrix Biol* 2006;**25**:355–364. https://doi.org/10.1016/j.matbio.2006.05.006
21. Bargal R, Cormier-Daire V, Ben-Neriah Z *et al.* Mutations in DDR2 gene cause SMED with short limbs and abnormal calcifications. *Am J Hum Genet* 2009;**84**:80–84. https://doi.org/10.1016/j.ajhg.2008.12.004
22. Cormier-Daire V. Spondylo-epi-metaphyseal dysplasia. *Best Pract Res Clin Rheumatol* 2008;**22**:33–44. https://doi.org/10.1016/j.berh.2007.12.009
23. Xu L, Jensen H, Johnston JJ *et al.* Recurrent, activating variants in the receptor tyrosine kinase DDR2 cause Warburg-Cinotti syndrome. *Am J Hum Genet* 2018;**103**:976–983. https://doi.org/10.1016/j.ajhg.2018.10.013
24. Kumari P, Singh SK, Raman R. A novel non-coding RNA within an intron of CDH2 and association of its SNP with non-syndromic cleft lip and palate. *Gene* 2018;**658**:123–128. https://doi.org/10.1016/j.gene.2018.03.017
25. Nikopensius T, Jagomägi T, Krjutskov K *et al.* Genetic variants in COL2A1, COL11A2, and IRF6 contribute risk to nonsyndromic cleft palate. *Birth Defects Res A Clin Mol Teratol* 2010;**88**:748–756. https://doi.org/10.1002/bdra.20700
26. Green BL, Fasaye G-A, Samaranayake SG *et al.* Frequent cleft lip and palate in families with pathogenic germline CDH1 variants. *Front Genet* 2022;**13**:1012025. https://doi.org/10.3389/fgene.2022.1012025
27. Gupta N, Correa ARE, Jana M *et al.* Report of a novel homozygous nonsense DDR2 mutation in an Indian adult male with Spondylo-meta-epiphyseal dysplasia, short limb-abnormal calcification type. *J Pediatr Genet* 2019;**8**:153–156. https://doi.org/10.1055/s-0039-1683382
28. Ishorst N, Hölzel S, Greve C *et al.* Role of ZFHX4 in orofacial clefting based on human genetic data and zebrafish models. *Eur J Hum Genet* 2025;**33**:595–606. https://doi.org/10.1038/s41431-024-01775-9
29. Mork L, Crump G. Zebrafish craniofacial development: a window into early patterning. *Curr Top Dev Biol* 2015;**115**:235–269. https://doi.org/10.1016/bs.ctdb.2015.07.001
30. Brunet FG, Volff J-N, Schartl M. Whole genome duplications shaped the receptor tyrosine kinase repertoire of jawed vertebrates. *Genome Biol Evol* 2016;**8**:1600–1613. https://doi.org/10.1093/gbe/evw103
31. Siewert A, Reiz B, Krug C *et al.* Analysis of candidate genes for cleft lip ± cleft palate using murine single-cell expression data. *Front Cell Dev Biol* 2023;**11**:1091666. https://doi.org/10.3389/fcell.2023.1091666
32. Siewert A, Hoeland S, Mangold E *et al.* Combining genetic and single-cell expression data reveals cell types and novel candidate genes for orofacial clefting. *Sci Rep* 2024;**14**:26492. https://doi.org/10.1038/s41598-024-77724-9
33. Nauroy P, Hughes S, Naba A *et al.* The in-silico zebrafish matrisome: A new tool to study extracellular matrix gene and protein functions. *Matrix Biology* 2018;**65**:5–13. https://doi.org/10.1016/j.matbio.2017.07.001
34. Leslie EJ, Liu H, Carlson JC *et al.* A genome-wide association study of nonsyndromic cleft palate identifies an etiologic missense variant in GRHL3. *Am J Hum Genet* 2016;**98**:744–754. https://doi.org/10.1016/j.ajhg.2016.02.014
35. Mariadoss AVA, Wang C-Z. Exploring the cellular and molecular mechanism of Discoidin domain receptors (DDR1 and DDR2) in bone formation, regeneration, and its associated disease conditions. *Int J Mol Sci* 2023;**24**:14895. https://doi.org/10.3390/ijms241914895
36. Erlebacher A, Filvaroff EH, Gitelman SE *et al.* Toward a molecular understanding of skeletal development. *Cell* 1995;**80**:371–378. https://doi.org/10.1016/0092-8674(95)90487-5
37. Carafoli F, Hohenester E. Collagen recognition and transmembrane signalling by discoidin domain receptors. *Biochim Biophys Acta* 2013;**1834**:2187–2194. https://doi.org/10.1016/j.bbapap.2012.10.014
38. Sunadome K, Erickson AG, Kah D *et al.* Directionality of developing skeletal muscles is set by mechanical forces. *Nat Commun* 2023;**14**:3060. https://doi.org/10.1038/s41467-023-38647-7
39. Hall JG. Importance of muscle movement for normal craniofacial development. *Journal of Craniofacial Surgery* 2010;**21**:1336–1338. https://doi.org/10.1097/SCS.0b013e3181ebcd4f
40. Seegmiller RE, Foster C, Burnham JL. Understanding chondrodysplasia (cho): a comprehensive review of cho as an animal model of birth defects, disorders, and molecular mechanisms. *Birth Defects Res* 2019;**111**:237–247. https://doi.org/10.1002/bdr2.1473
41. Diogo R, Hinits Y, Hughes SM. Development of mandibular, hyoid and hypobranchial muscles in the zebrafish: homologies and evolution of these muscles within bony fishes and tetrapods. *BMC Dev Biol* 2008;**8**:24. https://doi.org/10.1186/1471-213X-8-24
42. Schilling TF, Kimmel CB. Musculoskeletal patterning in the pharyngeal segments of the zebrafish embryo. *Development* 1997;**124**:2945–2960. https://doi.org/10.1242/dev.124.15.2945
43. Cox LL, Cox TC, Moreno Uribe LM *et al.* Mutations in the epithelial cadherin-p120-catenin complex cause Mendelian non-syndromic cleft lip with or without cleft palate. *Am J Hum Genet* 2018;**102**:1143–1157. https://doi.org/10.1016/j.ajhg.2018.04.009
44. Richardson RJ, Dixon J, Jiang R *et al.* Integration of IRF6 and Jagged2 signalling is essential for controlling palatal adhesion and fusion competence. *Hum Mol Genet* 2009;**18**:2632–2642. https://doi.org/10.1093/hmg/ddp201
45. Logan SM, Ruest LB, Benson MD *et al.* Extracellular matrix in secondary palate development. *Anat Rec* 2020;**303**:1543–1556. https://doi.org/10.1002/ar.24263
46. McLaren W, Gil L, Hunt SE *et al.* The Ensembl variant effect predictor. *Genome Biol* 2016;**17**:122. https://doi.org/10.1186/s13059-016-0974-4
47. Carafoli F, Bihan D, Stathopoulos S *et al.* Crystallographic insight into collagen recognition by Discoidin domain receptor 2. *Structure* 2009;**17**:1573–1581. https://doi.org/10.1016/j.str.2009.10.012
48. Westerfield M. The zebrafish book. Westerfield M. (ed.), A guide for the laboratory use of zebrafish (Danio rerio). In: *The zebrafish book. A guide for the laboratory use of zebrafish (Danio rerio)*, 4th edn. Eugene: Univ. of Oregon Press, 2000.
49. Thisse C, Thisse B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat Protoc* 2008;**3**:59–69. https://doi.org/10.1038/nprot.2007.514

50. Lange M, Granados A, Vijay Kumar S *et al.* A multimodal zebrafish developmental atlas reveals the state-transition dynamics of late-vertebrate pluripotent axial progenitors. *Cell* 2024;**187**:6742–6759.e17. https://doi.org/10.1016/j.cell.2024.09.047
51. Moreno-Mateos MA, Vejnar CE, Beaudoin J-D *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* 2015;**12**:982–988. https://doi.org/10.1038/nmeth.3543
52. Kroll F, Powell GT, Ghosh M *et al.* A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *eLife* 2021;**10**:e59683. https://doi.org/10.7554/eLife.59683
53. Meeker ND, Hutchinson SA, Ho L *et al.* Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Bio Techniques* 2007;**43**:610–614. https://doi.org/10.2144/000112619
54. Walker M, Kimmel C. A two-color acid-free cartilage and bone stain for zebrafish larvae. *Biotech Histochem* 2007;**82**:23–28. https://doi.org/10.1080/10520290701333558

Human Molecular Genetics, 2026, Vol. 35, 12, 1–17
https://doi.org/10.1093/hmg/ddag052
Original Article