Aus dem

CharitéCentrum 17 für Frauen-, Kinder- und Jugendmedizin
mit Perinatalzentrum und Humangenetik

Institut für Medizinische Genetik und Humangenetik
Direktor: Prof. Dr. med. Stefan Mundlos

Habilitationsschrift

Untersuchung des diagnostischen Potenzials der Computer-gestützten fazialen Phänotypisierung und der Hochdurchsatzsequenzierung für die klinische Genetik

zur Erlangung der Lehrbefähigung für das Fach Humangenetik

vorgelegt dem Fakultätsrat der Medizinischen Fakultät
Charité-Universitätsmedizin Berlin

von

Dr. med. Martin Atta Mensah

Eingereicht: **November 2023**
Dekan: **Prof. Dr. med. Joachim Spranger**

Für Theodor, Helene, Jakob und Hannah

Inhaltsverzeichnis

Abkürzungen

aCGH	array comparative genomic hybridisation, Microarray-basierte vergleichende genomische Hybridisierung
ACMG	American College of Medical Genetics and Genomics
ALDH1A2	*Aldehyde Dehydrogenase 1 Family, Member A2*
AMP	Association for Molecular Pathology
AUROC	area under the receiver operating characteristic curve, Fläche unter der Isosensitivitätskurve
BHLHA9	*Basic Helix-Loop-Helix Family, Member A9*
BOQA	Bayesian Ontology Querying for Accurate Comparison
BPTAS	Brachyphalangie-Polydaktylie-Tibia-Hypo/Aplasie-Syndrom
BTRC	*Beta-Transducin Repeat-Containing Protein*
CADD	combined annotation dependent depletion
CFTR	*Cystic Fibrosis Transmembrane Conductance Regulator*
DNS	Desoxyribonukleinsäure
DONSON	*Downstream Neighbour of SON*
DPCD	*Homolog of Mouse Deleted in Primary Ciliary Dyskenesia*
EJP-RD	European Joint Programme on Rare Diseases
EU	Europäische Union
FA	Fanconi-Anämie
FBXW4	*F-Box and WD40 Domain Protein 4*
FFS	Femoro-Faziales Syndrom
FGFR1	*Fibroblast Growth Factor Receptor 1*
FGFR2	*Fibroblast Growth Factor Receptor 2*
GLI3	*Glioma-associated-Kruppel Family Member 3*
HMGB1	*High Mobility Group Box 1*
HOXD13	*Homeobox D13*
HPO	Human Phenotype Ontology
IDR	intrinsically disordered region, intrinsisch ungeordnete Region
INSERM	Institut national de la santé et de la recherche médicale
MGS	Meier-Gorlin-Syndrom
MISSLA	Microcephaly, Short Stature and Limb Abnormality disorder
MMS	Mikrozephalie-Mikromelie-Syndrom
MPD	microcephalic primordial dwarfism,

	mikrozephaler primordialer Kleinwuchs
NGP	Next Generation Phenotyping, Computer-gestützte Phänotypisierung
NGS	Next Generation Sequencing, Hochdurchsatzsequenzierung
NHS	National Health Service, nationaler Gesundheitsdienst des Vereinigten Königreichs
PEDIA	Prioritization of exome data by image analysis
POLL	*DNA-Polymerase lambda*
PP4	unterstützendes (“supporting”) Kriterium 4 der ACMG/AMP Richtlinien zur Variantenklassifizierung
SCID	severe combined immunodeficiency, schwerer kombinierter Immundefekt
SEMA3D	*Semaphorin 3D*
SHH	*Sonic Hedgehog Signaling Molecule*
SLOS	Smith-Lemli-Opitz-Syndrom
SLS	Seckle-like syndrome, Seckel-ähnliches Syndrom
STRING	search tool for recurring instances of neighbouring genes
SVM	Support-Vector-Maschine
Translate-NAMSE	Translate-Projekt des nationalen Aktionsbündnisses für Menschen mit seltenen Erkrankungen
UBA2	*Ubiquitin-like Modifier-Activating Enzyme 2*
WGS	Whole Genome Sequencing, Ganzgenomsequenzierung

1. Einleitung

1.1 Seltene Erkrankungen als Herausforderung für Gesundheitssysteme

Seltene Erkrankungen stellen eine heterogene Gruppe von Störungen der Gesundheit dar. Sie umfassen sowohl angeborene als auch erworbene Krankheiten und können jedes Organsystem betreffen. Es gibt mehrere tausend seltene Erkrankungen, die Datenbank *Orphadata* des französischen *Institut national de la santé et de la recherche médicale (INSERM)* listet 9382 klinische Entitäten auf (INSERM and Orphanet consortium 1999). Schätzungen gehen davon aus, dass 39% bis 80% davon eine genetische Ursache haben (Jackson et al. 2018; Ferreira 2019).

Genaue Zahlen sind nur schwer zu bestimmen, da keine weltweit einheitliche Definition für seltene Erkrankungen existiert (Richter et al. 2015). Eine solche ist auch schwer zu fassen, da dieselbe Erkrankung in unterschiedlichen Epochen bzw. Erdteilen selten oder häufig sein kann.

So ist die autosomal-rezessive Sichelzellanämie in Subsahara-Afrika mit einer Inzidenz von bis zu 2,6% häufig (GBD 2021 Sickle Cell Disease Collaborators 2023). Dies entspricht einer Heterozygotenfrequenz (Anteil der Anlageträger an der Gesamtbevölkerung) von bis zu 27%. In der mitteleuropäischen Bevölkerung existiert die Krankheit hingegen fast nicht (Mañú Pereira et al. 2023). Ob es sich um eine seltene Erkrankung handelt, hängt also von der betrachteten Region ab.

Die Zahl der jährlichen Masernfälle in England und Wales schwankte zwischen 1940 und 1966 zwischen 200,000 und 800,000, um in den Jahren danach aufgrund der Einführung und Verfügbarmachung von gegen das auslösende RNA-Virus geeigneten Impfstoffen bis zur Jahrtausendwende auf nahezu 0 zu sinken (Berche 2022). Die Masern wurden folglich von einer häufigen zu einer sehr seltenen Erkrankung.

Einen umgekehrten Weg könnte die autosomal-rezessive Mukoviszidose, die auf biallelische Mutationen im *CFTR*-Gen zurückzuführen ist, nehmen. Die Prävalenz der Erkrankung hat aufgrund eines durch die Entwicklung neuer Therapien deutlich verlängerten Überlebens seit den 1950er Jahren kontinuierlich zugenommen und wird dies voraussichtlich auch weiter tun (Elborn 2016). Eine seltene Erkrankung, die ehemals nur bei Kindern gesehen wurde (die dann daran verstarben), entwickelt sich also aufgrund des medizinischen Fortschritts zu einer relativ häufigen Erkrankung, die auch in der Erwachsenenbevölkerung vorkommt.

In der Europäischen Union (EU) gilt eine Erkrankung im rechtlichen Sinne als selten, wenn ihre Prävalenz bei weniger als 1:2.000 liegt (Moliner and Waligora 2017). Aufgrund der erwähnten Vielzahl von seltenen Erkrankungen ist ihre minimale kumulative Gesamtprävalenz allerdings erheblich höher und wird auf 1,5% bis 6,2% geschätzt (Walker et al. 2017; Chiu et al. 2018; Ferreira 2019) (***Abb. 1***). Umgerechnet auf die 450 Millionen Einwohner der EU bedeutet dies, dass hier mindestens 6,7 bis 27,7 Millionen Menschen von einer seltenen Erkrankung betroffen sind. Ein nur scheinbarer Widerspruch, der auch als Paradox der seltenen Erkrankungen bezeichnet wird: Seltene Erkrankungen sind häufig. Folglich stellt die Gruppe der seltenen Erkrankungen Gesundheitssysteme vor erhebliche Herausforderungen.

Mehrere politische Initiativen wurden gestartet, um Diagnostik und Therapie seltener Erkrankungen zu verbessern. Zu nennen sind beispielsweise Programme wie das *European Joint Programme on Rare Diseases* (EJP-RD) und das *Translate-Projekt des Nationalen Aktionsbündnisses für Menschen mit seltenen Erkrankungen* (Translate-NAMSE) auf Unions- bzw. Bundesebene (Choukair et al. 2021; Druschke et al. 2021; Rillig et al. 2022).

Das EJP-RD hat sich z.B. die

1. “Verbesserung der Integration, der Wirksamkeit, der Produktion und der sozialen Auswirkungen der Forschung über seltene Erkrankungen durch die Entwicklung, Demonstration und Förderung des europaweiten/weltweiten Austauschs von Forschungs- und klinischen Daten, Materialien, Verfahren, Wissen und Know-how”
2. “Umsetzung und Weiterentwicklung eines effizienten Modells der finanziellen Unterstützung für alle Arten der Forschung über seltene Erkrankungen (Grundlagenforschung, klinische Forschung, epidemiologische Forschung, Sozialforschung, Wirtschaftsforschung, Gesundheitsfürsorge) in Verbindung mit einer beschleunigten Nutzung der Forschungsergebnisse zum Nutzen der Patienten.”

als Hauptziele gesetzt (Julkowska and EJP RD Initiative 2019).

Das Erreichen dieser Ziele benötigt und umfasst die Entwicklung und klinische Validierung neuer diagnostischer Technologien.

1.2 Klassische und reverse Phänotypisierung als Schlüssel zur Diagnosestellung

Die Diagnostik erblicher seltener Erkrankungen und damit des Großteils aller seltenen Erkrankungen erfordert eine detaillierte Kenntnis genetischer Syndrome. Nur wenn die Zuordnung eines klinischen Phänotyps (der Gesamtheit aller körperlichen, geistigen und laborchemischen Besonderheiten eines Patienten) zu einem bestimmten Syndrom gelingt, kann allerdings eine klinische Diagnose gestellt werden (Schulze and McMahon 2004; Ferreira et al. 2018). D.h., für die Sicherung der Diagnose eines genetischen Syndroms braucht es stets beides: Ein passendes Muster an Symptomen und den Nachweis einer Mutation in einem passenden Gen (bzw. in einer genomischen Region).

Die Testung der genetischen Information kann zu diesem Zweck z.B. durch eine Sequenzierung erfolgen. Die klassische Methode zur Sequenzierung von Kernsäuren nach Sanger verwendet zusätzlich zu den einfachen Desoxyribonukuleotiden einen geringen Anteil von Didesoxyribonukleotiden bei der Amplifikation der DNS, welcher dann entsprechend der anhängenden Base zu einem sequenzspezifischen Kettenabbruch führt (Kettenabbruch Methode)(Sanger, Nicklen, and Coulson 1977). Die dadurch entstehenden Amplifikate lassen sich in einem Gel auftrennen und aus ihrer Länge die Basenabfolge ablesen. Das Verfahren war zwar bei seiner Einführung revolutionär, aber eine rasche, gleichzeitige Testung mehrerer Gene ist mittels der klassischen Methode nach Sanger nicht möglich bzw. zu teuer.

In der Vergangenheit ging daher der Prozess der klinischen Phänotypisierung einer genetischen Testung voraus. Ziel war es, die Zahl der zu testenden Gene möglichst klein zu halten. Allerdings barg diese Vorgehensweise aus mehrerlei Gründen das Risiko, eine tatsächlich vorliegende seltene genetische Erkrankung nicht zu diagnostizieren:

- der Patient zeigt nicht alle oder nicht die typischen Zeichen der Erkrankung
- die gesuchte Erkrankung ist dem Untersucher unbekannt
- die gesuchte Erkrankung ist in der Fachliteratur nicht definiert
- die Erkrankung ist beschrieben, aber die genetische Grundlage ist unbekannt
- das gewählte Testverfahren kann die Art der ursächlichen Mutation nicht erfassen

Mit der Einführung der Hochdurchsatzsequenzierung (*next generation sequencing*, NGS) hat sich dies geändert (Shendure 2011; Bourchany et al. 2017).

Verschiedene Verfahren existieren. Bei einer etablierten Methode werden z.B. DNS-Fragmente auf einer festen Oberfläche verankert und dann mithilfe einer Polymerase schrittweise komplementäre Elemente synthetisiert. Dabei werden je nach anhängender Base mit einem spezifischen fluoreszierenden Marker gekennzeichnete Nukleotide eingesetzt, die sich reversibel an die naszierende DNS binden. Die Wellenlänge des emittierten Lichts kann von einer Kamera erfasst werden. Durch das Aufbringen mehrerer Cluster von DNS Fragmenten auf eine feste Oberfläche ist die parallele fotografische Erfassung der von den einzelnen Clustern emittierten Wellenlängen nach dem Einbau jeweils einer weiteren Base möglich. Mithilfe von Computern kann automatisiert aus der entstehenden Bildfolge die Sequenz der untersuchten DNS-Fragmente ermittelt werden. Auf diese Weise können innerhalb kurzer Zeit riesige Mengen von DNS sequenziert werden (Canard and Sarfati 1994). Jüngste Methoden erlauben sogar die noch schnellere Sequenzierung noch größerer Fragmente, indem diese als einzelne DNS-Moleküle mithilfe eines elektrischen Feldes durch kleinste Öffnungen in einer Membran (i.d.R. geeignete Proteine in einer Lipidmembran, sog. Nanoporen) getrieben werden. Dabei wird der elektrische Widerstand über der Nanopore gemessen, der für jede der vier Basen einen spezifischen Wert hat und daher das Ablesen der Basensequenz erlaubt (Kasianowicz et al. 1996; Stoddart et al. 2009).

NGS ermöglicht als Exomsequenzierung die parallele, rasche und kostengünstige Testung sämtlicher proteinkodierender Abschnitte aller ca. 25.000 menschlichen Gene und als Genomsequenzierung sogar die Testung nahezu des vollständigen ca. 2 x 3 Milliarden Basen umfassenden humanen Genoms (Shendure 2011; Wright et al. 2015). Dauerte die erste Sequenzierung des menschlichen Genoms in den 90er Jahren im Rahmen des *Human Genome Projects* noch mehr als ein Jahrzehnt (und auch nur deswegen nicht noch länger, weil neuere Methoden die klassischen Verfahren während der Durchführung des Projekts ablösten) (McPherson et al. 2001; Venter et al. 2001), ist man inzwischen in der Lage Datenbanken tausender menschlicher Genome zu generieren und kontinuierlich zu erweitern (1000 Genomes Project Consortium et al. 2015; Karczewski et al. 2020; Halldorsson et al. 2022).

Da es aufgrund seiner Leistungsfähigkeit im Gegensatz zur klassischen Sequenzierung nach Sanger die parallele Testung aller menschlichen Gene ermöglicht, kann mittels NGS durch die Identifikation ursächlicher, pathogener Mutationen in krankheitsassoziierten Genen theoretisch auch dann eine molekulare Diagnose gestellt werden, wenn klinisch kein eindeutiger Verdacht vorgelegen hat, sich untypische oder unspezifische Zeichen aber dem mutierten Gen zuordnen lassen. Weil es sich um eine Umkehrung des bisher etablierten Vorgehens handelt (bisher: erst phänotypischer Verdacht, dann molekulargenetische Bestätigung; nun: erst molekulargenetischer Verdacht, dann phänotypische Validierung) wird dieses Vorgehen auch als reverse Phänotypisierung bezeichnet. (Schulze and McMahon

2004; Uliana and Percesepe 2016; de Goede et al. 2016; Landini et al. 2020; Swietlik et al. 2020; Seltzsam et al. 2022; Musante et al. 2022; Best et al. 2022; Solomon et al. 2023)

Außerdem kann nach dem Aufbau von Patientenkohorten mit ähnlichen Phänotypen mithilfe des NGS nach noch unbekannten Gen-Krankheitsassoziationen gesucht werden (Gilissen et al. 2012; Koboldt et al. 2013; Huang et al. 2015; Todd et al. 2015; Deciphering Developmental Disorders Study 2015; Pillay et al. 2022; Ibañez et al. 2022).

1.3 Genomsequenzierung als allgemeiner Test für die klinische Humangenetik

NGS hat in Form der Exomsequenzierung inzwischen Eingang in die klinische Routinediagnostik gefunden (Iglesias et al. 2014; Wright et al. 2015; Petrovski et al. 2019; Arts et al. 2019; Wells et al. 2022; Wright et al. 2023). Die Genomsequenzierung befindet sich derzeit im Übergang von einem hauptsächlichen Forschungsinstrument zu einem Werkzeug der Routinediagnostik (Lappalainen et al. 2019; Stranneheim et al. 2021; Shickh et al. 2021; S. Marwaha, Knowles, and Ashley 2022; Bowling et al. 2022; Halldorsson et al. 2022).

Die Genauigkeit, mit der sich genetische Varianten mittels NGS-basierten Testverfahren identifizieren lassen, schwankt allerdings in Abhängigkeit von dem zu detektierenden Variantentyp und der verwendeten Technologie (Koboldt 2020; Weißbach et al. 2021). Wenig fraglich erscheint dabei die Möglichkeit, kleine kodierende Varianten (single nucleotide variants, SNVs) mittels eines Exoms zu identifizieren. Aber schon die Zuverlässigkeit der Detektion von Insertionen bzw. Deletionen von Oligonukleotiden (sog. Indels) gelingt nur unsicher. (Jiang, Turinsky, and Brudno 2015)

Kopienzahlvarianten, die mehrere Kilobasen groß sind, und für deren Detektion bisher eine (mitunter hochauflösende) Mikroarray-basierte vergleichende genomische Hybridisierung (*array comparative genomic hybridisation*, aCGH) durchgeführt werden musste, können mit einem Exom nur mit eingeschränkter Präzision detektiert werden (wie hoch diese genau ist, ist unklar, da bisher kein Verfahren existiert, dass alle CNVs eines Genoms detektieren könnte). Dafür müssen die Varianten idealerweise mehrere Exons umfassen und deren Bruchpunkte sich in Exons befinden, was, da das Exom weniger als 2% des menschlichen Erbguts ausmacht, nur ausnahmsweise gegeben ist. (Manheimer et al. 2018; Gordeeva et al. 2021; Coutelier et al. 2022)

Die Bestimmung der Länge von repetitiven Elementen ist mittels short-read NGS ebenfalls nur eingeschränkt möglich (Tang et al. 2017; Weißbach et al. 2021).

Dies muss - insbesondere für den klinischen Einsatz der Technologie - bei allen Vorteilen, die ein NGS bietet, beachtet werden.

1.4 Strategien zur Interpretation von NGS-Daten

Beim NGS fallen sehr große Mengen an Daten an. Bildlich gesprochen werden, wo früher nur ein Gentest erfolgte, nun über 20.000 Tests gleichzeitig durchgeführt. Eine händische Durchsicht der Testergebnisse ist daher unmöglich. Spezielle Software zur Aufbereitung, Filterung und Priorisierung der identifizierten Varianten ist folglich notwendig (DePristo et al. 2011; Robinson et al. 2014; Wang et al. 2015; Shamseldin et al. 2017; Kernohan et al. 2018; Cipriani et al. 2020; Coutelier et al. 2022).

Für die Filterung der detektierten Varianten bietet sich u.a. ein Abgleich mit den parentalen Genotypen mittels Testung der Eltern (entweder als unmittelbarer Trio-NGS-Ansatz oder als Sanger-basierte Segregationsanalyse) an. Wird eine seltene Variante bei einem Patienten, nicht aber bei dessen Eltern gefunden (*de novo* Mutation) und ist der Patient das einzige und erste von einem Syndrom betroffene Familienmitglied, gibt es guten Grund zu prüfen, ob die Variante ursächlich ist. Wird eine Variante hingegen bei einem Indexpatienten und einem nicht-betroffenen Elternteil gefunden und das Gen, welches die Variante trägt, ist mit einem autosomal-dominanten Erbgang assoziiert, ist es wenig wahrscheinlich, dass die Variante ursächlich ist.

Außerdem können die Allelfrequenzen der gefundenen Varianten unter der Annahme, dass in der gesunden Allgemeinbevölkerung häufige Varianten nicht mit seltenen erblichen Erkrankungen assoziiert sind, zur Filterung und Priorisierung genutzt werden (Lek et al. 2016; Gudmundsson et al. 2022).

Auch die Art der Mutation kann zur Filterung und Priorisierung von Varianten herangezogen werden. Da synonyme Varianten nicht zu einem Austausch einer Aminosäure im kodierten Protein führen, ist es wenig wahrscheinlich (aber nicht unmöglich, weil sich hinter einer scheinbar synonymen Variante z.B. eine tatsächliche Spleiß-Variante verbergen kann), dass diese zu einer veränderten Funktion und damit Schädigung des Genprodukts führen. Missense-Varianten führen hingegen zu einem Austausch einer Aminosäure, ob dieser die Funktion des kodierten Proteins beeinflusst, hängt allerdings von vielen Faktoren ab (z.B. Position der Aminosäure im Protein, chemische Eigenschaften der Referenzaminosäure im Gegensatz zur alternativen Aminosäure). Trunkierende Varianten - also solche, die ein Protein verkürzen (z.B. Nonsense-Mutationen, die zu einem vorzeitigen Stopp-Codon führen, oder Rasterschubmutationen, die zu einer fehlerhaften Aminosäuresequenz und einem veränderten Stopp-Codon führen) haben allerdings sehr wahrscheinlich einen relevanten Einfluss auf die Funktion des kodierten Proteins. Daher erhalten z.B. synonyme Varianten weniger Gewicht als Missense-Varianten und diese wiederum weniger als trunkierende Varianten.

Weiterhin kann der Grad der evolutionären Konservierung eines Locus wichtige Hinweise auf dessen Bedeutung für die Gesundheit geben. Hochkonservierte Loci sind dabei eher verdächtig, funktionell relevant zu sein (Pollard et al. 2010). Ist eine bestimmte Aminosäure in Orthologen eines bestimmten menschlichen Gens auch bei den anderen betrachteten Spezies identisch, so gibt dies einen Hinweis auf ihre funktionelle Relevanz. Wird an der entsprechenden Aminosäureposition je nach Spezies hingegen eine andere Aminosäure kodiert, ist es weniger wahrscheinlich, dass diese funktionell relevant ist (zumindest ist eine Variabilität der Aminosäurefolge an dieser Position des Proteins mit dem Leben vereinbar). So ist z.B. das Phenylalanin an Position 508 des *CFTR*-Gens, welches bei der häufigsten Mutation bei mitteleuropäisch-stämmigen Mukoviszidosepatienten deletiert ist (ΔF508 auch F508del), stark evolutionär konserviert (Rishishwar et al. 2012; Ong and Ramsey 2023). Es (und auch die umgebende Aminosäurefolge) findet sich sogar beim Zebrafisch.

Entscheidend für die sinnvolle Priorisierung und Bewertung der Varianten ist allerdings die oben erwähnte Phänotypisierung, also die Prüfung der Übereinstimmung des beim Patienten vorliegenden Phänotyps mit dem mit der Variante (bzw. dem betroffenen Gen) assoziierten Syndrom (Smedley and Robinson 2015; Jacobsen et al. 2022).

Das *American College of Medical Genetics and Genomics* (ACMG) und die *Association for Molecular Pathology* (AMP) haben entsprechende Kriterien für Filterung, Priorisierung und Bewertung von genetischen Varianten als konsentierte Leitlinien herausgegeben (Kazazian, Boehm, and Seltzer 2000; C. S. Richards et al. 2008; S. Richards et al. 2015).

Die ACMG/AMP-Richtlinien definieren acht verschiedene Informationsquellen für die Bewertung von Varianten: *1)* Populationsdaten, *2)* rechnergestützte und prädiktive Daten, *3)* funktionelle Daten, *4)* Segregationsdaten, *5)* De-novo-Daten, *6)* allelische Daten, *7)* andere Datenbanken und *8)* andere Quellen. Anhand dieser werden zwölf Kategorien definiert, die für die Bewertung einer Variante als benigne sprechen, welche drei Evidenzgraden (“stand-alone”, “strong” und “supporting”) zugeordnet werden können. Außerdem werden 16 Kategorien definiert, die für die Pathogenität einer detektierten Variante sprechen. Diese verteilen sich auf vier Evidenzgrade (“very strong”, “strong”, “moderate”, “supporting”). Entsprechend eines festgelegten Punktesystems können die Kriterien verwendet werden, um die klinische Relevanz einer Variante auf einer fünfstufigen Skala zu bestimmen: *1)* benigne, 2) wahrscheinlich benigne, 3) unklare Signifikanz, 4) wahrscheinlich pathogen, 5) pathogen). Eine seltene, wahrscheinliche Null-Variante, die bei bestätigter Vater- und Mutterschaft der Eltern de novo entstanden ist, welche ein Gen betrifft, bei dem ein Funktionsverlust ein bekannter Pathomechanismus ist, wird bspw. als pathogen klassifiziert (1 strong pathogenic + very strong pathogenic → pathogen). Eine Variante, die in einer der großen Exomdatenbanken mit einer Allelfrequenz in der gesunden Allgemeinbevölkerung von mehr als 5% gelistet ist, wird als benigne bewertet (stand alone benigne → benigne). (S. Richards et al. 2015)

Die ACMG/AMP-Richtlinien stellen inzwischen den internationalen Standard bei der NGS-Datenbefundung bei dem Verdacht auf pathogene Keimbahnmutationen dar und sollten entsprechend auch für die Ausgestaltung und Verwendung entsprechender Analysesoftware berücksichtigt werden (Cristofoli et al. 2021; Lesmann, Klinkhammer, and Krawitz 2023).

1.5 Schwierigkeiten der automatisierten reversen Phänotypisierung

Die ACMG Kriterien wurden jedoch dafür kritisiert, teils wenig spezifische Kriterien zu verwenden, die keine ausreichende Objektivität bei der Klassifikation von genetischen Varianten gewährleisten. So fand sich eine Konkordanz von nur 34% bei der Bewertung von genetischen Varianten mittels der ACMG/AMP-Kriterien durch neun Labore, die sich auch nach einer gemeinsamen Konsenssuche der Labore nur auf 74% steigern ließ.(Amendola et al. 2016)

Es wurde vermutet, dass das unterstützende (“*supporting*”) Kriterium 4 der ACMG/AMP Richtlinien zur Klassifizierung von Varianten als pathogen (PP4) eine der Hauptursachen für die subjektive Anwendung der ACMG/AMP Richtlinien ist (Johnson et al. 2022): “*Der Phänotyp des Patienten oder die Familienanamnese ist hochspezifisch für eine Krankheit mit einer einzigen genetischen Ätiologie*” (S. Richards et al. 2015).

Eine informatische Auswertung der Patientendaten im Hinblick auf das Kriterium PP4 könnte die Konkordanz der ACMG-Kriterien auch über verschiedene Labore hinweg steigern. Bei der computergestützten Automatisierung der Befundung von NGS-Daten stellt dies jedoch eine erhebliche Hürde dar. PP4 verlangt im Grunde die reverse Phänotypisierung des Patienten. Für eine Automatisierung müssen folglich die verschiedenen Dimensionen eines klinischen Phänotyps, auch wenn sie nicht unmittelbar numerisch sind (insbesondere Besonderheiten der Anatomie und der Psyche), computerlesbar erfasst werden.

Für eine solche computerlesbare Erfassung auch von nicht-numerischen Wissensdomänen - wie bio-medizinischen Fakten - bieten sich Ontologien an. Dabei bezieht sich der Begriff “Ontologie” in der Informatik auf die formale Darstellung über die Struktur und Beziehungen von Entitäten in einem bestimmten Fachgebiet. Praktisch umgesetzt und visualisiert können Ontologien durch Knowledge-Graphen werden. Ein Knowledge-Graph ist eine digitale Datenstruktur, die semantische Typen, Eigenschaften und Beziehungen von Entitäten einer Wissensdomäne beschreibt.(Harris 2008; Haendel, Chute, and Robinson 2018). Beispiele für derartige Ontologien bzw. Knowledge-Graphen sind die *search tool for recurring instances of neighbouring genes* (STRING) *database* zur Erfassung von Protein-Protein-Interaktionen (bzw. Gen-Gen-Interaktionen) und die *Gene Ontology (Snel et al. 2000; Szklarczyk et al. 2023; Ashburner et al. 2000; Gene Ontology Consortium et al. 2023)*

Speziell für die Erfassung menschlicher Phänotypen, die bei genetischen Syndromen auftreten können, wurde die *Human Phenotype Ontology* (HPO) entwickelt (Robinson et al. 2008; Köhler et al. 2021). Die phänotypischen Besonderheiten werden in der HPO durch Begriffe repräsentiert (HPO-Terms). Die Position der HPO-Terms innerhalb des semantischen Netzwerkes definiert deren Beziehungen zueinander. Außerdem bietet sie eine Datenbank von Syndromen, welche mit diesen Phänotypen assoziiert sind, so dass bioinformatische Analysen in Form einer semantischen Ähnlichkeitssuche von spezieller Software zu diagnostischen Zwecken durchgeführt werden können (Köhler et al. 2019). Die HPO, welche über 15.247 Phänotypen enthält, die mehr als 7801 Erkrankungen zugeordnet sind, hat sich zum internationalen Standard für die digitale Erfassung klinischer Phänotypen entwickelt (Köhler et al. 2021).
Mit ihr können allerdings nur solche Besonderheiten erfasst werden, die auch bereits als HPO-Terms definiert sind. Darüber hinaus ist eine Repräsentation unterschiedlicher Schweregrade nur diskret (also stufenweise durch explizite Definition entsprechender HPO-Terms), aber nicht stetig möglich.

Die für PP4 geforderte Spezifität eines Phänotyps kann häufig durch die Erfassung auffälliger Merkmale des Gesichts erreicht werden, denn viele genetische Syndrome haben ein für sie charakteristisches Gesicht (Ferry et al. 2014; Solomon et al. 2023). Tatsächlich sind z.B. die typischen Gesichtszüge, die bei einem Patienten mit einem Down Syndrom beobachtet werden können und welche häufig wegweisend beim Stellen der Diagnose sind, Ausgangspunkt für die erste wissenschaftliche Beschreibung des Syndroms gewesen (Down 1866; Antonarakis et al. 2020).

Für eine spezifische faziale Phänotypisierung ist die HPO aber aus oben genannten Gründen nur bedingt geeignet. Schließlich existieren zwar HPO-Terms wie *Doll-like facies* (HP:0000295, puppenartiges Gesicht) oder *Coarse facial features* (HP:000280, grobe Gesichtszüge), aber es fehlen Begriffe, die den charakteristischen Aspekt eines für ein Syndrom typischen Gesichts spezifisch erfassen, wie z.B. "Noonan-artiges Gesicht" oder "Kabuki-typische Fazies". Da diese explizit angegeben, dem Untersucher also vor Beginn einer automatisierten Analyse von NGS-Daten bewusst sein müssten, wäre eine Erweiterung der HPO um solche Begriffe zum Zwecke der reversen Phänotypisierung auch nicht weiterführend - bedeutete dies doch, dass der Nutzer den charakteristischen Aspekt des Gesichts auch ohne die Hilfe der HPO bereits erkannt und erfasst hat.

Es bleibt nur, die einzelnen partiellen Phänotypen (z.B. absteigende Lidachsen, prominentes Kinn, oder schmales Lippenrot) zu benennen. Allerdings ist das Ausmaß der Ähnlichkeit eines Patientengesichts zum typischen fazialen Bild eines Syndroms variabel (es hängt u.a. von Alter, Geschlecht, und ethnischen Hintergrund ab) und mitunter nur sehr mild, so dass die charakteristische Zeichen einem Untersucher entgehen können.(Lumaka et al. 2017)

1.6 Maschinelles Sehen zur automatisierten fazialen Phänotypisierung

Um automatisiert einen fazialen Phänotyp erkennen und dessen Ausmaß bestimmen zu können, bieten sich als Ergänzung bzw. Alternative zu Ontologie-basierten Verfahren die Methoden des maschinellen Sehens an. Maschinelles Sehen verbindet die digitale Bildverarbeitung mit den Konzepten des maschinellen Lernens, so dass entsprechende Algorithmen nach einem Training selbstständig Bildinformationen klassifizieren können. Das maschinelle Sehen findet bei der Verarbeitung von medizinischen Bilddaten inzwischen breite Anwendung (Esteva et al. 2019) und auch eine Reihe darauf basierender Systeme zur Computer-gestützten fazialen Phänotypisierung wurden entwickelt (Boehringer et al. 2006; Vollmar et al. 2008; Boehringer et al. 2011; Ferry et al. 2014; Cerrolaza et al. 2016; Tu et al. 2018; Gurovich et al. 2019; Dudding-Byth et al. 2017; Hallgrímsson et al. 2020; Porras et al. 2021).

Die meistverwendete Anwendung, das neuronale Netz DeepGestalt, analysiert gewöhnliche Portraitfotos von Menschen mit seltenen Syndromen, die mit fazialen Auffälligkeiten einhergehen (Gurovich et al. 2019). Das System wurde auf mehr als 17,000 entsprechenden Fotos trainiert und kann für die Verwendung als Entscheidungsunterstützungssystem einem Patientenfoto eine Liste mit Verdachtsdiagnosen zuordnen. Dafür ordnet DeepGestalt der Fazies des Probanden für jedes Syndrom, für das es trainiert wurde, einen numerischen Wert (DeepGestalt Score) zu. Ein DeepGestalt Score kann zwischen 0 und 1 betragen, wobei 0 keine Übereinstimmung mit dem für ein bestimmtes Syndrom typischen Gesicht und 1 die größtmögliche Übereinstimmung ausdrückt. Die Liste der vorgeschlagenen Verdachtsdiagnosen ist in absteigender Reihenfolge nach den DeepGestalt Scores geordnet. DeepGestalt kann auch nutzerspezifisch trainiert werden, um die faziale Unterscheidbarkeit verschiedener Patientenkohorten z. B. bei der Definition neuer Syndrome zu messen (Knaus et al. 2018; Liehr et al. 2018; Vorravanpreecha et al. 2018; Martinez-Monseny et al. 2019; Pascolini et al. 2019; Mishima et al. 2019; Kruszka et al. 2019; Carli et al. 2019; Weiss et al. 2020; Staufner et al. 2020; Myers et al. 2020; Tekendo-Ngongang and Kruszka 2020; Mak et al. 2021; Zhang et al. 2022). Patientenfotos werden dazu zu eigens definierten Kohorten zusammengestellt und das Netzwerk versucht zu erlernen, diese zu unterscheiden. Das System gibt zur Beurteilung der Unterscheidbarkeit eine Wahrheitsmatrix und Isosensitivitätskurven aus.

Da diese Techniken zur Computer-gestützten fazialen Phänotypisierung wie das NGS theoretisch einen raschen parallelen Abgleich mit tausenden seltenen Diagnosen ermöglichen, werden sie auch als faziales next generation phenotyping (NGP) bezeichnet (Liehr et al. 2018; van der Donk et al. 2019).

1.7 Fragestellung

Ziel dieser Arbeit ist es festzustellen, wie NGS und faziales NGP genutzt werden können, um die oben genannten Limitationen der klassischen Phänotypisierung zu überwinden und so die Diagnostik seltener genetischer Erkrankungen zu verbessern.

Dazu wird am Beispiel des inzwischen mit dem Gen *Downstream-Neighbour-of-SON* (*DONSON*) assoziierten Mikrozephalie-Kleinwuchs-und-Gliedmaßenanomalie-Syndroms (microcephaly short stature and limb abnormality disorder, MISSLA) gezeigt, wie eine Exomsequenzierung genutzt werden kann, um eine in der Fachliteratur zum Untersuchungszeitpunkt nicht definierte Erkrankung zu identifizieren.

Die Fähigkeit des fazialen NGP-Systems *DeepGestalt*, MISSLA von der phänotypisch überlappenden Fanconi-Anämie zu unterscheiden, wird gemessen.

Ein Klassifikator für die Priorisierung von Exomdaten unter Einbindung von DeepGestalt wird vorgestellt. Seine Effizienz bei der automatisierten Analyse von Exomdaten - und damit seine Fähigkeit, auch dem Untersucher unbekannte Diagnosen vorzuschlagen - wird getestet.

DeepGestalts diagnostische Genauigkeit und die Zahl der dem System bekannten Syndrome werden gemessen. Die Zahl der von DeepGestalt unterstützten Syndrome sowie dessen Sensitivität und Spezifität werden bestimmt. Darüber hinaus wird untersucht, wie diese zur Identifikation von Patienten, insbesondere mit milden oder untypischen fazialen Phänotypen, weiter gesteigert werden können.

Schließlich wird untersucht, welche diagnostischen Vorteile eine Ganzgenomsequenzierung (whole genome sequencing, WGS) gegenüber etablierten genetischen Testverfahren bietet. Insbesondere wird geprüft, ob diese sich als einheitliche Methode zur Erfassung verschiedenster Mutationsarten eignet. Am Beispiel des Brachyphalangie-Polydaktylie-Tibia-Hypo/Aplasie-Syndroms (BPTAS) wird gezeigt wie WGS nicht nur in der Literatur nicht beschriebene Gen-Syndrom-Assoziation, sondern auch spezifische Varianten-Syndrom-Assoziationen identifizieren kann.

2. Eigene Arbeiten

2.1 Mikrozephalie, Kleinwuchs und Gliedmaßenanomalien aufgrund neuer autosomaler biallelischer DONSON-Mutationen bei zwei deutschen Geschwistern

Schulz S*, **Mensah MA***, de Vries H, Fröber R, Romeike B, Schneider U, Borte S, Schindler D, Kentouche K.

Microcephaly, short stature, and limb abnormality disorder due to novel autosomal biallelic DONSON mutations in two German siblings.

Eur J Hum Genet. 2018 Sep;26(9):1282-1287. DOI: 10.1038/s41431-018-0128-0

*contributed equally

In dieser Studie haben wir zwei Geschwister (weiblicher Fet und ein 9 Monate altes Mädchen) untersucht, die beide bei ansonsten unauffälliger Familienanamnese von einem angeborenen syndromalen Kleinwuchs - wenn auch mit verschiedenem Schweregrad - betroffen waren. Aufgrund der Verwandtschaft und weil die Phänotypen stark überlappten und diese Überlappungen auch spezifische Zeichen umfassten (radiale Strahlendefekte, Pachygyrie, charakteristische Fazies), nahmen wir das Vorliegen einer seltenen autosomal-rezessiven genetischen Ursache an.

Als zum Untersuchungszeitpunkt definierte Verdachtsdiagnose kam nach klassischer Phänotypisierung insbesondere eine Erkrankung aus der Gruppe der Fanconi-Anämien (FA) in Frage, die neben hämatologischen Auffälligkeiten mit einem Kleinwuchs und radialen Strahlendefekten einhergeht. Auf diese Erkrankungsgruppe deutete auch ein positiver Mitomycin-C-Test hin, der Zellzyklusdefekte, die typischerweise mit einer FA einhergehen, anzeigt. Die gezielte Testung der mit FA assoziierten Gene erbrachte allerdings ein unauffälliges Ergebnis. Die Differentialdiagnosen Thrombozytopenie-Radiusaplasie-Syndrom und Roberts-Syndrom konnten durch gezielte Testung ebenfalls nicht bestätigt werden. Da bei einer Schwester ein schwerer kombinierter Immundefekt (severe combined immunodefcicency, SCID) vorlag, wurden auch SCID-assoziierte Gene und per Microarrayanalyse das Vorliegen einer Deletion 22q11 (DiGeorge-Syndrom) geprüft. Die Ergebnisse waren genau wie ein Karyogramm unauffällig.

Mehrere der typischen Hürden der klassischen Phänotypisierung kamen als Erklärung für das Nicht-Identifizieren der genetischen Ursache in der Familie in Frage. Da sich die Schweregrade der Erkrankung bei den Schwestern unterschieden, war unklar, welches der charakteristische Phänotyp war und es erschien darüber hinaus aufgrund der unauffälligen genetischen Testergebnisse wahrscheinlich, dass es sich um eine den Untersuchern unbekannte zum Untersuchungszeitpunkt möglicherweise noch nicht definierte Erkrankung handelte.

Es wurde daher eine Exomsequenzierung einer der betroffenen Schwestern, beider Eltern und eines nicht-betroffenen Bruders durchgeführt. Eine Filterung der Daten nach seltenen, potentiell funktionell relevanten biallelischen Varianten, die mit dem vorliegenden Phänotyp segregierten, identifizierte zwei Varianten in *DONSON*(NM_017613.3) als potentielle Ursache: c.1047-2A>G p.(?) und c.1433C>T p.(Pro478Leu). Die Testung der anderen betroffenen Schwester und eines weiteren nicht-betroffenen Bruders bestätigte das Segregieren der biallelischen Trägerschaft mit der Erkrankung. Eine Exomsequenzierung konnte also potentiell pathogene Varianten in einem zum Untersuchungszeitpunkt uncharakterisierten Gen identifizieren.

Nachdem Reynolds et al. und Evrony et. al. *DONSON*-Mutationen als Ursache des mikrozephalen primordialen Kleinwuchses (MPD, *microcephalic primordial dwarfism*) bzw. des perinatal letalen Mikrozephalie-Mikromelie-Syndroms (MMS) beschrieben, erfolgte eine reverse Phänotypisierung und die bei den untersuchten Geschwistern identifizierten Varianten konnten als pathogen klassifiziert werden.

Interessanterweise zeigte der Fet MMS und das Mädchen MPD, so dass wir schlussfolgerten, dass es sich dabei nicht um zwei distinkte klinische Entitäten, sondern um unterschiedliche Ausprägungen nur eines Syndroms - MISSLA - handelt.

https://doi.org/10.1038/s41431-018-0128-0

https://doi.org/10.1038/s41431-018-0128-0

https://doi.org/10.1038/s41431-018-0128-0

https://doi.org/10.1038/s41431-018-0128-0

https://doi.org/10.1038/s41431-018-0128-0

https://doi.org/10.1038/s41431-018-0128-0

2.2 Differenzierung von MISSLA und Fanconi-Anämie durch computergestützte Bildanalyse und Präsentation von zwei neuen MISSLA-Geschwistern

Danyel M, Cheng Z, Jung C, Boschann F, Pantel JT, Hajjir N, Flöttmann R, Schulz S, Demuth I, Sheridan E, Mundlos S, Horn D, **Mensah MA**.

Differentiation of MISSLA and Fanconi anaemia by computer-aided image analysis and presentation of two novel MISSLA siblings.

Eur J Hum Genet. 2019 Dec;27(12):1827-1835. DOI: 10.1038/s41431-019-0469-3

MISSLA und FA ähneln sich in Bezug auf die körperlichen Merkmale und in den Laborbefunden (Zellzyklusdefekte) stark. 21 FA-assoziierte Gene sind bekannt. Bei mehreren Patienten, die nach klassischer Phänotypisierung die klinische Diagnose FA erhielten, fand sich später eine *DONSON*-Mutation. Es stellte sich daher die Frage, ob MISSLA als eine Form der FA und *DONSON* lediglich als 22. FA-assoziiertes Gen betrachtet werden sollte.

In dieser Studie untersuchten wir, ob es faziale Unterschiede zwischen FA und MISSLA gibt und ob und wie diese mittels NGP gemessen werden können. Außerdem berichten wir in dieser Studie über zwei weitere Geschwister (Bruder und Schwester) mit schwerem MISSLA und der zuvor beschriebenen Mutationen *DONSON*(NM_017613.3):c.1433C>T p.(Pro478Leu) sowie der neuen wahrscheinlich pathogenen Variante *DONSON*(NM_017613.3):c.661T>C p.(Trp221Arg) in compound heterozygotem Zustand.

Zur Messung fazialer Unterschiede wurden basierend auf DeepGestalt nutzerspezifische Klassifikatoren trainiert. Fünf Klassen von Portraitfotos wurden für die Testung mit einem Mehrklassen-Klassifikator verwendet. Neben Bildern von MISSLA- und FA-Patienten, Aufnahmen von Patienten mit dysmorphen Syndromen, Portraits von dazu nach Alter, Geschlecht und ethnischem Hintergrund passenden unauffälligen Kontrollen und Fotos von Patienten mit Smith-Lemli-Opitz-Syndrom (SLOS). Für alle Klassen wurden Durchschnittsgesichter berechnet. Darüber hinaus wurde getestet, ob ein binärer Klassifikator MISSLA- von FA-Bildern unterscheiden kann.

Es zeigte sich, dass der Klassifikator die faziale Erscheinung von MISSLA-Patienten von Portraitaufnahmen von FA-Patienten unterscheiden kann. MISSLA ließ sich dabei durch den Mehrklassen-Klassifikator ähnlich gut erkennen wie SLOS, für das eine charakteristische Fazies bekannt ist. Die berechneten Durchschnittsgesichter zeigten die für MISSLA und SLOS typischen Gesichtszüge. Interessanterweise ließen sich - wenn auch schlechter - ebenfalls die FA-Bilder und in geringerem Maße die dysmorphen und unauffälligen Kontrollbilder überzufällig gut klassifizieren.

Wir schlussfolgerten, dass MISSLA als eigenständige von der FA unabhängige Diagnose betrachtet werden sollte. Außerdem zeigten die Ergebnisse das Potenzial, welches ein faziales NGP bei der Charakterisierung und in der Diagnostik seltener, genetisch-bedingter, syndromaler Erkrankungen haben kann.

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

https://doi.org/10.1038/s41431-019-0469-3

2.3 PEDIA: Priorisierung von Exomdaten durch Bildanalyse

Hsieh TC*, **Mensah MA***, Pantel JT, Aguilar D, Bar O, Bayat A, Becerra-Solano L, Bentzen HB, Biskup S, Borisov O, Braaten O, Ciaccio C, Coutelier M, Cremer K, Danyel M, Daschkey S, Eden HD, Devriendt K, Wilson S, Douzgou S, Đukić D, Ehmke N, Fauth C, Fischer-Zirnsak B, Fleischer N, Gabriel H, Graul-Neumann L, Gripp KW, Gurovich Y, Gusina A, Haddad N, Hajjir N, Hanani Y, Hertzberg J, Hoertnagel K, Howell J, Ivanovski I, Kaindl A, Kamphans T, Kamphausen S, Karimov C, Kathom H, Keryan A, Knaus A, Köhler S, Kornak U, Lavrov A, Leitheiser M, Lyon GJ, Mangold E, Reina PM, Carrascal AM, Mitter D, Herrador LM, Nadav G, Nöthen M, Orrico A, Ott CE, Park K, Peterlin B, Pölsler L, Raas-Rothschild A, Randolph L, Revencu N, Fagerberg CR, Robinson PN, Rosnev S, Rudnik S, Rudolf G, Schatz U, Schossig A, Schubach M, Shanoon O, Sheridan E, Smirin-Yosef P, Spielmann M, Suk EK, Sznajer Y, Thiel CT, Thiel G, Verloes A, Vrecar I, Wahl D, Weber I, Winter K, Wiśniewska M, Wollnik B, Yeung MW, Zhao M, Zhu N, Zschocke J, Mundlos S, Horn D, Krawitz PM

PEDIA: prioritization of exome data by image analysis.

Genet Med. 2019 Dec;21(12):2807-2814. DOI: 10.1038/s41436-019-0566-2

*contributed equally

Bei der Exomanalyse werden phänotypische Informationen zur Priorisierung der Varianten verwendet. Dazu wird der Phänotyp üblicherweise in Form von HPO-Terms kodiert und mittels geeigneter Algorithmen wird eine Liste passender Verdachtsdiagnosen und damit assoziierter Gene erstellt. Varianten, die in diesen Genen liegen und aufgrund ihrer molekularen und populationsgenetischen Eigenschaften potentiell pathogen erscheinen, werden dem Nutzer bevorzugt angezeigt. Diese Priorisierung ist sehr hilfreich, allerdings noch verbesserungsfähig.

In dieser Studie testeten wir, ob ein faziales NGP als zusätzliche phänotypische Informationsquelle bei der Priorisierung von Exomdaten dienen kann, um die Sensitivität der Exomanalyse zu steigern.

Für die HPO-Term basierte phänotypische Analyse verwendeten wir den Phenomizer, den Bayesian-Ontology-Querying-for-Acurate-Comparison-(BOQA)-Algorithmus und FeatureMatch. Zur Bewertung der molekularen Eigenschaften der Varianten wurde der *combined-annotation-dependent-depletion*-(CADD)-Score und für das faziale NGP DeepGestalt verwendet. Die erzielten numerischen Scores wurden mittels einer Support-Vector-Maschine (SVM) zu einem Wert, dem Prioritization-of-exome-data-by-image analysis-

(PEDIA)-Score integriert. Als Trainingsdaten dienten Fotos von Patienten mit einer seltenen, dysmorphen Erkrankung und gesicherter molekularer Diagnose und Exomdaten unauffälliger, gesunder Probanden, denen die pathogenen Mutationen der Patientenkohorte hinzugefügt wurden. Die Sensitivität von DeepGestalt, des CADD-Scores, der üblichen Exomauswertung (CADD + Phenomizer) und des PEDIA-Scores wurden gemessen.

Es zeigte sich, dass die Sensitivität der bloßen (ohne zusätzliche phänotypische oder Sequenzinformationen) automatisierten Analyse eines Patientenfotos durch DeepGestalt ähnlich groß war wie die Sensitivität der Exomdatenanalyse nur mittels CADD (ohne phänotypische Informationen). Ein Ansatz zur Exomdatenpriorisierung, der phänotypische (HPO-Terms) und genotypische Informationen (CADD-Scores) zusammenführte, hatte eine deutlich höhere Sensitivität. Die höchste Sensitivität erzielte allerdings der PEDIA-Score.

Diese Ergebnisse zeigen, dass die Integration von Systemen zum fazialen NGP einen zusätzlichen Nutzen bei der automatisierten Analyse von Exomdaten haben kann.

ARTICLE | Genetics inMedicine

Open

PEDIA: prioritization of exome data by image analysis

A full list of authors and affiliations appears at the end of the paper.

Purpose: Phenotype information is crucial for the interpretation of genomic variants. So far it has only been accessible for bioinformatics workflows after encoding into clinical terms by expert dysmorphologists.

Methods: Here, we introduce an approach driven by artificial intelligence that uses portrait photographs for the interpretation of clinical exome data. We measured the value added by computer-assisted image analysis to the diagnostic yield on a cohort consisting of 679 individuals with 105 different monogenic disorders. For each case in the cohort we compiled frontal photos, clinical features, and the disease-causing variants, and simulated multiple exomes of different ethnic backgrounds.

Results: The additional use of similarity scores from computer-assisted analysis of frontal photos improved the top 1 accuracy rate by more than 20–89% and the top 10 accuracy rate by more than 5–99% for the disease-causing gene.

Conclusion: Image analysis by deep-learning algorithms can be used to quantify the phenotypic similarity (PP4 criterion of the American College of Medical Genetics and Genomics guidelines) and to advance the performance of bioinformatics pipelines for exome analysis.

Genetics in Medicine (2019) 21:2807–2814; https://doi.org/10.1038/s41436-019-0566-2

Keywords: deep learning; computer vision; dysmorphology; variant prioritization; exome diagnostics

INTRODUCTION

Worldwide, more than half a million children born per year have a rare genetic disorder that is suitable for diagnostic evaluation by exome sequencing. This test's unprecedented diagnostic yield is contrasted by the time requirement for variant interpretation. Making phenotypic information—the observable, clinical presentation—computer-readable is key to solving this problem and important for providing clinicians with a much-needed tool for diagnosing genetic syndromes.[1]

To date, the most advanced exome prioritization algorithms combine deleteriousness scores for variants with semantic similarity searches of the clinical description of a patient.[2] The Human Phenotype Ontology (HPO) has become the *lingua franca* for this purpose.[3] However, a facial gestalt for which no term exists and that is simply described as "characteristic" for a certain disease is not suitable for these computational approaches.

Beyond language, capturing indicative patterns through deep-learning approaches has recently gained attention in assessing facial dysmorphism.[4,5] Artificial neural networks measure the similarities of patient photos to hundreds of disease entities. We hypothesized that results of this next-generation phenotyping tool could be used similarly to deleteriousness scores on the molecular level. This would enable us to transition from the dichotomous PP4 criterion "matching phenotype" in the American College of Medical Genetics and Genomics (ACMG) guidelines for variant interpretation to a quantifiable one.[6,7]

We therefore developed an approach to interpret sequence variants integrating results from the next-generation phenotyping tool DeepGestalt. By this means the clinical presentation of an individual is not only assessed by a human expert clinician, but also by using an artificial intelligence approach on the basis of frontal photographs. In short, we call this approach prioritization of exome data by image analysis (PEDIA).

MATERIALS AND METHODS

We compiled a cohort comprising 679 individuals with frontal facial photographs and clinical features documented in HPO terminology.[3] The diagnoses of all individuals have previously been confirmed molecularly and are suitable for analysis by exome sequencing. In total, the cohort covers 105 different monogenic syndromes linked to 181 different genes. Of the individuals in this cohort, 446 were published and 233 have not been previously reported (see PMID column in Supplementary Table 1).

The study was approved by the ethics committees of the Charité–Universitätsmedizin Berlin and of the University

Correspondence: Peter M. Krawitz (pkrawitz@uni-bonn.de)
These authors contributed equally: Tzung-Chien Hsieh, Martin A. Mensah

Submitted 13 December 2018; accepted: 23 May 2019
Published online: 5 June 2019

Hospital Bonn. Written informed consent was given by the patients or their guardians, including permission to publish photographs. Easy to understand, transparent information with both text and illustrations about the pattern recognition in our algorithm that processes personal data in the form of 2D portrait photographs can be found at https://www.pedia-study.org/documents. Through technical and organizational measures (privacy by design), we process the photos and the data obtained from them in the least identifiable manner necessary for achieving the purpose. This respects the data minimization principle of data being adequate, relevant, and limited.

In addition to the PEDIA data set, we analyzed a subset of the DeepGestalt study. By removing disorders that are confirmed by tests other than exome sequencing, such as Down syndrome (Supplementary Table 2), we ended up with 260 of 329 cases from the DeepGestalt set.[5]

The facial images were analyzed with DeepGestalt, a deep convolutional neural network trained on more than 17,000 patient images.[5] The results of this analysis are gestalt scores that quantify the similarity to 216 different rare phenotypes per individual. These vectors can also be used to identify duplicates in the DeepGestalt training set and test set without the need to access the original photos. To avoid overfitting, we excluded all cases of the PEDIA cohort from a DeepGestalt model that we used for benchmarking. It is noteworthy that the version of DeepGestalt available at Face2Gene will not yield the same results when photos of the PEDIA cohort are reanalyzed because it is built as a framework that aims to learn from every solved case.

In addition to the image analysis, we performed semantic similarity searches with the annotated HPO terms by three different tools: Feature Match (FDNA), Phenomizer, and Bayesian Ontology Querying for Accurate Comparisons (BOQA).[8,9] HPO terms for all published cases as well as the clinical notes in the electronic health records were independently extracted by two data curators. All terms that did not occur in both lists were revisited by a third curator (see Fig. 1a and Supplementary Table 1). The similarity scores from image analysis as well as semantic similarity searches were mapped to genes by mim2gene and morbidmap from OMIM.[10] If there were several syndromes linked to a gene, the highest gestalt and feature scores were selected for this gene.

Exome sequencing data was not available for the vast majority of cases. Therefore, we spiked in the disease-causing variant of each case into randomly selected exomes of healthy

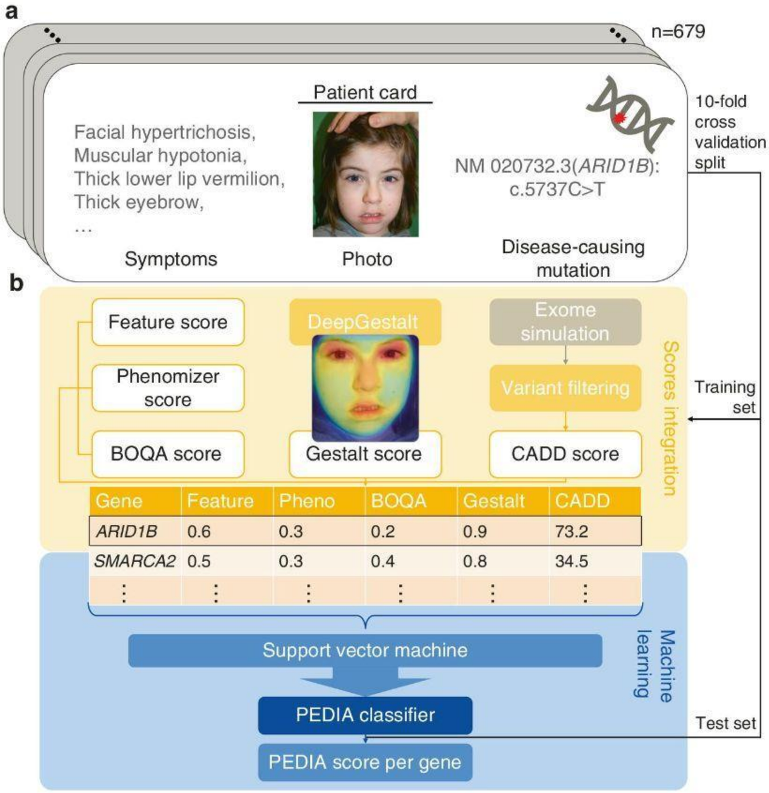


Fig. 1 Prioritization of exome data by image analysis (PEDIA): cohort and classification approach. (**a**) Clinical features, facial photograph, and pathogenic variant of one individual of the PEDIA cohort. In total the cohort consists of 679 cases with monogenic disorders that are suitable for a diagnostic workup by exome sequencing. (**b**) Clinical features, images, and exome variants were evaluated separately and integrated to a single score by a machine learning approach. The disease-causing gene is shown at the top of the list.

individuals of different ethnicities from the 1000 Genomes Project.[11] All sequence variants were then filtered as described by Wright et al. and scored for deleteriousness with CADD.[12,13] Per gene, the variant with the highest CADD score was used, regardless of the genotype. This heuristic was chosen to maximize the sensitivity also for compound heterozygous cases where the second hit in a recessive disease gene achieves only a relatively low CADD score.

For each case this procedure resulted in a table with rows for genes and the five different scores in the columns (Fig. **1b**). All five scores per line as well as the Boolean label disease gene "true" or "false" (i.e., the vector) were used to train a classifier that yields a single value per gene, the PEDIA score, that can be used for prioritization (Fig. **1b**). A detailed description of preprocessing and filtering, as well as all the annotated data, can be found in our code repository.

We used a support vector machine (SVM) to prioritize the genes based on the five scores for each case. To benchmark our approach, we performed tenfold cross-validation. First, we split the PEDIA cohort into ten groups, ensuring that a certain disease gene was included only in one of ten groups. By this means, we avoided overfitting, in case the same disease-causing variant occurred in two different individuals (Supplementary Fig. 1). We used a linear kernel on the five scores to train the SVM and selected the hyperparameter C in the range from 2^{-6} to 2^{12} by performing internal fivefold cross-validation on the training set. The C with the highest top 1 accuracy was selected for training a linear SVM. We further benchmarked the performance of each case in the test set with this model. The distance of each gene to the hyperplane—defined as the PEDIA score—was used to rank the genes for the case. If the disease-causing gene was at the first position, we called it a top 1 match, or if it was among the first ten genes, we considered it a top 10 match.

For the 260 cases from the DeepGestalt publication test set, where exome diagnostics would be applicable, we randomly selected cases from the PEDIA cohort with the same diagnosis and added the CADD and the feature scores per case (see column C in Supplemental Table 1). The cases in the PEDIA cohort with the same pathogenic variant as already assigned to the DeepGestalt test set were removed from the training set. Then we trained the classifier on the PEDIA cohort and tested it on the DeepGestalt publication test set. The experiment was repeated ten times with random selection. By this means we studied how the publicly available portraits of the DeepGestalt test set would improve the performance when used in exome analysis with the PEDIA approach. However, it has to be emphasized that both approaches solve different multiclass classification problems (MCPs), the first tool operating on phenotypes and the second on genes. The difficulty of the task is not only characterized by the number of classes and the distinguishability of the different entities but also by the information available for the classification. For both MCPs the maximum number of classes can be estimated from OMIM by querying with the HPO term "abnormal facial shape", yielding around 700 disorders and genes with disease-causing variants. As there is additional and nonredundant information available from the molecular level for PEDIA, it achieves better top 1 and top 10 accuracies.

CODE AVAILABILITY

All training data as well as the classifier are available at https://github.com/PEDIA-Charite/PEDIA-workflow. The trained PEDIA model is provided as a service that is ready to use at https://pedia-study.org.

RESULTS

The performance of a prioritization tool can be assessed by the proportion of cases for which the correct diagnosis or disease gene is placed at the first position or among the first ten suggestions (top 1 and top 10 accuracy). The composition of the test set has an influence on the accuracy because some disease phenotypes are easier to recognize, and some gene variants are more readily identified as deleterious. The setup of the PEDIA cohort, which is comprehensively documented in the Supplementary Appendix, therefore aims at emulating the whole spectrum of cases that could be analyzed with DeepGestalt and diagnosed by exome sequencing.

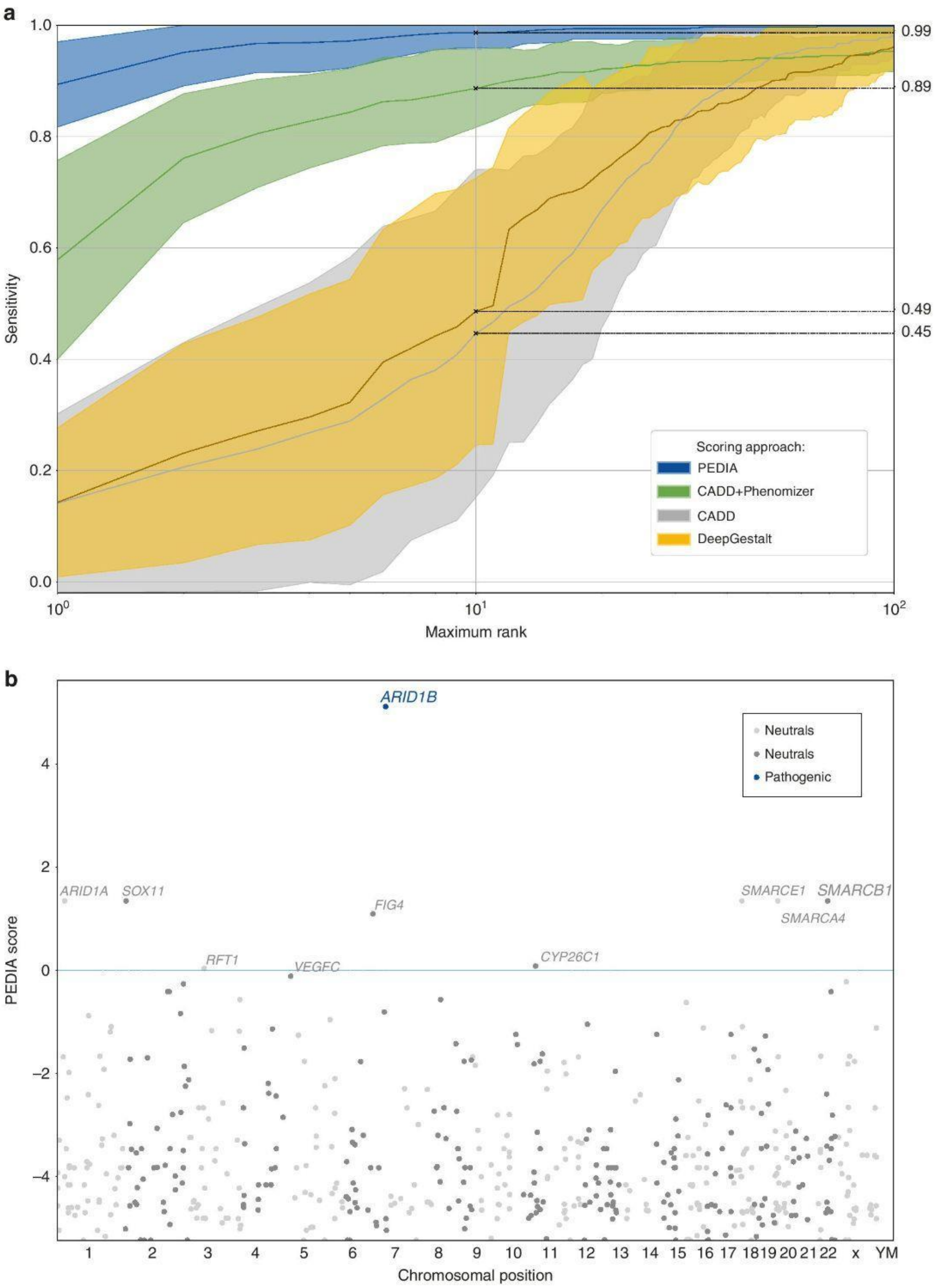
When only CADD scores are used for variant ranking, the disease-causing gene is in the top 10 in less than 45% of all tested cases. The top 10 accuracy increases up to 63–94%, when different semantic similarity scores based on HPO feature annotations are included (Supplementary Table 3).

The additional information from frontal photos of cases pushes the correct disease gene to the top 10 in 99% of all PEDIA cases (Fig. **2a**). Particularly striking is the performance gain for the top 1 accuracy rate from 36–74% without DeepGestalt scores to 86–89% including the scores from image analysis (Supplementary Table 3).

The distribution of the PEDIA scores does not differ using exomes with different ethnic backgrounds (Supplementary Fig. 2).

Although the top 10 accuracies of DeepGestalt scoring on the phenotype level and PEDIA scoring on the gene level cannot be compared directly, both approaches operate on a similar number of classes (Fig. 2). Adding suitable molecular information to 260 cases from the DeepGestalt publication test set confirms our results in the PEDIA cohort by achieving a top 10 accuracy rate of 99% (Supplementary Table 2).

The value of a frontal photograph is demonstrated by a case with Coffin–Siris syndrome (shown in Fig. 1): the characteristic facial features are relatively mild, so the correct diagnosis is only listed as the third suggestion by DeepGestalt. Among all the variants encountered in the exome, the disease-causing gene *ARID1B* would only achieve rank 27, if scored by the molecular information alone. However, combined with the phenotypic information, the PEDIA approach lists this gene as the first candidate (Fig. **2b**).



Although the diagnosis of the illustrated case could be molecularly confirmed by a directed single-gene test in other instances where the facial gestalt is more indicative, syndromic disorders often puzzle clinicians due to their high phenotypic variability. In the Deciphering Developmental Disorders (DDD) project many syndromes were diagnosed only after exome sequencing.[14] Still, the top 10 accuracy rate of 49% that DeepGestalt can achieve for phenotypes linked to genes is impressive (Fig. **2a**). The contribution from the different sources of evidence to the PEDIA score is also

Fig. 2 Performance readout and visualization of test results for a representative prioritization of exome data by image analysis (PEDIA) case. (**a**) For each case the exome variants are ordered according to four different scoring approaches, solely by a molecular deleteriousness score (CADD), by a score from image analysis (DeepGestalt), by a combination of a molecular deleteriousness score and a clinical feature–based semantic similarity score (CADD +Phenomizer), or the PEDIA score that includes all three levels of evidence. The sensitivity of the prioritization approach depends on the number of genes that are considered in an ordered list. The top 1 and top 10 accuracy rates correspond to the intersection of the curves at maximum rank 1 and 10. Note that for benchmarking DeepGestalt on the gene level, syndrome similarity scores first have to be mapped to the gene level, resulting in a lower performance compared with the readout on a phenotype level, due to heterogeneity. The area under the curve is largest for PEDIA scoring. (**b**) The disease-causing gene of the case depicted in Fig. 1 achieves the highest PEDIA score and molecularly confirms the diagnosis of Coffin–Siris syndrome. Other genes associated with similar phenotypes, such as Nicolaides–Baraitser syndrome, also achieved high scores for gestalt but not for variant deleteriousness.

reflected by the relative weight of the deleteriousness of the pathogenic variant (0.44), all feature-based scores combined (0.25), and the results from image analysis by DeepGestalt (0.31) that can be derived from a linear SVM model. The information contained in a frontal photograph of a patient therefore goes beyond what clinical terms can capture. The top 1 and top 10 accuracies are reported for all combinations of scores in the Supplementary Table 3.

DISCUSSION

The guidelines for variant classification in the laboratory follow a qualitative heuristic that combines distinct types of evidence (functional, population, phenotype, etc.). Interestingly, it is also compatible with Bayesian statistics[7] and the advantage of such a framework is that continuous evidence types can be integrated into the classification system. While in silico predictions about a variant's pathogenicity have a relatively long history in bioinformatics and machine learning, the quantification of phenotypic raw data such as facial images with artificial intelligence systems has just begun: the PEDIA approach uses scores from DeepGestalt for gene prioritization in combination with quantitative scores from the molecular level in Mendelian disorders identifiable by exome sequencing.

Interestingly, the ethnicity, which affects the number of variant calls or the deleterious variant load, had minor influence on the performance of PEDIA. Although the total number of variants detected by reference-guided sequencing in individuals of African descent is considerably higher than in individuals of European or Asian descent, the distribution of the CADD scores for rare variants is comparable (Supplementary Figs. 3, 4). That means the rank that a gene achieves due to the molecular score and the corresponding scores from the phenotypic information is hardly affected by the background population (Supplementary Fig. 2).

With regard to the routine use in the laboratory we have learned three important lessons from specific subgroups or cases achieving lower PEDIA ranks:

1. Although DeepGestalt, the convolutional neural network used for image analysis, has been pretrained on real-world uncontrolled 2D images, patient photographs that were not frontal, of low resolution, had poor lightening and contrast, or contained artifacts such as glasses, yielded lower gestalt scores for the searched disorder. In one use case envisioned for PEDIA, the human expert in the lab will only receive the similarity scores from DeepGestalt, but not the original photograph. In this setting it is not clear whether low scores originate from a low-quality photograph or whether there is little dysmorphic signal indicative of a syndromic disorder. This potential problem could be addressed by providing gestalt scores from additional photographs.

2. Particularly rare diseases or recently described disorders, for which the classifier's representation is based on a smaller training set, show a lower performance, even if experienced dysmorphologists would consider them highly distinguishable. In a recent publication by Duddin-Byth et al. the machine learning approach showed the lowest accuracy for the disorder with the smallest number of training cases; however, so did humans.[15]

3. Disease-causing variants in genes that interact in a molecular pathway often result in highly similar phenotypes that are organized as series in OMIM and modeled as a single entity by DeepGestalt. Often there are subtle gene-specific differences in the gestalt and modeling the entire phenotypic series by a single class is not the theoretical optimum achievable with more cases.[16,17] This will especially diminish the performance of genes less frequently mutated in a molecular pathway. This is exemplified in the PEDIA cohort by Hyperphosphatasia with Mental Retardation Syndrome (HPMRS), where the least frequently mutated gene, *PGAP2*, shows the lowest performance. Likewise, this applies to microdeletion syndromes that can also be caused by pathogenic variants in single genes, such as Smith–Magenis syndrome, or an atypical clinical presentation with Kabuki syndrome (see e.g., case IDs 246245 and 204233 in Supplementary Table 1).[18]

It is noteworthy that these shortcomings are mainly due to the limited training data for these particular genes and that they will most likely be overcome by more molecularly confirmed cases. DeepGestalt and PEDIA are therefore built as frameworks that will be improved continuously with additional data. In general, the use of artificial intelligence in medical sciences raises new or exacerbates existing ethical and legal issues as repositories of combined genotype and phenotype data become crucial for the machine learning community.[19,20] Sharing portrait photos of individuals with rare diseases can be accomplished within the scope of even the most elaborate data privacy laws, such as the European Union General Data Protection Regulation 2016/679 (GDPR). The GDPR not only ensures the protection of individuals, but also the free movement of personal data, inter alia, for scientific research purposes.[21]

The interpretation of genetic variants is greatly facilitated by sequencing additional family members. Analogously, we hypothesize that the signal-to-noise ratio of next-generation phenotyping technologies can further be improved by including unaffected siblings or parents in the analysis.

We include and strive to include a wide variety of ethnicities, but European backgrounds are currently best represented, leading to best performance for this population. As the data set expands further, the algorithm will improve for currently underrepresented ethnicities.

Assistance with diagnosis of rare genetic disorders is highly valuable to clinicians, and by extension to the patients themselves and their families. Especially in inconclusive cases with findings of unknown clinical significance, additional evidence from computer-assisted analysis of medical imaging data could be a decisive factor.[13]

In conclusion, the PEDIA study documents that exome variant interpretation benefits from computer-assisted image analysis of facial photographs. By including similarity scores from DeepGestalt, we improved the top 10 accuracy rate significantly compared with state-of-the-art algorithms. Artificial intelligence–driven pattern recognition of frontal facial patient photographs is therefore an example of next-generation phenotyping technology that has proven its clinical value for the interpretation of next-generation sequencing data.[22]

SUPPLEMENTARY INFORMATION

The online version of this article (https://doi.org/10.1038/s41436-019-0566-2) contains supplementary material, which is available to authorized users.

ACKNOWLEDGEMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (KR 3985/7-3, KR 3985/6-1). M.A.M. is participant in the BIH Charité Junior Clinician Scientist Program funded by the Charité–Universitätsmedizin Berlin and the Berlin Institute of Health.

DISCLOSURE

P.M.K. and K.W.G. receive compensation as consultants for FDNA Inc. H.D.E., Y.H., G.N., O. Bar, O.S., Y.G., N.F. are employees of FDNA; T.K. is an employee of GeneTalk GmbH. The other authors declare no conflicts of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Biesecker LG, Green RC. Diagnostic clinical genome and exome sequencing. N Engl J Med. 2014;371:1170.
2. Pengelly RJ, et al. Evaluating phenotype-driven approaches for genetic diagnoses from exomes in a clinical setting. Sci Rep. 2017;7:13509.
3. Robinson PN, et al. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. Am J Hum Genet. 2008;83:610–615.
4. Ferry Q, et al. Diagnostically relevant facial gestalt information from ordinary photos. eLife. 2014;3:e02020.
5. Gurovich Y, et al. DeepGestalt—identifying rare genetic syndromes using deep learning. Nat Med. 2019;25:60–64.
6. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–424.
7. Tavtigian SV, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. Genet Med. 2018;20:1054–1060.
8. Köhler S, et al. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. Am J Hum Genet. 2009;85:457–464.
9. Bauer S, Köhler S, Schulz MH, Robinson PN. Bayesian ontology querying for accurate and noise-tolerant semantic searches. Bioinformatics. 2012;28:2502–2508.
10. Online Mendelian Inheritance in Man (OMIM). Baltimore, MD: McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins Universrity (Baltimore, MD); 2019. https://omim.org/.
11. The 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526:68.
12. Kircher M, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–315.
13. Wright CF, et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. Genet Med. 2018;20:1216–1223.
14. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017;542:433–438.
15. Dudding-Byth T, et al. Computer face-matching technology using two-dimensional photographs accurately matches the facial gestalt of unrelated individuals with the same syndromic form of intellectual disability. BMC Biotechnol. 2017;17:90.
16. Pantel JT, et al. Advances in computer-assisted syndrome recognition by the example of inborn errors of metabolism. J Inherit Metab Dis. 2018;41:533–539.
17. Knaus A, et al. Characterization of glycosylphosphatidylinositol biosynthesis defects by clinical features, flow cytometry, and automated image analysis. Genome Med. 2018;10:3.
18. Badalato L, et al. KMT2D p.Gln3575His segregating in a family with autosomal dominant choanal atresia strengthens the Kabuki/CHARGE connection. Am J Med Genet A. 2017;173:183–189.
19. Hallowell DPH, et al. Big data phenotyping in rare diseases: some ethical issues. Genet Med. 2019;21:272–274.
20. Mascalzoni D, et al. Are requirements to deposit data in research repositories compatible with the European Union's General Data Protection Regulation? Ann Intern Med. 2019;170:332–334.
21. Bentzen HB, Høstmælingen N. Balancing protection and free movement of personal data: the new European Union General Data Protection Regulation. Ann Intern Med. 2019;170:335–337.
22. Hennekam R, Biesecker LG. Next-generation sequencing demands next-generation phenotyping. Hum Mutat. 2012;33:884–886.

Tzung-Chien Hsieh, MS[1,2,3], Martin A. Mensah, MD[2,3], Jean T. Pantel, cand.med.[1,2,3], Dione Aguilar, MD[4], Omri Bar, MS[5], Allan Bayat, MD[6], Luis Becerra-Solano, MD[7], Heidi B. Bentzen, LLM[8], Saskia Biskup, MD[9], Oleg Borisov, MD[1], Oivind Braaten, MD[10], Claudia Ciaccio, MD[11], Marie Coutelier, MD[2], Kirsten Cremer, MD[12], Magdalena Danyel, MD[2], Svenja Daschkey, PhD[13], Hilda David Eden, MS[5], Koenraad Devriendt, MD[14], Sandra Wilson, MD[15], Sofia Douzgou, MD[16,17], Dejan Đukić, MS[1], Nadja Ehmke, MD[2], Christine Fauth, MD[18], Björn Fischer-Zirnsak, PhD[2], Nicole Fleischer, MS[5], Heinz Gabriel, MD[19], Luitgard Graul-Neumann, MD[2], Karen W. Gripp, MD[20], Yaron Gurovich, MS[5], Asya Gusina, MD[21], Nechama Haddad, MS[2], Nurulhuda Hajjir, MD[2], Yair Hanani, MS[5], Jakob Hertzberg, MS[2], Konstanze Hoertnagel, MD[9], Janelle Howell, MD[22], Ivan Ivanovski, MD[23], Angela Kaindl, MD[24], Tom Kamphans, PhD[25], Susanne Kamphausen, MD[26], Catherine Karimov, MD[27], Hadil Kathom, MD[28], Anna Keryan, MD[27], Alexej Knaus, PhD[1], Sebastian Köhler, PhD[29], Uwe Kornak, MD[2], Alexander Lavrov, MD[30], Maximilian Leitheiser, cand.med.[2], Gholson J. Lyon, MD[31], Elisabeth Mangold, MD[32], Purificación Marín Reina, MD[33], Antonio Martinez Carrascal, MD[34], Diana Mitter, MD[35], Laura Morlan Herrador, MD[36], Guy Nadav, MS[5], Markus Nöthen, MD[12], Alfredo Orrico, MD[37], Claus-Eric Ott, MD[2], Kristen Park, MD[38], Borut Peterlin, MD[39], Laura Pölsler, MD[18], Annick Raas-Rothschild, MD[40], Linda Randolph, MD[27], Nicole Revencu, MD[41], Christina Ringmann Fagerberg, MD[42], Peter Nick Robinson, MD[43], Stanislav Rosnev, cand.med.[2], Sabine Rudnik, MD[18], Gorazd Rudolf, MD[39], Ulrich Schatz, MD[18], Anna Schossig, MD[18], Max Schubach, PhD[3], Or Shanoon, MS[5], Eamonn Sheridan, MD[44], Pola Smirin-Yosef, MD[45], Malte Spielmann, MD[2], Eun-Kyung Suk, MD[46], Yves Sznajer, MD[47], Christian T. Thiel, MD[48], Gundula Thiel, MD[46], Alain Verloes, MD[49], Irena Vrecar, MD[39], Dagmar Wahl, MD[50], Ingrid Weber, MD[18], Korina Winter, MD[2], Marzena Wiśniewska, MD[51], Bernd Wollnik, MD[52], Ming W. Yeung, MS[1], Max Zhao, cand.med.[2], Na Zhu, PhD[2], Johannes Zschocke, MD[18], Stefan Mundlos, MD[2], Denise Horn, MD[2] and Peter M. Krawitz, MD[1]

[1]Institute of Genomic Statistics and Bioinformatics, University of Bonn, Bonn, Germany. [2]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Medical Genetics and Human Genetics, Berlin, Germany. [3]Berlin Institute of Health (BIH), Berlin, Germany. [4]Centro de Cáncer de Mama, Tecnológico de Monterrey, Monterrey, Mexico. [5]FDNA Inc., Boston, MA, USA. [6]Rigshospitalet, Department of Neurology, Copenhagen, Denmark. [7]Unidad de Investigación Médica en Medicina Reproductiva, Mexico City, Mexico. [8]Centre for Medical Ethics, Faculty of Medicine and the Norwegian Research Center for Computers and Law, Faculty of Law, University of Oslo, Oslo, Norway. [9]CeGaT GmbH, Tübingen, Germany. [10]Faculty of Medicine, Department of Medical Genetics, University of Oslo, Blindern, Oslo, Norway. [11]Developmental Neurology Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy. [12]Department of Human Genetics, University Hospital of Bonn, Bonn, Germany. [13]Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [14]Department of Human Genetics, KU Leuven, Leuven, Belgium. [15]Department of Human Genetics, University of Hamburg, Hamburg, Germany. [16]Manchester Centre for Genomic Medicine, St Mary's Hospital, Central Manchester University Hospitals NHS Foundation Trust Manchester Academic Health Sciences Centre, Manchester, United Kingdom. [17]School of Biological Sciences, Division of Evolution and Genomic Sciences, University of Manchester, Manchester, United Kingdom. [18]Division of Human Genetics, Medical University of Innsbruck, Innsbruck, Austria. [19]Center for Genomics and Transcriptomics, Eberhard Karls University of Tübingen, Tübingen, Germany. [20]A. I. duPont Hospital for Children, Wilmington, DE, USA. [21]National Research and Applied Medicine Centre 'Mother and Child'', Minsk, Belarus. [22]Lineagen, Salt Lake City, Utah, USA. [23]Clinical Genetics Unit, AUSL-IRCCS Reggio Emilia, Reggio Emilia, Italy. [24]Center for Chronically Sick Children (Sozialpädiatrisches Zentrum, SPZ), Charité - Universitätsmedizin Berlin, Berlin, Germany. [25]GeneTalk, Bonn, Germany. [26]University Hospital Magdeburg, Magdeburg, Germany. [27]Children's Hospital of Los Angeles, Los Angeles, CA, USA. [28]Department of Pediatrics, Medical University of Sofia, Sofia, Bulgaria. [29]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, NeuroCure Clinical Research Center, Berlin, Germany. [30]Research Institute of Medical Genetics of Russian Academy of Medical Sciences, Moscow, Russian Federation. [31]Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, Woodbury, New York, USA. [32]Institute of Human Genetics, University of Bonn, Bonn, Germany. [33]Hospital General Universitario De Valencia, Valencia, Spain. [34]Hospital General De Requena, Servicio Pediatría, Spain. [35]University Hospital Leipzig, Leipzig, Germany. [36]Hospital Universitario Miguel Servet, Zaragoza, Spain. [37]Azienda Ospedaliera Universitaria Senese, Siena, Italy. [38]Department of Pediatrics and Neurology, University of Colorado School of Medicine, Colorado, Aurora, USA. [39]Clinical Institute of Medical Genetics, University Medical Centre Ljubljana, Ljubljana, Slovenia. [40]The Danek Gertner Institute of Human Genetics, Sheba Medical Center, Tel-Hashomer, Israel. [41]Center for Human Genetics, University Hospital, Université Catholique de Louvain, Brussels, Belgium. [42]Odense University Hospital, Odense, Denmark. [43]The Jackson Laboratory for Genomic Medicine, Farmington,

CT, USA. [44]School of Medicine, University of Leeds, Leeds, United Kingdom. [45]Genomic Bioinformatics Laboratory, Department of Molecular Biology, Ariel University, Ariel, Israel. [46]Center for Prenatal Diagnosis and Human Genetics, Berlin, Germany. [47]Cliniques universitaires Saint Luc UCL, Bruxelles, Belgium. [48]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg FAU, Erlangen, Erlangen, Germany. [49]Hopital Robert Debré, Paris, France. [50]Center for Human Genetics and Laboratory Diagnostics Dr. Klein, Dr. Rost and Colleagues, Martinsried, Germany. [51]Poznañ University of Medical Sciences, Poznañ, Poland. [52]University Medical Center Göttingen, Göttingen, Germany

2.4 Effizienz der computergestützten fazialen Phänotypisierung (DeepGestalt) bei Personen mit und ohne genetisches Syndrom: Studie zur diagnostischen Genauigkeit

Pantel JT*, Hajjir N*, Danyel M, Elsner J, Abad-Perez AT, Hansen P, Mundlos S, Spielmann M, Horn D, Ott CE, **Mensah MA**.

Efficiency of Computer-Aided Facial Phenotyping (DeepGestalt) in Individuals With and Without a Genetic Syndrome: Diagnostic Accuracy Study.

J Med Internet Res. 2020 Oct 22;22(10):e19263. DOI: 10.2196/19263

Die meisten Studien zur Genauigkeit von DeepGestalt und anderen fazialen NGP-Systemen untersuchen die Sensitivität, d.h. die Fähigkeit des Systems, Personen, die von einem seltenen, dysmorphen Syndrom betroffen sind, die korrekte Diagnose zuzuordnen. Um die Genauigkeit der Systeme - und damit ihren potentiellen Nutzen in der klinischen Routine - beurteilen zu können, ist allerdings auch eine Evaluation der Spezifität, d.h. der Fähigkeit des Systems, unauffällige Probanden als solche zu erkennen, notwendig. Aus technischen Gründen verfügt DeepGestalt nicht über eine Klasse "unauffälliges Gesicht". Um die Spezifität von DeepGestalt zu beurteilen, überprüften wir in dieser Studie daher, ob sich die jeweils höchsten Gestalt Scores, die den einzelnen Bildern zugeordnet wurden, im Mittel zwischen Bildern von Betroffenen und dazu nach Alter, Geschlecht und ethnischem Hintergrund passenden, unauffälligen Kontrollbildern unterschieden. Außerdem überprüften wir, ob sich eine SVM, die als binärer Klassifikator (unauffällig vs. dysmorph) auf DeepGestalts Mehr-Klassen-Output trainiert wurde, zur Steigerung der Spezifität einsetzen ließe. Wir testeten auch, wie viele Diagnosen DeepGestalt vergeben kann und ob und wie DeepGestalts Sensitivität und Spezifität von den jeweiligen Diagnosen abhingen. Ferner überprüften wir, ob der ethnische Hintergrund einer Person Einfluss auf die Genauigkeit des Systems hat.

DeepGestalt schlug insgesamt 238 Syndrome vor. Die Wahrscheinlichkeit, mit der ein Syndrom vorgeschlagen wurde, schwankte dabei erheblich. 11 Syndrome wurden bei mehr als 60% der Kontrollbilder in den Ergebnislisten aufgeführt. Am häufigsten fand sich (Falsch-Positiven-Rate >80%) der Vorschlag eines Fragilen-X-Syndroms. Auch die Sensitivität war syndromabhängig, unter den getesteten Syndromen war sie am besten für das Treacher-Collins-Syndrom und am schlechtesten für das Loeys-Dietz-Syndrom. Gestalt Scores von Bildern von Betroffenen waren im Mittel höher als bei unauffälligen Kontrollen. Allerdings fand sich eine relevante Überschneidung der Scoreverteilungen beider Gruppen. Nach Transformation mit einer SVM konnte diese Überlappung reduziert werden. Der ethnische Hintergrund hatte keinen wesentlichen Einfluss auf die Genauigkeit des Systems.

Original Paper

Efficiency of Computer-Aided Facial Phenotyping (DeepGestalt) in Individuals With and Without a Genetic Syndrome: Diagnostic Accuracy Study

Jean Tori Pantel[1,2*]; Nurulhuda Hajjir[1,3*]; Magdalena Danyel[1,4]; Jonas Elsner[1]; Angela Teresa Abad-Perez[1]; Peter Hansen[1,5], PhD; Stefan Mundlos[1,6], Prof Dr; Malte Spielmann[6,7], Prof Dr; Denise Horn[1], Prof Dr; Claus-Eric Ott[1], MD; Martin Atta Mensah[1,8], MD

[1]Institute of Medical Genetics and Human Genetics, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

[2]Institute for Genomic Statistics and Bioinformatics, University Hospital Bonn, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany

[3]Klinik für Pädiatrie mit Schwerpunkt Gastroenterologie, Nephrologie und Stoffwechselmedizin, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

[4]Berlin Center for Rare Diseases, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health, Berlin, Germany

[5]The Jackson Laboratory for Genomic Medicine, Farmington, CT, United States

[6]RG Development & Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany

[7]Institute of Human Genetics, University of Lübeck, Lübeck, Germany

[8]Berlin Institute of Health, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Martin Atta Mensah, MD
Institute of Medical Genetics and Human Genetics
Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health
Berlin
Germany
Phone: 49 30 450 569 132
Fax: 49 30 450 569 914
Email: martin-atta.mensah@charite.de

Abstract

Background: Collectively, an estimated 5% of the population have a genetic disease. Many of them feature characteristics that can be detected by facial phenotyping. Face2Gene CLINIC is an online app for facial phenotyping of patients with genetic syndromes. DeepGestalt, the neural network driving Face2Gene, automatically prioritizes syndrome suggestions based on ordinary patient photographs, potentially improving the diagnostic process. Hitherto, studies on DeepGestalt's quality highlighted its sensitivity in syndromic patients. However, determining the accuracy of a diagnostic methodology also requires testing of negative controls.

Objective: The aim of this study was to evaluate DeepGestalt's accuracy with photos of individuals with and without a genetic syndrome. Moreover, we aimed to propose a machine learning–based framework for the automated differentiation of DeepGestalt's output on such images.

Methods: Frontal facial images of individuals with a diagnosis of a genetic syndrome (established clinically or molecularly) from a convenience sample were reanalyzed. Each photo was matched by age, sex, and ethnicity to a picture featuring an individual without a genetic syndrome. Absence of a facial gestalt suggestive of a genetic syndrome was determined by physicians working in medical genetics. Photos were selected from online reports or were taken by us for the purpose of this study. Facial phenotype was analyzed by DeepGestalt version 19.1.7, accessed via Face2Gene CLINIC. Furthermore, we designed linear support vector machines (SVMs) using Python 3.7 to automatically differentiate between the 2 classes of photographs based on DeepGestalt's result lists.

Results: We included photos of 323 patients diagnosed with 17 different genetic syndromes and matched those with an equal number of facial images without a genetic syndrome, analyzing a total of 646 pictures. We confirm DeepGestalt's high sensitivity (top 10 sensitivity: 295/323, 91%). DeepGestalt's syndrome suggestions in individuals without a craniofacially dysmorphic syndrome followed a nonrandom distribution. A total of 17 syndromes appeared in the top 30 suggestions of more than 50% of nondysmorphic images. DeepGestalt's top scores differed between the syndromic and control images (area under the receiver operating characteristic [AUROC] curve 0.72, 95% CI 0.68-0.76; *P*<.001). A linear SVM running on DeepGestalt's result vectors showed stronger differences (AUROC 0.89, 95% CI 0.87-0.92; *P*<.001).

Conclusions: DeepGestalt fairly separates images of individuals with and without a genetic syndrome. This separation can be significantly improved by SVMs running on top of DeepGestalt, thus supporting the diagnostic process of patients with a genetic syndrome. Our findings facilitate the critical interpretation of DeepGestalt's results and may help enhance it and similar computer-aided facial phenotyping tools.

(J Med Internet Res 2020;22(10):e19263) doi: 10.2196/19263

KEYWORDS

facial phenotyping; DeepGestalt; facial recognition; Face2Gene; medical genetics; diagnostic accuracy; genetic syndrome; machine learning

Introduction

Background

Although individual genetic diseases are rare, they collectively affect an estimated 5% of a population [1]. Thus, these diseases represent a major challenge for health care systems, as it usually requires highly specialized knowledge to propose a specific genetic diagnosis. Assessing the facial phenotypes of patients with genetic syndromes is key to this diagnostic process [2]. Traditionally performed by a physician, the advents of computer vision and machine learning in medicine enable rapid and automated assessment of a patient's facial traits [3,4]. Numerous facial phenotyping systems have been developed with the potential to aid the diagnostic processes in medical genetics [5-12]. DeepGestalt, the neural network behind Face2Gene CLINIC, which was trained on more than 17,106 images, is thus far the best-investigated and most convenient to use application [11]. Several studies assessed the algorithm's sensitivity, suggesting that it is of a certain quality [11,13-38]. These tests predominantly analyzed images of patients diagnosed with a genetic disorder known to show characteristic facial features. This appears reasonable as DeepGestalt is designed to identify such syndromes. However, it might introduce a bias in conclusions of the system's everyday clinical use since not all individuals seen in a real-life setting belong to the group of patients included in previous studies of DeepGestalt. This may be because (1) the featured syndrome is yet to be analyzed by the system; (2) an individual features a syndrome not associated with a characteristic facies; or (3) an individual has no syndrome at all.

In addition to such evaluations of DeepGestalt's sensitivity, there is a need for studies on its specificity when tested on individuals without craniofacial dysmorphism. As DeepGestalt is not designed to suggest the class label "inconspicuous face" [11], evaluating its clinical specificity is not too trivial a task. Some studies tested the ability of DeepGestalt's methodology to distinguish between facial images with and without a genetic syndrome by constructing user-specific neural networks trained on healthy control images and on images of limited numbers of well-selected genetic disorders using Face2Gene RESEARCH [20,26-28,30,32,34,39-41]. Their results suggested that neural networks such as DeepGestalt may have the potential to differentiate between the 2 classes and may thus be used in diagnosing patients in medical genetics. Such a test could be applied at different stages of the diagnostic process. Patients who want to know if genetic counseling is necessary could use it as a triage test to check whether a suspicion of a genetic disease is justified. Physicians and other medical professionals could similarly use such a test on patients suspected of having a genetic syndrome to narrow down the range of possible diagnoses. Geneticists could use it as an add-on test to further confirm a diagnosis, for example, in the presence of a variant of unknown significance.

Objectives

We aimed to systematically benchmark DeepGestalt's power to discern images of individuals with a dysmorphic genetic syndrome from images of healthy control individuals. For this purpose, we tested the basic prerequisite for the diagnostic usefulness of DeepGestalt, that is, to yield different scores in persons with a conventionally established diagnosis of a genetic syndrome than in persons without a genetic syndrome (H_1: $\mu_{syndromic} \neq \mu_{healthy}$). We also determined DeepGestalt's capacity to distinguish those images by measuring its area under the receiver operating characteristic (AUROC) curve. Furthermore, we aimed to develop and test a machine learning–based approach to improve DeepGestalt's accuracy.

Methods

Selection and Analysis of Portrait Photos

Study Design

To be included in this study, portrait photos had to depict the entire frontal face (from hairline to chin showing both eyes) and no artifact other than glasses. To achieve a vertical positioning of the face, the images were cropped and rotated if necessary. A convenience sample of online accessible images was collected between September 2019 and December 2019, using a methodology adjusted from Ferry et al [8]. Pictures photographed by us were taken at the 2018 meeting of the

Elterninitiative Apertsyndrom und Verwandte Fehlbildungen eV, a parents' initiative on Apert syndrome and related disorders in Germany, after obtaining written informed consents as approved by the ethics committee of the Charité – Universitätsmedizin Berlin (EA2/190/16). Image inclusion was planned before conducting analysis by DeepGestalt. A sample size of the positive and negative class of 105 (N=210) was calculated using G*Power, version 3.1.9.7 (effect size 0.5; α=.05; power 0.95; allocation ratio 1).

Defining Reference Phenotypes

Only images of individuals reported to be clinically or molecularly diagnosed with a genetic syndrome were labeled as syndromic. When no syndrome was reported and no facial gestalt suggestive of a syndrome was observed, as judged by physicians working in medical genetics, images were labeled as "healthy."

Computer-Aided Facial Phenotyping

Computer-aided facial phenotyping was performed using DeepGestalt version 19.1.7, accessed via Face2Gene CLINIC (FDNA Inc). Neither the class labels nor diagnoses were passed to DeepGestalt. No other phenotypic information but 1 portrait photo per case was entered into the system. DeepGestalt's training set was tested not to contain duplicates of images used in this study, as described previously [42].

Danyel Cohort

The Danyel cohort, originally described by Danyel et al [30], comprises 116 healthy control images.

Syndromic Cohort

This cohort comprises frontal facial images of 17 syndromes. We planned to collect the same number of images for each of these syndromes. A total of 16 of these syndromes were chosen from the 201 distinct suggestions in DeepGestalt's top 30 results lists of the Danyel cohort. Syndromes of different frequencies ranging from 76% (frequently suggested) to 1% (rarely suggested) were selected. In descending order of frequency, these syndromes are as follows: Fragile X syndrome (OMIM: #300624), Angelman syndrome (OMIM: #105830), Rett syndrome (OMIM: #312750), Phelan-McDermid syndrome (OMIM: #606232), Klinefelter syndrome, Beckwith-Wiedemann syndrome (OMIM: #130650), 22q11.2 deletion syndrome (OMIM: #611867), Sotos syndrome (OMIM: #117550), Noonan syndrome (OMIM: PS163950), Loeys-Dietz syndrome (OMIM: PS609192), Williams-Beuren syndrome (OMIM: #194050), Rubinstein-Taybi syndrome (OMIM: PS180849), achondroplasia (OMIM: #100800), Wolf-Hirschhorn syndrome (OMIM: #194190), Pallister-Killian syndrome (OMIM: #601803), and Treacher Collins syndrome (OMIM: PS154500). In addition, we chose Apert syndrome (OMIM: #101200), which was not implied in the Danyel cohort.

Matched Control Cohort

Each photo of the syndromic cohort was matched to an image of an individual without a genetic syndrome by age, sex, and ethnicity to build a cohort of an equal number of control images.

Statistical Evaluation and Classification Experiments

Face2Gene CLINIC returns DeepGestalt's top 30 syndrome suggestions. DeepGestalt associates each suggestion with a Gestalt score [11]. The syndrome suggestions' frequencies, scores, and ranks were statistically evaluated.

Feature Extraction and Vector Construction

All images were labeled by class (syndromic vs healthy). Vectors were built to hold an attribute for any of the syndromes suggested at least once in DeepGestalt's top 30 suggestions. To construct a vector for a given photo, the 30 highest Gestalt scores were assigned to their respective attributes; and the remaining attributes were set to 0 (s. matrix.txt in Multimedia Appendix 1).

Classification

To differentiate between syndromic and healthy portrait photos, we trained linear support vector machines (SVMs) using the LinearSVM class of scikit-learn, version 0.21.3, with default parameters in Python 3.7. To avoid overfitting, training and testing were performed using a leave-1-out classification scheme. Since ethnic background is a possible confounder of DeepGestalt [15,22,26,29,33], we designed classification experiments based on all images, images of White persons, and those of persons with other ethnicities, to benchmark the influence of ethnicity on SVM performance.

To test a possible influence of the number of top ranks considered, classification of all images was run 30 times with the number of considered top Gestalt ranks, ranging from 1 to 30.

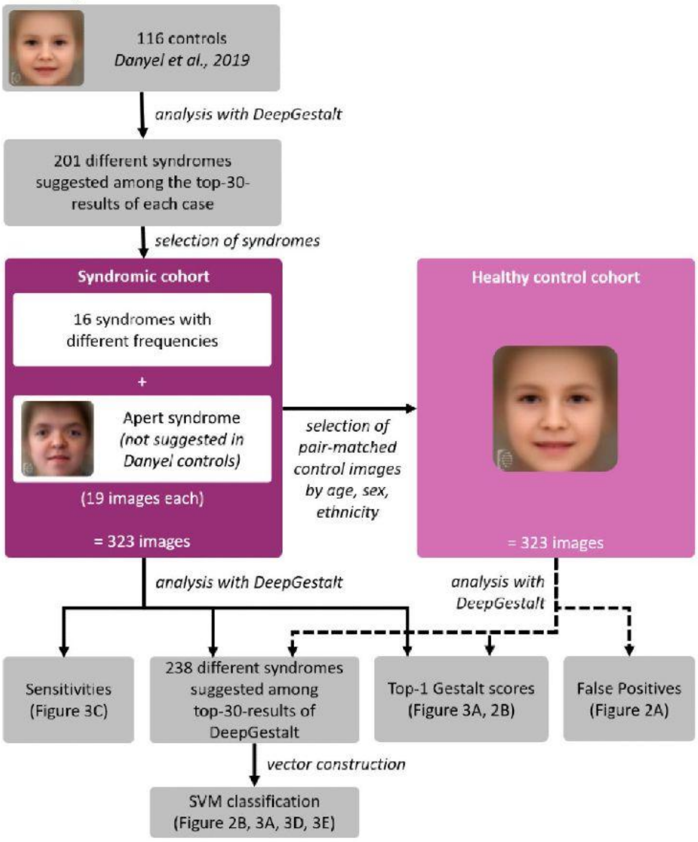
Statistical Analysis

Scores of the syndromic and healthy control cohort were tested to be different using a 2-sided, independent Welch *t* test. Difference of receiver operating characteristics (ROCs) was tested using a DeLong test. Classification performance was assessed using Matthews correlation coefficient (MCC). All statistical tests were performed in Python 3.7; the code can be found in Multimedia Appendix 1.

Data and Code Availability

The data and code can be found in Multimedia Appendix 1. For reasons of data protection, all data were cumulated (where possible), deidentified, and minimized. Facial images depicted in Figure 1 show computer-generated composite masks and not real individuals. In Multimedia Appendix 1, file data.txt describes the diagnosis, age, sex, and ethnicity of persons in the analyzed set of images; and file matrix.txt contains DeepGestalt's output vectors as used for this study. Files differentiator.py and reproduce.py may be used for reproducing the statistical results of this study. Further information may be found in file readme.txt (Multimedia Appendix 1).

Figure 1. Workflow of classification experiments.


116 controls
Danyel et al., 2019
analysis with DeepGestalt
201 different syndromes suggested among the top-30-results of each case
selection of syndromes
Syndromic cohort
16 syndromes with different frequencies
+
Apert syndrome (not suggested in Danyel controls)
(19 images each)
= 323 images
selection of pair-matched control images by age, sex, ethnicity
Healthy control cohort
= 323 images
analysis with DeepGestalt
analysis with DeepGestalt
Sensitivities (Figure 3C)
238 different syndromes suggested among top-30-results of DeepGestalt
Top-1 Gestalt scores (Figure 3A, 2B)
False Positives (Figure 2A)
vector construction
SVM classification (Figure 2B, 3A, 3D, 3E)

Results

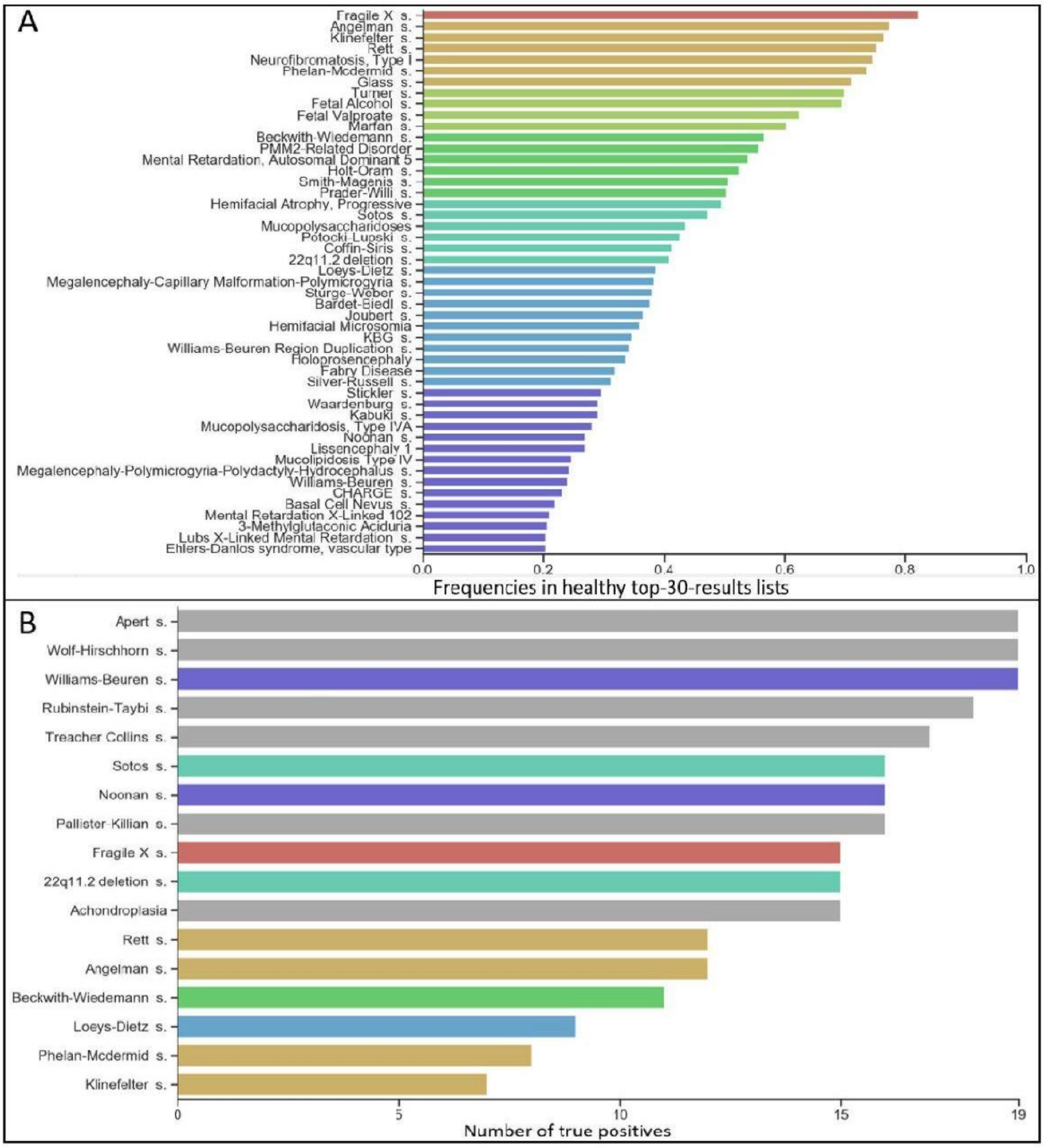
Included Images

We could include 19 images for each of the 17 syndromes in the syndromic cohort. A total of 83% (272/323) of these images were of White persons (file data.txt of Multimedia Appendix 1). Images from the syndromic cohort were matched to 323 images forming the matched control cohort, resulting in a total number of 646 analyzed photos (Figure 1).

Frequencies and Scores of Suggested Syndromes in Control Individuals

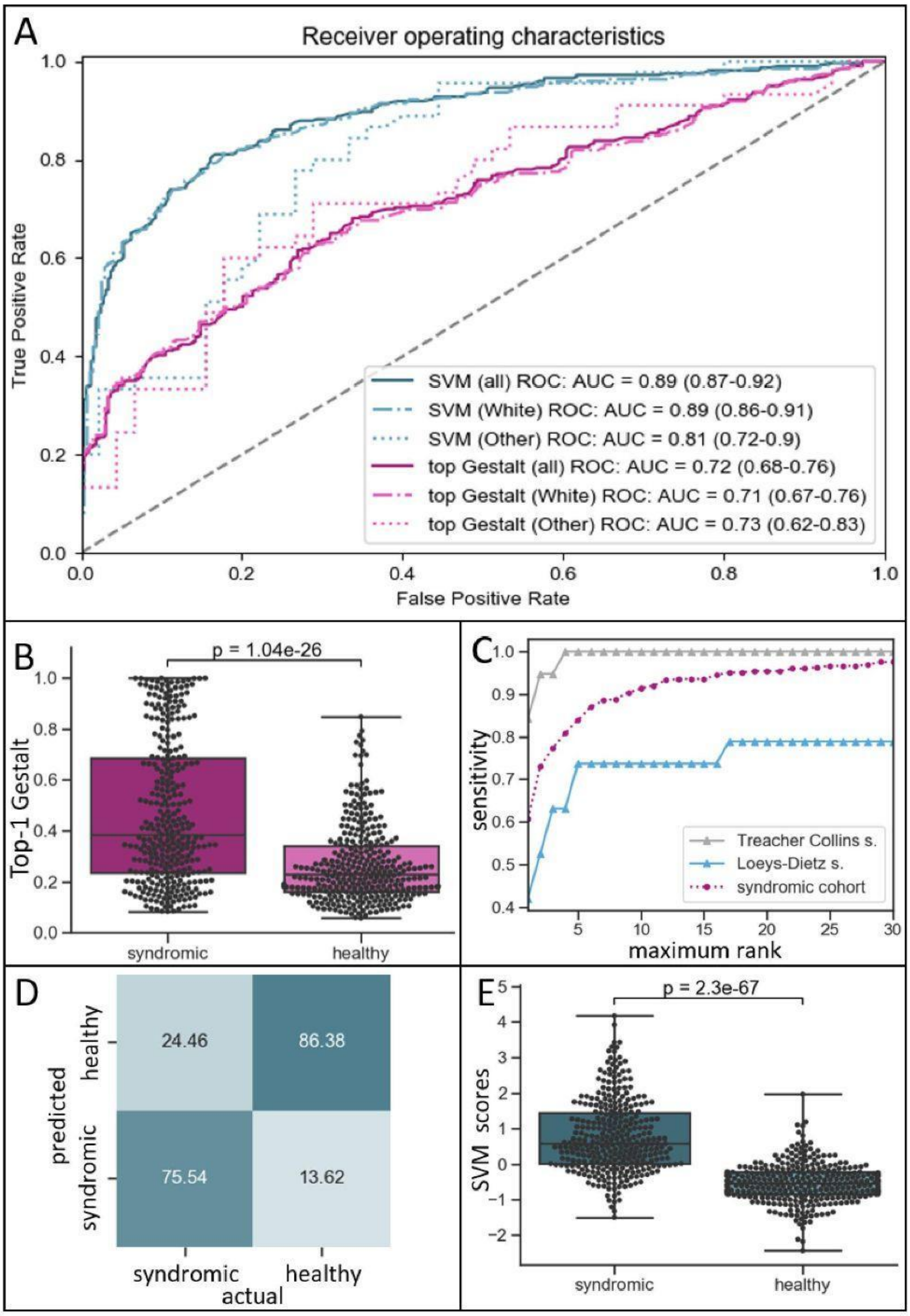
DeepGestalt suggested 238 different syndromes among the top 30 suggestions of the matched control cohort. One syndrome was suggested in more than 80% of the cases (Fragile X syndrome, 82%), 6 syndromes in 70%-80% of the cases; 4 syndromes in 60%-70% of the cases; 6 syndromes in 50%-60% of the cases; 6 syndromes in 40%-50% of the cases; 11 syndromes in 30%-40% of the cases; 15 syndromes in 20%-30% of the cases; 29 syndromes in 10%-20% of the cases; and 160 syndromes at least once in less than 10% of the cases (Figure 2A).

Figure 2. (A) Frequency of syndromes suggested by DeepGestalt in more than 20% of the matched control cohort's top-30-results lists. Colors indicate frequency percentages. (B) Number of images correctly classified as "syndromic"; colors relate to (A) and gray indicates <20%.



The highest first-rank Gestalt score of the matched control cohort amounted to 0.85, and the lowest, to 0.06, with a mean of 0.27 (SD 0.15). First-rank Gestalt scores of the syndromic cohort (highest 1.0; lowest 0.08; mean 0.47, SD 0.28) and the matched control cohort appeared to be separable with an AUROC of 0.72 (95% CI 0.68-0.76) (Figure 3A). Notably, this was found for both tested ethnic groups (Figure 3A, Multimedia Appendix 2), White persons only (AUROC 0.71, 95% CI 0.67-0.76; *P*<.001), and persons of other ethnicities only (AUROC 0.71, 95% CI 0.62-0.83; *P*<.001). Separability of the 2 cohorts is evident and significant (*P*<.001), as shown in Figure 3B.

Figure 3. (A) Receiver operating characteristic (ROC) curves: dashed line indicates random ROC curve; note that support vector machine (SVM) scores yield higher areas under the ROC curves (AUROCs) than their respective raw first-rank Gestalt scores. (B) Distribution of first-rank Gestalt scores in the syndromic cohort and the matched control cohort (healthy). (C) Sensitivities of DeepGestalt (X-axis: number of considered top ranks). Dark-purple circles: average of syndromic cohort; gray triangles: 19 images with Treacher-Collins syndrome; blue triangles: 19 images with Loeys-Dietz syndrome. (D) Distribution of SVM scores in the syndromic cohort and the matched control cohort; note: improved separability as compared to B. (E) SVM classification results based on the entire matched control cohort and syndromic cohort (threshold SVM score: 0).



Sensitivity of DeepGestalt

DeepGestalt's average top 10 sensitivity in the syndromic cohort amounted to 91%, varying between the 17 tested syndromes (Figure 3C, Multimedia Appendix 3). Interestingly, DeepGestalt was sensitive independent of ethnicity (White persons only, 90%; persons of other ethnicities only, 97%). A total of 7 syndromes reached a top 10 sensitivity of 100% (Fragile X, Noonan, Phelan-McDermid, Rett, Sotos, Treacher-Collins, and Williams-Beuren syndromes). DeepGestalt performed worst for Loeys-Dietz syndrome, with a top 10 sensitivity of 74% (Figure 3C).

Performance of the SVM

Sensitivities of binary SVM classification differed between syndromes (Figure 2B). All images of individuals with Apert syndrome, Wolf-Hirschhorn syndrome, and Williams-Beuren syndrome were correctly classified as being syndromic. The SVM performed worst on the 19 images of individuals with Klinefelter syndrome, correctly classifying only 7 of them as syndromic.

Binary SVM classification of DeepGestalt's output achieved an increased separability of syndromic images and healthy controls as compared to top Gestalt scores with an AUROC of 0.89 (95% CI 0.87-0.92) (Figure 3A). Again, this was true in both tested ethnic groups (Figure 3A), for photos of White persons (AUROC 0.88, 95% CI 0.86-0.91; *P*<.001) and those of persons of other ethnicities (AUROC 0.79, 95% CI 0.62-0.83). However, difference in ROCs was not significant in the latter (*P*=.13). SVM classification performance improved with an increasing number of considered ranks. Using the top 30 Gestalt scores showed the best MCC (0.63), as shown in Multimedia Appendix 4, with a sensitivity of 75.54% and a specificity of 86.38% (Figure 3D). Separability was significant (*P*<.001) (Figure 3E).

Discussion

Classification of Images of Individuals Without a Genetic Syndrome

To our knowledge, this is the first study to systematically analyze DeepGestalt's behavior on portrait photos of individuals without a genetic syndrome. For these images, we show that DeepGestalt's syndrome suggestions follow an interesting distribution. Certain syndromes are implied as differential diagnoses with a considerably high likelihood. Among these were Fragile X, Klinefelter, Rett, and Angelman syndromes, which were suggested in more than 3 quarters of the matched control cohort. In contrast, syndromes such as Treacher-Collins syndrome and Wolf-Hirschhorn syndrome were implied very rarely.

DeepGestalt cannot assign the class label "inconspicuous." Yet, DeepGestalt's scores are used to help judge the presence of a given syndrome. Based on a high maximum Gestalt score, a user could assume that the individual depicted in an entered image is likely to have a syndrome. Likewise, one is tempted to assume that a low maximum Gestalt score makes an underlying syndrome unlikely. Indeed, the mean of first-rank Gestalt scores is higher in images depicting syndromic facies than in images of individuals without a genetic syndrome. Similarly, scores higher than 0.85 appear to be specific indicators of a syndromic facies, and those lower than 0.08 are not suggestive of a genetic syndrome. However, these specific values are very rare. Gestalt scores alone are only fairly sufficient for judging the presence or absence of a genetic syndrome with facial dysmorphism since the distributions of the highest Gestalt scores of the syndromic and matched control cohort greatly overlap. We show that this problem can be reduced by considering both top Gestalt scores and the actual list of suggested syndrome matches. The boost in discriminatory power is illustrated by the increase of the respective AUROCs. Although DeepGestalt cannot directly assess the presence/absence of a syndromic facies, machine learning–based tools (eg, SVMs) built on top of DeepGestalt may be used for this purpose.

It is noteworthy that we achieved promising results with a comparably low number of samples and a low complexity classification model with default hyperparameters. We assume that the quality and complexity of future classifiers will improve as more data will become available. Increasing the number of top ranks considered for vector construction increased the performance of the SVM. However, the number of DeepGestalt's suggestions accessible via Face2Gene CLINIC is limited to 30 suggestions. We hypothesize that using more than just the 30 top ranks for vector construction might further boost classification performance. We classified DeepGestalt's output to predict the presence of a syndromic facies. We also suggest evaluating classification performance based on DeepGestalt's input vectors.

Potential Confounders

Until now, differences in the diagnostic performance of DeepGestalt, which arise due to the ethnicity of the person depicted, have been evaluated using DeepGestalt's sensitivity. Studies of earlier versions of DeepGestalt showed that its sensitivity is dependent on the ethnic background in certain syndromes [15,22]. Studies of more recent versions of DeepGestalt suggested that ethnicity had no major influence on its sensitivity [26,29]. In our set of syndromic images, DeepGestalt's sensitivity is remarkably high, which is in line with the previous studies highlighting DeepGestalt's good general sensitivity [11,36,42]. This high sensitivity of DeepGestalt was confirmed for both groups of images, those of White persons and those of persons of other ethnicities. Improvement of distinguishability of images of individuals with and without a genetic syndrome appeared to be stronger in the group of photos of White persons than in the group of photos of persons of other ethnicities. However, we assume that this is caused by the limited sample size of images of non-White persons in our data set. We believe that our approach is also applicable to populations comprising predominantly other ethnicities.

The SVM had difficulties classifying images of patients with syndromes that were frequently suggested in healthy controls. Possible explanations for DeepGestalt's output to be similar in controls and individuals with these syndromes could be as follows: (1) such syndromes have only mild characteristic facial features; (2) they have a typical facial gestalt, which is present only in some but not all affected individuals; or (3) they have no typical facies at all. For example, not all patients with Loeys-Dietz syndrome exhibit distinctive facial features [43], and the facial appearances of males with Klinefelter syndrome show no commonly observed characteristics [44].

Further Research

Further research is necessary to determine DeepGestalt's capacity to distinguish individuals with and without a genetic syndrome when combined with other sources of information, such as genetic test results and nonfacial phenotypic information. We suggest including additional scores that are based on both phenotype and genotype (eg, prioritization of exome data by image analysis [PEDIA] scores [42]) in future classifiers of the presence/absence of a syndromic facies.

The increasing use and quality of facial phenotyping software in clinical genetics should also be accompanied by an ethical evaluation of these systems [45]. This affects issues such as the automation of medical diagnostic action, the sharing of (potentially identifiable) data, and a potentially altered doctor-patient relationship. In particular, a systematic analysis of the patient perspective on the use of computer-aided facial analysis methodologies in clinical genetics is lacking so far.

We believe that our findings will help improve future versions of DeepGestalt and similar systems and are crucial when interpreting Face2Gene's results in the clinical routine. In particular, we recommend providing users with the false-positive rates of each suggested syndrome.

Conclusion

DeepGestalt is a computer-aided facial phenotyping tool that showed promising results for detecting a potentially syndromic facies. It yields higher first-rank scores in individuals with a genetic syndrome than in those without a diagnosis of a genetic syndrome. Its output may be classified to improve this detection. The exact stage to use DeepGestalt during the diagnostic makeup of individuals with a suspected genetic syndrome remains to be determined. Primarily, it should be used by expert geneticists.

Acknowledgments

We thank the members of the Elterninitiative Apertsyndrom und Verwandte Fehlbildungen eV, a parents' initiative on Apert syndrome and related disorders in Germany, for the contribution of their images, and Yaron Gurovich and Nicole Fleischer of FDNA Inc for technical assistance in checking DeepGestalt's training set for duplicate images used in this study. MAM is a participant in the BIH Charité Digital Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health. We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Funds of Charité – Universitätsmedizin Berlin.

Authors' Contributions

JTP, NH, and MAM designed the study. JTP, NH, MD, JE, ATAP, and MAM collected the data. SM, MS, DH, and CEO provided insights that were critical for the interpretation of data. MAM implemented the Python code with support from PH. PH and MAM performed the statistical analysis. JTP, NH, CEO, and MAM wrote the manuscript with approval of all the authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Code and data.
[ZIP File (Zip Archive), 137 KB-Multimedia Appendix 1]

Multimedia Appendix 2

(A) Distribution of first-rank Gestalt scores for the images of White persons in the syndromic cohort and the matched control cohort (healthy). (B) Distribution of first-rank Gestalt scores for the images of persons with other ethnicities in the syndromic cohort and the matched control cohort (healthy).
[PNG File , 97 KB-Multimedia Appendix 2]

Multimedia Appendix 3

DeepGestalt's sensitivities: purple circles indicate the average of the entire syndromic cohort; for other symbols/coloring, see respective subfigure title.
[PNG File , 208 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Performance of the SVM on the entire syndromic cohort and matched control cohort: X-axis number of top-rank Gestalt score used for vector construction per case. MCC: Matthews correlation coefficient. Note: rising tendency.
[PNG File , 46 KB-Multimedia Appendix 4]

References

1. Jackson M, Marks L, May GHW, Wilson JB. The genetic basis of disease. Essays Biochem 2018 Dec 03;62(5):643-723 [FREE Full text] [doi: 10.1042/EBC20170053] [Medline: 30509934]
2. Hart TC, Hart PS. Genetic studies of craniofacial anomalies: clinical implications and applications. Orthod Craniofac Res 2009 Aug;12(3):212-220 [FREE Full text] [doi: 10.1111/j.1601-6343.2009.01455.x] [Medline: 19627523]
3. Xie Q, Faust K, Van Ommeren R, Sheikh A, Djuric U, Diamandis P. Deep learning for image analysis: Personalizing medicine closer to the point of care. Crit Rev Clin Lab Sci 2019 Jan;56(1):61-73. [doi: 10.1080/10408363.2018.1536111] [Medline: 30628494]
4. Dias R, Torkamani A. Artificial intelligence in clinical and genomic diagnostics. Genome Med 2019 Nov 19;11(1):70 [FREE Full text] [doi: 10.1186/s13073-019-0689-8] [Medline: 31744524]
5. Boehringer S, Vollmar T, Tasse C, Wurtz RP, Gillessen-Kaesbach G, Horsthemke B, et al. Syndrome identification based on 2D analysis software. Eur J Hum Genet 2006 Oct;14(10):1082-1089 [FREE Full text] [doi: 10.1038/sj.ejhg.5201673] [Medline: 16773127]
6. Vollmar T, Maus B, Wurtz RP, Gillessen-Kaesbach G, Horsthemke B, Wieczorek D, et al. Impact of geometry and viewing angle on classification accuracy of 2D based analysis of dysmorphic faces. Eur J Med Genet 2008;51(1):44-53. [doi: 10.1016/j.ejmg.2007.10.002] [Medline: 18054308]
7. Boehringer S, Guenther M, Sinigerova S, Wurtz RP, Horsthemke B, Wieczorek D. Automated syndrome detection in a set of clinical facial photographs. Am J Med Genet A 2011 Sep;155A(9):2161-2169. [doi: 10.1002/ajmg.a.34157] [Medline: 21815261]
8. Ferry Q, Steinberg J, Webber C, FitzPatrick DR, Ponting CP, Zisserman A, et al. Diagnostically relevant facial gestalt information from ordinary photos. Elife 2014 Jun 24;3:e02020 [FREE Full text] [doi: 10.7554/eLife.02020] [Medline: 24963138]
9. Cerrolaza JJ, Porras AR, Mansoor A, Zhao Q, Summar M, Linguraru MG. Identification of dysmorphic syndromes using landmark-specific local texture descriptors Internet. 2016 Presented at: IEEE 13th International Symposium on Biomedical Imaging (ISBI); 13-16 April 2016; Prague, Czech Republic. [doi: 10.1109/isbi.2016.7493453]
10. Tu L, Porras A, Boyle A, Linguraru M. Analysis of 3D Facial Dysmorphology in Genetic Syndromes from Unconstrained 2D Photographs Internet. In: Frangi A, Schnabel J, Davatzikos C, Alberola-López C, Fichtinger G, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018. MICCAI 2018. Lecture Notes in Computer Science, vol 11070. Cham: Springer; 2018:347-355.
11. Gurovich Y, Hanani Y, Bar O, Nadav G, Fleischer N, Gelbman D, et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 2019 Jan;25(1):60-64. [doi: 10.1038/s41591-018-0279-0] [Medline: 30617323]
12. Dudding-Byth T, Baxter A, Holliday EG, Hackett A, O'Donnell S, White SM, et al. Computer face-matching technology using two-dimensional photographs accurately matches the facial gestalt of unrelated individuals with the same syndromic form of intellectual disability. BMC Biotechnol 2017 Dec 19;17(1):90 [FREE Full text] [doi: 10.1186/s12896-017-0410-1] [Medline: 29258477]
13. Basel-Vanagaite L, Wolf L, Orin M, Larizza L, Gervasini C, Krantz ID, et al. Recognition of the Cornelia de Lange syndrome phenotype with facial dysmorphology novel analysis. Clin Genet 2016 May;89(5):557-563. [doi: 10.1111/cge.12716] [Medline: 26663098]
14. Gripp KW, Baker L, Telegrafi A, Monaghan KG. The role of objective facial analysis using FDNA in making diagnoses following whole exome analysis. Report of two patients with mutations in the BAF complex genes. Am J Med Genet A 2016 Jul;170(7):1754-1762. [doi: 10.1002/ajmg.a.37672] [Medline: 27112773]
15. Lumaka A, Cosemans N, Lulebo Mampasi A, Mubungu G, Mvuama N, Lubala T, et al. Facial dysmorphism is influenced by ethnic background of the patient and of the evaluator. Clin Genet 2017 Aug;92(2):166-171. [doi: 10.1111/cge.12948] [Medline: 27925162]
16. Hadj-Rabia S, Schneider H, Navarro E, Klein O, Kirby N, Huttner K, et al. Automatic recognition of the XLHED phenotype from facial images. Am J Med Genet A 2017 Sep;173(9):2408-2414. [doi: 10.1002/ajmg.a.38343] [Medline: 28691769]
17. Gardner OK, Haynes K, Schweitzer D, Johns A, Magee WP, Urata MM, et al. Familial Recurrence of 3MC Syndrome in Consanguineous Families: A Clinical and Molecular Diagnostic Approach With Review of the Literature. Cleft Palate Craniofac J 2017 Nov;54(6):739-748. [doi: 10.1597/15-151] [Medline: 27356087]
18. Valentine M, Bihm DCJ, Wolf L, Hoyme HE, May PA, Buckley D, et al. Computer-Aided Recognition of Facial Attributes for Fetal Alcohol Spectrum Disorders. Pediatrics 2017 Dec;140(6):e20162028. [doi: 10.1542/peds.2016-2028] [Medline: 29187580]
19. Knaus A, Pantel JT, Pendziwiat M, Hajjir N, Zhao M, Hsieh T, et al. Characterization of glycosylphosphatidylinositol biosynthesis defects by clinical features, flow cytometry, and automated image analysis. Genome Med 2018 Jan 09;10(1):3 [FREE Full text] [doi: 10.1186/s13073-017-0510-5] [Medline: 29310717]
20. Liehr T, Acquarola N, Pyle K, St-Pierre S, Rinholm M, Bar O, et al. Next generation phenotyping in Emanuel and Pallister-Killian syndrome using computer-aided facial dysmorphology analysis of 2D photos. Clin Genet 2018 Feb;93(2):378-381. [doi: 10.1111/cge.13087] [Medline: 28661575]

21. Zarate YA, Smith-Hicks CL, Greene C, Abbott M, Siu VM, Calhoun ARUL, et al. Natural history and genotype-phenotype correlations in 72 individuals with SATB2-associated syndrome. Am J Med Genet A 2018 Apr;176(4):925-935. [doi: 10.1002/ajmg.a.38630] [Medline: 29436146]
22. Pantel JT, Zhao M, Mensah MA, Hajjir N, Hsieh T, Hanani Y, et al. Advances in computer-assisted syndrome recognition by the example of inborn errors of metabolism. J Inherit Metab Dis 2018 May;41(3):533-539 [FREE Full text] [doi: 10.1007/s10545-018-0174-3] [Medline: 29623569]
23. Ferreira CR, Altassan R, Marques-Da-Silva D, Francisco R, Jaeken J, Morava E. Recognizable phenotypes in CDG. J Inherit Metab Dis 2018 May;41(3):541-553 [FREE Full text] [doi: 10.1007/s10545-018-0156-5] [Medline: 29654385]
24. Jiang Y, Wangler MF, McGuire AL, Lupski JR, Posey JE, Khayat MM, et al. The phenotypic spectrum of Xia-Gibbs syndrome. Am J Med Genet A 2018 Jun;176(6):1315-1326 [FREE Full text] [doi: 10.1002/ajmg.a.38699] [Medline: 29696776]
25. Graul-Neumann LM, Mensah MA, Klopocki E, Uebe S, Ekici AB, Thiel CT, et al. Biallelic intragenic deletion in MASP1 in an adult female with 3MC syndrome. Eur J Med Genet 2018 Jul;61(7):363-368. [doi: 10.1016/j.ejmg.2018.01.016] [Medline: 29407414]
26. Vorravanpreecha N, Lertboonnum T, Rodjanadit R, Sriplienchan P, Rojnueangnit K. Studying Down syndrome recognition probabilities in Thai children with de-identified computer-aided facial analysis. Am J Med Genet A 2018 Sep;176(9):1935-1940. [doi: 10.1002/ajmg.a.40483] [Medline: 30070762]
27. Martinez-Monseny A, Cuadras D, Bolasell M, Muchart J, Arjona C, Borregan M, et al. From gestalt to gene: early predictive dysmorphic features of PMM2-CDG. J Med Genet 2019 Apr;56(4):236-245. [doi: 10.1136/jmedgenet-2018-105588] [Medline: 30464053]
28. Pascolini G, Fleischer N, Ferraris A, Majore S, Grammatico P. The facial dysmorphology analysis technology in intellectual disability syndromes related to defects in the histones modifiers. J Hum Genet 2019 Aug;64(8):721-728. [doi: 10.1038/s10038-019-0598-0] [Medline: 31086247]
29. Mishima H, Suzuki H, Doi M, Miyazaki M, Watanabe A, Matsumoto T, et al. Evaluation of Face2Gene using facial images of patients with congenital dysmorphic syndromes recruited in Japan. J Hum Genet 2019 Aug;64(8):789-794. [doi: 10.1038/s10038-019-0619-z] [Medline: 31138847]
30. Danyel M, Cheng Z, Jung C, Boschann F, Pantel JT, Hajjir N, et al. Differentiation of MISSLA and Fanconi anaemia by computer-aided image analysis and presentation of two novel MISSLA siblings. Eur J Hum Genet 2019 Dec;27(12):1827-1835. [doi: 10.1038/s41431-019-0469-3] [Medline: 31320746]
31. Pascolini G, Valiante M, Bottillo I, Laino L, Fleischer N, Ferraris A, et al. Striking phenotypic overlap between Nicolaides-Baraitser and Coffin-Siris syndromes in monozygotic twins with ARID1B intragenic deletion. Eur J Med Genet 2020 Mar;63(3):103739. [doi: 10.1016/j.ejmg.2019.103739] [Medline: 31421289]
32. Kruszka P, Hu T, Hong S, Signer R, Cogné B, Isidor B, et al. Phenotype delineation of ZNF462 related syndrome. Am J Med Genet A 2019 Oct;179(10):2075-2082 [FREE Full text] [doi: 10.1002/ajmg.a.61306] [Medline: 31361404]
33. Fung JLF, Rethanavelu K, Luk H, Ho MSP, Lo IFM, Chung BHY. Coffin-Lowry syndrome in Chinese. Am J Med Genet A 2019 Oct;179(10):2043-2048. [doi: 10.1002/ajmg.a.61323] [Medline: 31400053]
34. Weiss K, Lazar HP, Kurolap A, Martinez AF, Paperna T, Cohen L, et al. The CHD4-related syndrome: a comprehensive investigation of the clinical spectrum, genotype-phenotype correlations, and molecular basis. Genet Med 2020 Feb;22(2):389-397. [doi: 10.1038/s41436-019-0612-0] [Medline: 31388190]
35. Zarate YA, Bosanko KA, Gripp KW. Using facial analysis technology in a typical genetic clinic: experience from 30 individuals from a single institution. J Hum Genet 2019 Dec;64(12):1243-1245. [doi: 10.1038/s10038-019-0673-6] [Medline: 31551534]
36. Narayanan DL, Ranganath P, Aggarwal S, Dalal A, Phadke SR, Mandal K. Computer-aided Facial Analysis in Diagnosing Dysmorphic Syndromes in Indian Children. Indian Pediatr 2019 Dec 15;56(12):1017-1019 [FREE Full text] [Medline: 31884430]
37. Latorre-Pellicer A, Ascaso Á, Trujillano L, Gil-Salvador M, Arnedo M, Lucia-Campos C, et al. Evaluating Face2Gene as a Tool to Identify Cornelia de Lange Syndrome by Facial Phenotypes. Int J Mol Sci 2020 Feb 04;21(3):1042 [FREE Full text] [doi: 10.3390/ijms21031042] [Medline: 32033219]
38. Arora V, Puri RD, Bijarnia-Mahay S, Verma IC. Expanding the phenotypic and genotypic spectrum of Wiedemann-Steiner syndrome: First patient from India. Am J Med Genet A 2020 May;182(5):953-956. [doi: 10.1002/ajmg.a.61534] [Medline: 32128942]
39. Carli D, Giorgio E, Pantaleoni F, Bruselles A, Barresi S, Riberi E, et al. NBAS pathogenic variants: Defining the associated clinical and facial phenotype and genotype-phenotype correlations. Hum Mutat 2019 Jun;40(6):721-728. [doi: 10.1002/humu.23734] [Medline: 30825388]
40. Staufner C, Peters B, Wagner M, Alameer S, Barić I, Broué P, et al. Defining clinical subgroups and genotype-phenotype correlations in NBAS-associated disease across 110 patients. Genet Med 2020 Mar;22(3):610-621. [doi: 10.1038/s41436-019-0698-4] [Medline: 31761904]

41. Myers L, Anderlid B, Nordgren A, Lundin K, Kuja-Halkola R, Tammimies K, et al. Clinical versus automated assessments of morphological variants in twins with and without neurodevelopmental disorders. Am J Med Genet A 2020 May 12;182(5):1177-1189. [doi: 10.1002/ajmg.a.61545] [Medline: 32162839]
42. Hsieh T, Mensah MA, Pantel JT, Aguilar D, Bar O, Bayat A, et al. PEDIA: prioritization of exome data by image analysis. Genet Med 2019 Dec;21(12):2807-2814 [FREE Full text] [doi: 10.1038/s41436-019-0566-2] [Medline: 31164752]
43. MacCarrick G, Black JH, Bowdin S, El-Hamamsy I, Frischmeyer-Guerrerio PA, Guerrerio AL, et al. Loeys-Dietz syndrome: a primer for diagnosis and management. Genet Med 2014 Aug;16(8):576-587 [FREE Full text] [doi: 10.1038/gim.2014.11] [Medline: 24577266]
44. Bird RJ, Hurren BJ. Anatomical and clinical aspects of Klinefelter's syndrome. Clin Anat 2016 Jul;29(5):606-619. [doi: 10.1002/ca.22695] [Medline: 26823086]
45. Martinez-Martin N. What Are Important Ethical Implications of Using Facial Recognition Technology in Health Care? AMA J Ethics 2019 Mar 01;21(2):E180-E187 [FREE Full text] [doi: 10.1001/amajethics.2019.180] [Medline: 30794128]

Abbreviations

AUROC: area under the receiver operating characteristic
MCC: Matthews correlation coefficient
PEDIA: prioritization of exome data by image analysis
ROC: receiver operating characteristic
SVM: support vector machine

Edited by G Eysenbach; submitted 10.04.20; peer-reviewed by T Liehr, G Pascolini, M Pradhan, D Szinay; comments to author 12.06.20; revised version received 26.06.20; accepted 26.07.20; published 22.10.20

Please cite as:
Pantel JT, Hajjir N, Danyel M, Elsner J, Abad-Perez AT, Hansen P, Mundlos S, Spielmann M, Horn D, Ott CE, Mensah MA
Efficiency of Computer-Aided Facial Phenotyping (DeepGestalt) in Individuals With and Without a Genetic Syndrome: Diagnostic Accuracy Study
J Med Internet Res 2020;22(10):e19263
URL: http://www.jmir.org/2020/10/e19263/
doi: 10.2196/19263
PMID: 33090109

2.5 Genomsequenzierung in Familien mit angeborenen Gliedmaßenfehlbildungen

Elsner J*, **Mensah MA***, Holtgrewe M, Hertzberg J, Bigoni S, Busche A, Coutelier M, de Silva DC, Elçioglu N, Filges I, Gerkes E, Girisha KM, Graul-Neumann L, Jamsheer A, Krawitz P, Kurth I, Markus S, Megarbane A, Reis A, Reuter MS, Svoboda D, Teller C, Tuysuz B, Türkmen S, Wilson M, Woitschach R, Vater I, Caliebe A, Hülsemann W, Horn D, Mundlos S, Spielmann M.

Genome sequencing in families with congenital limb malformations.

Hum Genet. 2021 Aug;140(8):1229-1239. DOI: 10.1007/s00439-021-02295-y

*contributed equally

Neben einer verbesserten Phänotypisierung mittels NGP-Techniken könnte auch der umfassende Einsatz der neuen Sequenziertechnologien (d.h. die Testung des gesamten Genoms statt nur des Exoms) den diagnostischen Nutzen des NGS verbessern. Während eine Exomsequenzierug nur die ca. 1,5% des menschlichen Genoms, die für Proteine kodieren, untersucht, erfasst ein WGS das gesamte Erbgut. Ein wesentliches Problem bei der Bewertung nicht-proteinkodierender Varianten ist allerdings das Fehlen eines auf diese anwendbaren genetischen Codes: Man kann zwar die Aminosäurefolge eines Proteins aus der Basenfolge des zugehörigen Gens und so auch funktionelle Konsequenzen proteinkodierender Varianten vorhersagen, eine solche präzise Vorhersage der funktionellen Konsequenzen ist mit dem aktuellen Wissensstand für nicht-proteinkodierende Sequenzen allerdings nicht möglich. Die Genomsequenzierung verspricht über die Testung nicht-proteinkodierender Varianten hinaus auch die verbesserte Erfassung struktureller Varianten (Deletionen, Duplikationen, Translokationen, Inversionen, Repeatlängenveränderungen). Dies könnte eine der Hürden der klassischen genetischen Diagnostik, die Auswahl des richtigen Testverfahrens, lösen.

In dieser Studie untersuchten wir den zusätzlichen diagnostischen Nutzen des WGS an einer Kohorte von 69 Fällen mit angeborenen Extremitätenfehlbildungen, bei denen die genetische Routinediagnostik keine ursächliche Mutation identifizieren konnte. Dazu sequenzierten wir die Genome von 64 Trios, 1 Duo und 4 Einzelfällen. Für die Auswertung von kodierenden und strukturellen Varianten verwendeten wir die Analysesoftware VarFish, für die Filterung und Priorisierung von nicht-kodierenden Varianten kombinierten wir molekulare, populationsgenetische, phänotypische und murine Daten aus der wissenschaftlichen Literatur bzw. entsprechenden Datenbanken und definierten ein potentielles, mit der Entwicklung der Extremitäten assoziiertes Regulom.

In 12 der 69 Fälle konnten relevante Varianten identifiziert werden. Dazu zählten pathogene Einzelnukleotidvarianten in den bereits bekannten Krankheitsgenen *BHLHA9*, *FGFR1*, *FGFR2* und *GLI3* und in dem Kandidatengen *UBA2*, das wir als Krankheitsgen bestätigen konnten. Des Weiteren zählten dazu potentiell pathogene Varianten in den von uns neu identifizierten Kandidatengenen *ALDH1A2*, *HMGB1* und *SEMA3D:* die *de novo* missense Variante *SEMA3D*(NM_152754.3):c.1918G>A p.(Asp640Asn) bei einem Patienten mit Kleinwuchs und fehlenden Endgliedern des 5. Strahls, die mit einer Syndaktylie segregierende Frameshift-Mutation *ALDH1A2*(NM_003888.4):c.35delT p.(Val12Glyfs*31) und die *de novo* Frameshift-Mutation *HMGB1*(NM_002128.7):c.551_554delAGAA p.(Lys184Argfs*44) bei einem Feten mit komplexer Extremitätenfehlbildung. Außerdem umfasste dies eine pathogene Repeatexpansion in *HOXD13* sowie zwei komplexe Varianten (*SHH*-Locus und SHFM3-Locus).

Das von uns definierte Extremitätenregulom umfasste 0,24% des Genoms. Nicht-proteinkodierende Einzelnukleotidvarianten, die wir als pathogen einschätzen, konnten allerdings nicht identifiziert werden.

Die Ergebnisse zeigen, dass die Genomsequenzierung das Potenzial hat, als ein umfassender Test für die Testung verschiedener Mutationsarten zu dienen, aber auch, dass die klinische Interpretation von nicht-proteinkodierenden Sequenzvarianten noch nicht möglich ist.

Human Genetics (2021) 140:1229–1239
https://doi.org/10.1007/s00439-021-02295-y

ORIGINAL INVESTIGATION

Genome sequencing in families with congenital limb malformations

Jonas Elsner[1] · Martin A. Mensah[1,2] · Manuel Holtgrewe[3] · Jakob Hertzberg[4] · Stefania Bigoni[5] · Andreas Busche[6] · Marie Coutelier[1,7] · Deepthi C. de Silva[8] · Nursel Elçioglu[9,10] · Isabel Filges[11] · Erica Gerkes[12] · Katta M. Girisha[13] · Luitgard Graul-Neumann[1] · Aleksander Jamsheer[14] · Peter Krawitz[15] · Ingo Kurth[16] · Susanne Markus[17] · Andre Megarbane[18] · André Reis[19] · Miriam S. Reuter[19] · Daniel Svoboda[20] · Christopher Teller[21] · Beyhan Tuysuz[22] · Seval Türkmen[1,23] · Meredith Wilson[24] · Rixa Woitschach[25] · Inga Vater[26] · Almuth Caliebe[26] · Wiebke Hülsemann[27] · Denise Horn[1] · Stefan Mundlos[1,4] · Malte Spielmann[4,26,28]

Received: 18 March 2021 / Accepted: 10 May 2021 / Published online: 22 June 2021

Abstract

The extensive clinical and genetic heterogeneity of congenital limb malformation calls for comprehensive genome-wide analysis of genetic variation. Genome sequencing (GS) has the potential to identify all genetic variants. Here we aim to determine the diagnostic potential of GS as a comprehensive one-test-for-all strategy in a cohort of undiagnosed patients with congenital limb malformations. We collected 69 cases (64 trios, 1 duo, 5 singletons) with congenital limb malformations with no molecular diagnosis after standard clinical genetic testing and performed genome sequencing. We also developed a framework to identify potential noncoding pathogenic variants. We identified likely pathogenic/disease-associated variants in 12 cases (17.4%) including four in known disease genes, and one repeat expansion in *HOXD13*. In three unrelated cases with ectrodactyly, we identified likely pathogenic variants in *UBA2*, establishing it as a novel disease gene. In addition, we found two complex structural variants (3%). We also identified likely causative variants in three novel high confidence candidate genes. We were not able to identify any noncoding variants. GS is a powerful strategy to identify all types of genomic variants associated with congenital limb malformation, including repeat expansions and complex structural variants missed by standard diagnostic approaches. In this cohort, no causative noncoding SNVs could be identified.

Introduction

The repertoire of diagnostic tests in human genetics is as diverse as the types of genetic alterations they were developed to detect (Berisha et al. 2020). Through the development of Next Generation Sequencing technologies (NGS) sequencing has become several orders of magnitude faster and cheaper. This has led to an enormous increase in the efficiency of genetic testing (Levy and Myers 2016). NGS quickly found its way from research applications to the clinic: today, panel and exome sequencing are elements of the routine diagnostics in genetic medicine (Deciphering Developmental Disorders Study 2017). Despite these significant advances, classical genetic testing methods such as chromosomal microarray analysis (CMA) and Sanger sequencing remain part of the standard diagnostic arsenal. This is because NGS-based gene panels often do not detect structural variants such as inversions and translocations, or fail to determine repeat lengths (Berisha et al. 2020). The goal of detecting all types of genetic variation in a single test can theoretically be achieved by short-read based genome sequencing (GS) (Xue et al. 2015). While there are some very encouraging proof of concept studies for the use of GS in individuals with intellectual disability (Lindstrand et al. 2019), GS is not yet part of the clinical routine and there is a lack of systematic studies on the benefits of such tests for individuals with congenital malformations.

A major limitation of panel and exome sequencing approaches is that they usually do not cover 98% of the genome which is noncoding, and are, hence, unable to detect deep intronic splice variants or intergenic regulatory

Jonas Elsner and Martin A. Mensah shared first authorship.

Stefan Mundlos and Malte Spielmann shared senior authorship.

✉ Stefan Mundlos
stefan.mundlos@charite.de

✉ Malte Spielmann
malte.spielmann@uksh.de

Extended author information available on the last page of the article

variants. Therefore, over 40% of individuals with genetic diseases receive no molecular diagnosis after standard testing (Gilissen et al. 2014). This is likely because the noncoding sequence has largely been ignored despite most nucleotides and single nucleotide variants being noncoding. The two main challenges that currently hamper the medical interpretation of noncoding variants are the poor understanding of the "regulatory code" of the noncoding genome and the large number of noncoding variants in each individual that renders classical functional work-up strategies impossible.

In this study, we aimed to determine the diagnostic potential of GS as a comprehensive one-test-for-all strategy in a cohort of 69 unsolved patients with congenital limb malformations. We also attempted to develop a framework to prioritize the large number of noncoding variants identified in the GS studies by combining mouse genetic and human functional epigenetic data with in vivo-validated enhancer sequences.

Materials and methods

Study design

Patients affected with malformations of two limbs, or two individuals from a family, each affected with a malformation of at least one limb were recruited (Supplementary Fig. 1). Exclusion criteria included a molecularly established genetic diagnosis, a suspected diagnosis of amniotic band syndrome, or an isolated fifth finger clinodactyly. A convenience set of samples was collected from the patients of the Department of Hand Surgery of the Katholisches Kinderkrankenhaus Wilhelmstift Hamburg and the Institute of Medical Genetics and Human Genetics of the Charité (IMG)—Universitätsmedizin Berlin. This sample-set was compiled with cases that were sent to the IMG by external physicians for diagnostic purposes. The sample-set was fixed before conducting GS.

Included patients

We included 69 patients in this study (Supplementary Table 1). We sequenced the index case and both parents in 64 cases, the index and one parent in one case, and only the index in four cases (parental DNA was not available for testing). In one case, we additionally sequenced a sibling. In five cases, one parent showed a limb malformation comparable to the index. In one case featuring ectrodactyly and apparently unaffected parents and grandparents, a maternal grand-uncle was affected, who was also sequenced. In 60 cases no family member other than the index was reported to show a limb malformation.

Phenotyping and conventional genetic testing

Limb malformations were phenotyped based on photographs and radiographs by a panel of medical professionals including expert clinical geneticists. Phenotypes were described as per Human Phenotype Ontology (HPO) terminology.

Based on a patient's phenotype, genes were selected for sequencing by medical geneticists. Sample preparation and Sanger-sequencing were performed using standard procedures. High resolution (1 M oligo) CMA was performed as described previously (Flöttmann et al. 2018).

Genome sequencing and variant calling

Paired-end PCR-free GS was performed by Macrogen Inc. (South-Korea) using a HiSeq X Ten platform. DNA preparation, sequencing, and sequence data processing were performed according to Macrogen's standard protocol (coverage 30× and read length 150 bp).

The FASTQ files were transferred to the Core Unit Bioinformatics of the Berlin Institute of Health (CUBI) for variant calling. Files were further processed and securely stored in the System for Omics Data Analysis and Retrieval (SODAR) (Nieminen et al. 2020). GATK HC was used to call simple nucleotide variants, while structural variants were called using Delly2, PopDel, and ERDS/SV2. Afterwards, variants were processed and annotated by the VarFish platform (Holtgrewe et al. 2020). Variants were mapped according to the hg19 reference genome.

Variant filtration

Coding SNVs and SVs

Each index case was filtered as a singleton, regardless of the availability of family data. If parental samples were available, a trio-based filtration approach was additionally performed. Male-index trio-cases were also filtered for hemizygous X-chromosomal variants.

Simple nucleotide variant filtration was performed on the VarFish platform (Version v0.17.2) (Holtgrewe et al. 2020). We filtered GS data for non-synonymous exonic and splice variants using default settings for read depth, allelic balance, and read quality. Allele counts were set as described in Supplementary Table 2. SNVs that passed filtration were exported as variant calling files (VCF). For evaluation of variant pathogenicity, VCFs were uploaded to MutationDistiller and Exomiser. The first ten results were exported for semi-automated, in-depth analysis (see Supplementary Fig. 2 for details). We also tested for truncating or probable

LoF (CADD > 20) variants affecting the same gene with a pLI > 0.9 in at least two independent patients.

Structural variant filtration was performed as described in Supplementary Table 3. The minimal size for structural variant filtration was 1500 bp. Whenever we obtained more than 30 structural variants after initial filtering, we increased the number of minimally covered informative reads from 2 to 5. Each SV passing filtration was judged manually with the information provided by the IGV-Browser and UCSC (see Supplementary Fig. 3 for details).

All findings were evaluated at weekly clinical meetings.

Analysis of noncoding variants: limb regulome

We defined a limb-specific potential regulome to filter and interpret non-coding variants. For this purpose, we created a list of 1719 genes involved in embryonic limb development based on data from the Mouse Genome Informatics (MGI) database and entries in OMIM. We defined the human limb regulome as the following: 1. all conserved (phyloP > 1,3) variants, 2. those located within the same topologically associating domain (TAD) (as determined in human fibroblasts (Dixon et al. 2012)) as a limb gene, 3. those that were marked by an H3K27 acetylation peak in human limb buds (Cotney et al. 2013). We also included the validated enhancer elements of the VISTA Enhancer Browser. Coding and noncoding SNVs were filtered for rare variants (frequency < 0.1%) and were considered to be potentially affecting the same regulatory element if they were either less than 300 bp apart or positioned in the same established enhancer.

Whenever we identified rare, potentially pathogenic heterozygous coding variants of genes associated with a recessive limb phenotype in individuals featuring at least a partial overlap with that phenotype, we also screened for *in trans* conserved non-coding variants with a MAF < 3% affecting the same TAD.

Results

We collected a cohort of 69 individuals affected with congenital limb malformations. All individuals had previously gone through our clinical genetics routine pipeline including clinical examination, candidate gene testing, and CMA. We then performed GS as a comprehensive one-test-for-all strategy.

In total, we identified 333,163,643 single nucleotide variants (SNVs) among the 69 sequenced index patients, of which 7,020,766 were either coding or flanking coding elements by 10 bp or less. 326,142,877 were noncoding SNVs, of which 19,362 were rare (gnomAD AF < 0.01). 21,369 of the coding SNV calls were classified by Jannovar to be of at least moderate relevance (missense and truncating), and 1429 to be of high relevance (truncating only). VarFish filtering returned 5761 SNVs. Filtering by Exomizer and MutationTaster identified 433 potentially pathogenic coding calls among these, of which 174 were high-quality calls suitable for further evaluation.

30,062, SNVs resulted in the potential loss-of-function of probably haploinsufficient genes. 49 of these affected the same gene in two unrelated index individuals (Supplementary Fig. 2).

We also analyzed the structural variants in 68 of the 69 index patients. 55 cases were filtered as trios with unaffected parents and moderate filter settings. Five were filtered as trios with another affected relative and moderate filter settings. Stricter filter settings were chosen for 9 trios because moderate filter settings produced an unmanageable amount of SV calls. 3 cases were analyzed as singletons. Individuals I1, I2, I3 did not yield any results, I4 was excluded from the SV analysis because too many SVs were called even with stricter filter settings due to poor data quality.

Of the 1,555,426 SVs, 633 SVs passed the filtering by VarFish, of which 222 were inversions, 288 deletions, 76 duplications, and 47 breakpoints of potential translocations.

417 of these SVs were excluded because they were of poor calling quality or because they were inherited from an unaffected parent. We then manually inspected the remaining 216 SVs. Segregation analysis in the parents was performed by qPCR after comparing candidate CNVs with known limb genes according to the Human Phenotype Ontology, cross-species phenotype comparison, mouse models, gene expression data (Cao et al. 2019), limb enhancer elements (Visel et al. 2009), and the local topological associating domain (TAD) architecture of the locus (Dixon et al. 2012; Cao et al. 2019). As a result, we identified 30 promising variants (Supplementary Fig. 3).

Variants in four known, limb malformation associated genes

We identified pathogenic variants in established disease genes in four individuals (I5, I6, I7, I8, Supplementary Fig. 4), which we confirmed by Sanger sequencing. These included a missense variant in *FGFR1,* already described in the literature (Muenke et al. 2014), and three previously undescribed variants in the genes *FGFR2*, *GLI3*, and *BHLHA9*. In all four cases, we classified the variants as (likely) pathogenic according to the criteria of the American College of Medical Genetics and Genomics (ACMG), based on the type of variant and the phenotype of the patient. All variants were inherited (note that the mother of I8, was not radiographically phenotyped, which is necessary to diagnose mild *FGFR2*-associated phenotypes (Flöttmann et al. 2015)).

Repeat expansions of *HOXD13*

In individual I9, featuring brachy-poly-syndactyly, we detected a repeat expansion by eight alanines on the *HOXD13* allele inherited from his affected mother, already described as pathogenic in the literature (Brison et al. 2014), and a polymorphic repeat expansion by only one alanine on the paternal *HOXD13* allele (Supplementary Fig. 5). These findings were confirmed by conventional *HOXD13* microsatellite analysis.

Structural variants at known disease loci

In individual I10 with bilateral upper and lower limb ectrodactyly, we identified an inversion of 105 kb (chr10: 103,321,526–103,426,609) flanked by two deletions (chr10:103,319,219–103,321,525 and chr10:103,426,610–103,436,718) at the split-hand foot malformation locus 3 (SHFM3) on chr10q24 inherited from his unaffected mother (Fig. 1, Supplementary Fig. 6). His affected great-uncle also carried the inversion. The variant overlaps with the most common duplications associated with ectrodactyly (de Mollerat et al. 2003; Klopocki et al. 2012). The minimal overlapping region of pathogenic SHFM3 duplications includes *BTRC*, *POLL,* and *DPCD* (Holder-Espinasse et al. 2019). The inversion described here is copy number neutral, suggesting that positional effects rather than gene dosage might be responsible for the phenotype. It includes a topologically associating domain boundary (Holder-Espinasse et al. 2019) and is likely to change the enhancer landscape at the SHFM3 locus leading to *FGF8* misregulation causing ectrodactyly.

In individual I11, featuring bilateral mirror-image polydactyly of the hands and feet (Fig. 2a), CMA had detected a 300 kb amplification on chr7q36.1. We initially classified

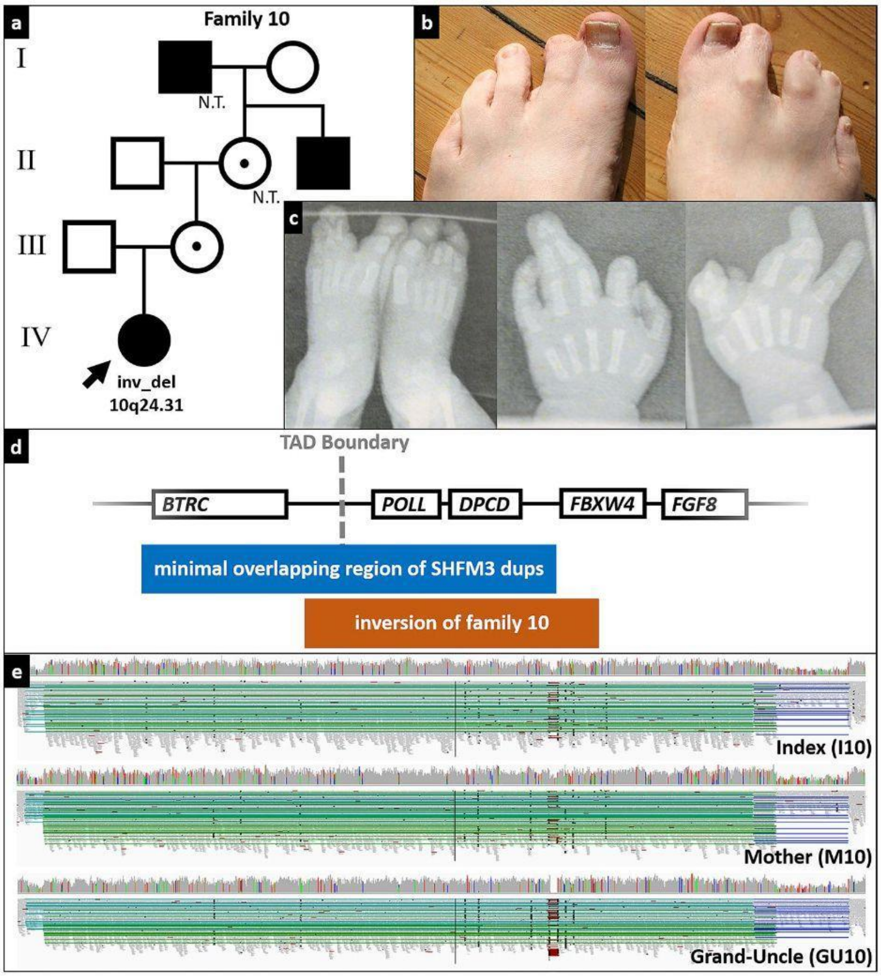


Fig. 1 Inversion-deletion at SHFM3 locus: **a** pedigree, N.T. not tested. **b** feet of grand-uncle (II-3). **c** hands and feet of the index patient (IV-1). **d** genomic architecture of SHFM3. **e** GS data of the family, note the presence of an inversion (chr10: 103,321,526–103,426,609) flanked by deletions (chr10:103,319,219–103,321,525 and chr10:103,426,610–103,436,718) on either site

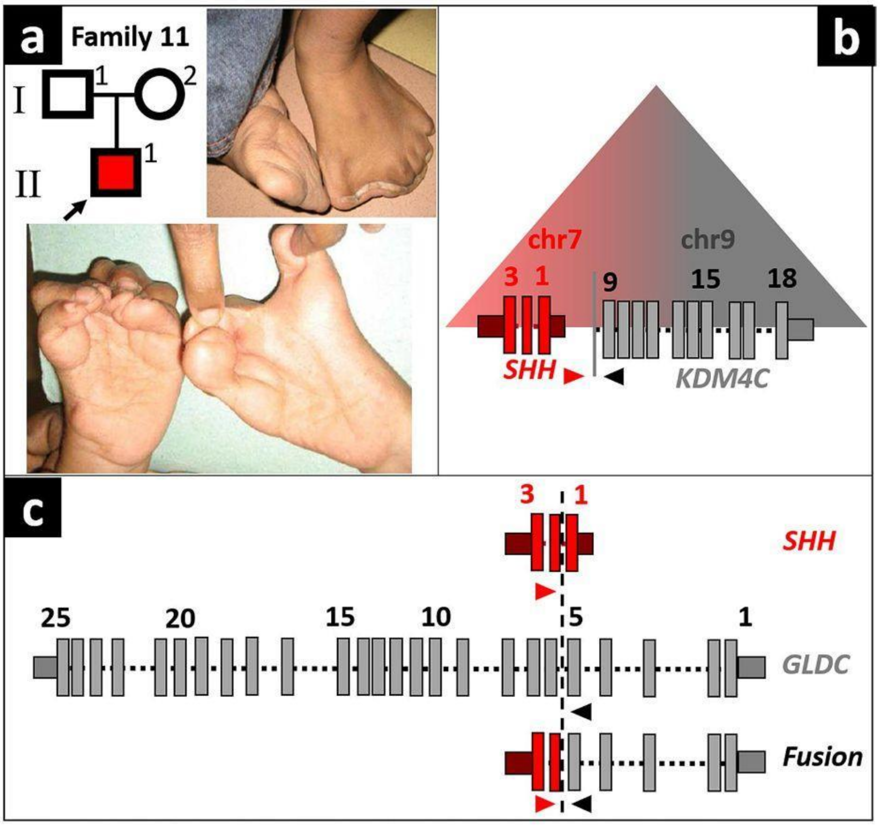


Fig. 2 **a** Pedigree and phenotype of individual I11. **b** Potential neo-TAD at the fusion site. **c** Breakpoint and fusion sites between regions from chr7 and chr9

the variant as a variant of unknown significance, because the individual's phenotype did not match that of an individual with muscular hypertrophy, reported to have a similarly sized and positioned duplication (Kroeldrup et al. 2012).

The amplification was identified again using GS. However, sequencing revealed that it was part of a complex structural variant containing two overlapping duplications (dup1 and dup2) at chr7q36.1 (Supplementary Fig. 7). The smaller dup2 shares the central breakpoint with dup1. The distal breakpoint of dup2 is positioned within dup1 in intron 1 of *SHH*, at chr7:155,603,964. GS data showed that the duplicons were not positioned *in tandem*, but are both fused to sequences originating from chr9p24.1. The distal breakpoint of dup2 was fused to intron 5 of *GLDC* and the distal breakpoint of dup1 to intron 8 of *KDM4C* (Fig. 2c; Supplementary Fig. 7). Analysis of the parents by Sanger sequencing showed that the structural variant occurred de novo in individual I11.

The mirror-image polydactyly of individual I11 shows striking phenotypic overlap with Laurin–Sandrow syndrome, which is caused by duplications of the *SHH* regulator ZRS, positioned in intron 5 of *LMBR1* on chr7p36.3, resulting in ectopic expression of *SHH* in the embryonic limb (Lohan et al. 2014). Both duplications do not include the ZRS and duplications of *SHH* itself have not been described to cause Laurin–Sandrow syndrome. However, a duplicated fragment containing *SHH* that is inserted into another domain, as observed in the de novo SV of I11, makes an ectopic expression of *SHH* in the embryonic limb very likely. We assume that the formation of an *SHH-KDM4C* neo-TAD, resulting in the misregulation of *SHH* by *KDM4C*-enhancers in the limb mesenchyme (note the known expression of KDM4C in embryonic vertebrate limb buds) is the most likely explanation for such an ectopic SHH-expression (Fig. 2b).

Therefore, we re-classified the complex SV involving *SHH* in individual I11 with bilateral mirror-image polydactyly as causative.

Establishing *UBA2* as a novel disease gene

We also identified variants in new candidate genes. Two unrelated individuals with isolated split hand malformation featured different heterozygous frameshift variants in the *ubiquitin-like modifier-activating enzyme 2* (*UBA2*) (Fig. 3a, b). Individual I12 harboured the de novo variant NM_005499.3(*UBA2*):c.1355_1356delTG;p.(Val452Alafs*6).

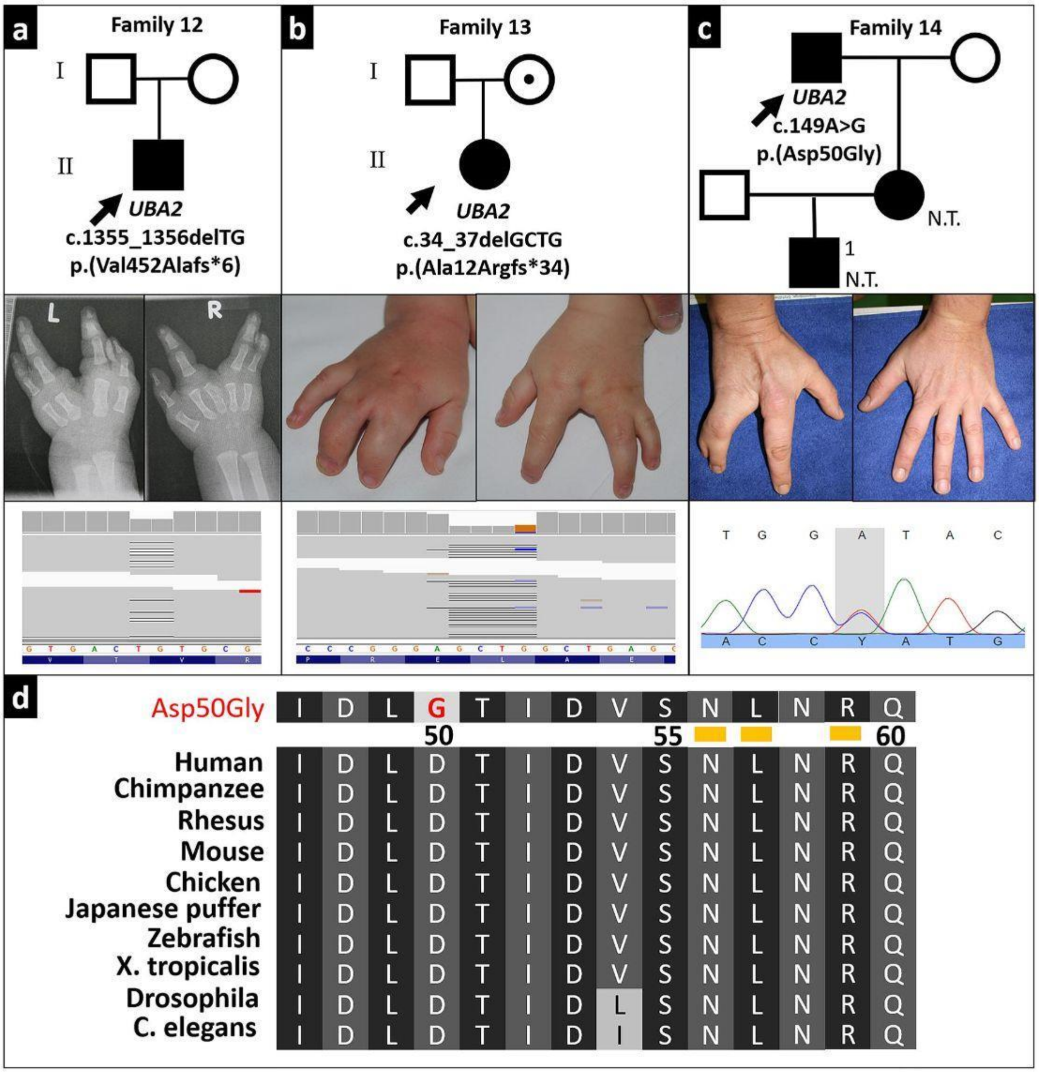


Fig. 3 *UBA2* variants and ectrodactyly. **a–c** Patients with likely pathogenic UBA2 variants *upper panels*: pedigrees, N.T.: not tested; *middle panels*: characteristic limb malformations, *lower panels*: sequencing data. **d** conservation of Asp50 mutated in individual I14, numbers indicate amino acid residues, *yellow bars* highlight positions tested by Olsen et al. to cause loss of function when substituted by alanine (Olsen et al. 2010)

Individual I13 inherited the variant NM_005499.3(*UBA 2*):c.34_37delGCTG;p.(Ala12Argfs*34) from his apparently unaffected mother (no radiographs of her hands were available).

We classified the pathogenicity of these variants according to the ACMG guidelines. Both are null variants of *UBA2* which has a pLI-score of 1 (PVS1). c.1355_1356delTG occurred de novo in an individual with a negative family history (PS2). The variants are absent from the 1000 Genomes Project and the Exome Aggregation Consortium databases (PM2) and were predicted to be pathogenic by MutationTaster (PP3). *UBA2* variants have recently been described in individuals with ectrodactyly (Chowdhury et al. 2014; Abe et al. 2018; Yamoto et al. 2019; Aerden et al. 2020). Hence, we regarded UBA2 as a disease-associated gene and these variants as pathogenic (1PVS(+1PS)+1PM+1PP).

These findings prompted subsequent Sanger sequencing of *UBA2* in 24 unrelated families with ectrodactyly, who have been tested negative for variants in the established SHFM loci/genes. In one individual (I14) with unilateral split-hand malformation (Fig. 3c), we identified the missense variant NM_005499.3(*UBA2*):c.149A > G;p (Asp50Gly). The daughter of I14 and her son also feature ectrodactyly (PP4), but were unavailable for testing. Asp50 is part of a consecutive 15 amino acid sequence (Ile47 to Phe61) shared amongst all nephrozoan *UBA2* orthologues (Fig. 3d). Olsen et al. showed that variants of residues (Asn56Ala, Leu57Ala, Arg59Ala) of this element result in loss of UBA2 function, and found that the very residue mutated in individual I14, Asp50, forms hydrogen bonds with Asn177 and Thr178 essential for proper UBA2 folding and thus its function (PS3) (Olsen et al. 2010). The variant was also absent in the databases (PM2) and predicted to be pathogenic by MutationTaster (PP3). Hence, we classified these variants as likely pathogenic according to the ACMG's guidelines (1PS + 1PM + 2PP).

Novel candidate genes

Our analysis also revealed several novel, high-confidence candidate genes associated with limb defects. In individual I15 featuring severe mirror image foot polydactyly, we found a de novo frameshift variant in the gene encoding the high mobility group box 1 protein (HMGB1) (Supplementary Fig. 8). NM_002128.7(*HMGB1*):c.551_554delAGAA;p.(Lys184Argfs*44) leads to the replacement of the protein's entire C-terminal 30-residue acidic tail by 41 other unrelated residues. The tail is normally formed by an Asp/Glu-repeat element, which is highly conserved among *HMGB1* orthologues. This repeat element stabilizes HMGB1's secondary structure and is crucial for its DNA-bending capacity (Belgrano et al. 2013; Anggayasti et al. 2020). The variant is not only absent from the databases but also no variant listed in gnomAD contains an amino acid residue except Glu or Asp in the acidic tail domain. *HMGB1* has a pLI score of 1. In mouse and zebrafish studies, HMGB1 has been shown to regulate digit number during embryonic limb development by interacting with WNT, BMP, and SHH (Itou et al. 2011). We, therefore, consider *HMGB1* to be a novel candidate gene for mirror image foot polydactyly.

Individual I16, who featured short stature, absent distal phalanges of the 5th fingers and toes, and dysplastic middle phalanges of the toes carried a de novo missense variant in the gene encoding semaphorin 3D (*SEMA3D*) (Supplementary Fig. 9). NM_152754.3(*SEMA3D*):c.1918G > A;p.(Asp640Asn) is absent from the 1000 Genomes Project database and is listed only 4 times in the Genome Aggregation Database (gnomAD). The Asp640 residue in the immunoglobulin-like domain of *SEMA3D* is highly conserved amongst vertebrates. The variant is predicted to be pathogenic by MutationTaster. SEMA3D regulates neural crest cell differentiation and is involved in the organogenesis of the heart (Sanchez-Castro et al. 2015), parathyroid gland (Singh et al. 2019), and, notably, limbs (Govindan et al. 2016). We, therefore, consider it a candidate gene for short stature with limb abnormalities.

In individual I17 we identified a paternally inherited frameshift variant in the *aldehyde dehydrogenase 1 family member A2* gene (*ALDH1A2*) encoding retinaldehyde dehydrogenase 2 (Supplementary Fig. 10). Both, the patient and her father feature isolated cutaneous syndactyly of the fingers III–IV and the toes II–III. The variant NM_003888.4(*ALDH1A2*):c.35delT;p.(Val12Glyfs*31) is absent from the databases and is predicted to be disease-causing by MutationTaster. ALDH1A2 is a direct target of HOXA13 and plays a key role in vertebrate digit development by regulating, in particular, interdigital programmed cell death (Shou et al. 2013). Rescued *ALDH1A2* knockout mice show reduced interdigital cell death and thus impaired digit separation during limb development resulting in syndactyly (Zhao et al. 2010). It is, therefore, likely that *ALDH1A2*:c.35delT caused the phenotype of syndactyly in individual I17 and her father.

Identification of noncoding variants

So far, the interpretation of disease-related variation has been focused on protein-coding DNA and the identification of variants that directly result in the disruption of specific gene functions.

Here, we aimed to develop a framework to prioritize a large number of noncoding variants from GS studies, by combining mouse genetic and human functional epigenetic data with in vivo-validated enhancer sequences. We then defined a limb-specific regulome that we used to filter all noncoding variants (Materials and Methods). Our potentially disease-relevant limb-specific regulome consists of 5,591,007 sites covering in total 7,294,220 bp, i.e. 0.24% of the human genome (hg19).

Overall, we identified 19,362 rare noncoding SNVs in the 69 index patients, of which 143 were located within the limb regulome (Fig. 4). First, we focused on the de novo variants and identified 6 calls located in potential regulatory elements. Two variants were excluded because they were called in cases with (likely) pathogenic coding or structural variants.

Individual I18 presenting with bilateral syndactyly of fingers II-V featured the de novo call chr1:41948304AAG > A in intron 2 of *EDN2*. The position shows increased acetylation of H3K27 in human limb buds. The 2 bp deletion also removes one element of a 6-AG-repeat whose length is not conserved in vertebrates. Furthermore, *EDN2* encodes endothelin 2, a potent vasoconstrictor with no evident link to limb development.

Individual I19 showing upper limb amelia featured three calls (chr5:157285900CACGTGGG > C, chr5:157285909CTCGG > C, chr5:157285915CACAAC

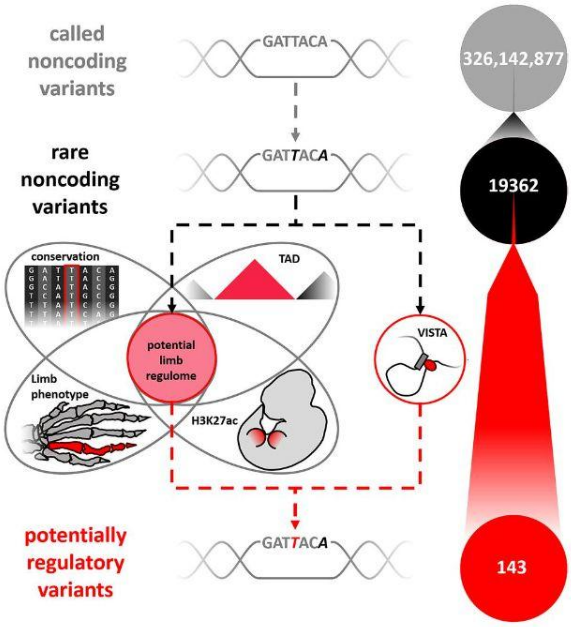


Fig. 4 Pipeline of noncoding data analysis

TG > C) referring to the same indel in intron 1 (15 bp downstream of the first exon–intron boundary) of *CLINT1*. However, *CLINT1* shows only a moderate pLI score (0.54) and there is no evidence other than increased H3K27ac marks of its promoter region in human limb buds linking it to limb development.

We were not able to identify any rare variants in validated VISTA enhancers that showed enhancer activity in the limb bud.

Next, we focused on noncoding variants that were located close to one another in more than one case. In total, 3425 rare noncoding variants in the unsolved cases were positioned 300 bp or less apart from a variant in another unsolved case. 16 of these calls were located within the limb regulome, but in five of these variants, the other variant was positioned outside of the limb regulome.

In two cases both variants were positioned within the limb regulome and within 300 bp: individual I21 and individual I22, both showing finger syndactyly, harbored the overlapping deletions chr22:24552064GGGGGCCGG GACTGGGGCCGGGACT > G and chr22:24552086ACT GGGGCCGGGG > A, respectively. The deletions are positioned in intron 29 of *CABIN1*, in an evolutionarily partially conserved element, that shows increased H3K27ac marks in human embryonic limb buds. However, both deletions were inherited from unaffected parents.

Eight of the close variants were double hits (i.e. we detected rare calls not in just one but two index patients at four positions of the potential limb regulome). However, none of these four pairs of index patients showed overlapping phenotypes.

We identified no coding variant of a known limb disease gene in trans with a conserved, rare noncoding variant of the same TAD.

In summary, despite extensive efforts, we were not able to identify any noncoding SNVs that showed convincing evidence to be causal in congenital limb malformations.

Discussion

In this study, we set out to determine the potential of GS as a comprehensive diagnostic tool to determine all kinds of genetic variants associated with congenital limb malformation.

In our cohort of patients with congenital limb malformations, GS was able to detect both previously described and novel causative genetic variants in already established limb malformation associated genes. In addition, it enabled the identification of three candidate genes and the independent verification of the novel disease gene *UBA2* for causing ectrodactyly (Yamoto et al. 2019). Our approach was able to detect SNVs and structural variants. Finally, GS proved to be a powerful strategy to identify genomic variants previously missed by most other approaches, including repeat expansions and complex structural variants. In total, we identified variants that we consider to be likely pathogenic/disease-associated in 12 of 69 cases (17.4%). This diagnostic yield is comparable to the recent landmark study conducted by the British National Health Service that used GS in cohorts with other congenital disease entities (Turro et al. 2020). A clear advantage of GS compared to most other technologies is the ability to detect copy number neutral variants and to gain position information on CNVs. In our cohort of only 69, we were able to detect two complex variants, an inversion at the *FGF8* locus and a translocated triplication including the *SHH* gene. Both were missed by standard technologies. Further research is necessary to clarify their exact pathomechanisms. The variants identified in the genes *HMGB1*, *SEMA3D* and *ALDH1A2* are all likely to cause loss of function. The genes were previously associated with vertebrate limb development in animal studies and the variants either arose de novo or segregate with the respective phenotype. However, we could not identify unrelated individuals featuring comparable variants in these candidate genes and similar phenotypes. Future research is necessary to identify such to establish the described candidates as disease genes. Our findings once again highlight the role of GS as an attractive one-test-for-all strategy for clinically very heterogeneous cohorts such as congenital malformation syndromes or intellectual disability (Gilissen et al. 2014; Turro et al. 2020). The total cost of the various conventional tests currently used in clinical routine far exceeds that of trio GS.

One of the main challenges of GS data is the medical interpretation of changes in the noncoding DNA. While most clinical GS studies tend to ignore noncoding SNVs (Gilissen et al. 2014) there are recent anecdotal reports of noncoding variants as the cause of Mendelian disorders (Lettice et al. 2003; Jeong et al. 2008; Albers et al. 2012; Bhatia et al. 2013; Weedon et al. 2014; Bae et al. 2014), although there is no established systematic approach, yet. Therefore, we set out to develop a framework to prioritize such noncoding variants associated with congenital limb malformation. We used a combinatorial approach of mouse and human epigenetic data, in vivo validated enhancer sequences, knock-out mice, and the recent knowledge about 3D genome folding, and the cis-regulatory architecture of the genome to define a limb regulome. This limb regulome consists of 0.24% of the genome and includes all known in vivo-validated limb enhancer elements. Contrary to our expectation, we could only identify candidate loci, but no definitely pathogenic noncoding variants. These findings are in stark contrast to our recent study where we demonstrated that CNVs affecting noncoding regulatory elements are a major cause of congenital limb malformations (Flöttmann et al. 2018).

While our results suggest that GS is sensitive to classical sequence variants, it is noteworthy that the method cannot detect epigenetic variants. Epimutations (e.g. imprinting defects) are known to cause inheritable human disease. However, to our knowledge, no epimutation has been linked to congenital limb malformation yet.

Tools for the analysis of GS data are continuously being developed further and the precision of algorithms to call structural variants can certainly be improved. We expect the diagnostic rate to increase steadily with the accuracy of the instruments invoked to analyze GS data.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00439-021-02295-y.

Acknowledgements We would like to thank all families for their collaboration and contribution to this project. J.E. is supported by a scholarship of the Studienstiftung des deutschen Volkes e.V. M.A.M. is a participant in the BIH Charité Digital Clinician Scientist Program founded by the late Prof. Duska Dragun and funded by the Charité-Universitätsmedizin Berlin and the Berlin Institute of Health.

Author contributions Jonas Elsner, Martin A. Mensah, Stefan Mundlos and Malte Spielmann designed the study. Jonas Elsner, Martin A. Mensah, Manuel Holtgrewe, Jakob Hertzberg and Marie Coutelier performed the variant filtration and processed data. Jonas Elsner, Martin A. Mensah and Malte Spielmann investigated and together with Stefan Mundlos and Denise Horn critically evaluated variants. Martin A. Mensah, Stefania Bigoni, Andreas Busche, Deepthi C. de Silva, Nursel Elçioglu, Isabel Filges, Erica Gerkes, Katta Girisha, Luitgard Graul-Neumann, Jane Hurst, Aleksander Jamsheer, Peter Krawitz, Ingo Kurth, Susanne Markus, Andre Megarbane, André Reis, Miriam Reuter, Daniel Svoboda, Christopher Teller, Beyhan Tuysuz, Seval Türkmen, Meredith Wilson, Rixa Woitschach, Inga Vater, Almuth Caliebe, Wiebke Hülsemann, Denise Horn, Stefan Mundlos and Malte Spielmann examined the patients and provided samples. Jonas Elsner and Martin A. Mensah visualised the workflow and results. Jonas Elsner, Martin A. Mensah, Stefan Mundlos and Malte Spielmann wrote and reviewed the manuscript with the approval of all authors. Stefan Mundlos and Malte Spielmann supervised the project.

Funding Open Access funding enabled and organized by Projekt DEAL. A.J. was supported by the grant from the Polish National Science Centre UMO-2016/22/E/NZ5/00270. M.S. is supported by grants from the Deutsche Forschungsgemeinschaft (DFG) (SP1532/3-1, SP1532/4-1 and SP1532/5-1), the Max Planck Foundation and the Deutsches Zentrum für Luft- und Raumfahrt (DLR 01GM1925).

Code and data availability Code and data (pseudonymized, grouped where possible, and minimized) are available upon reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval This study was approved by the ethics committees of the Charité-Universitätsmedizin Berlin (EA2/087/15) and the Medical Chamber (Ärztekammer) Hamburg and conducted in accordance with the declaration of Helsinki. The Charité's ethic commission served as the central IRB.

Consent to participate Participants or their legal guardians provided written informed consent for participation in this study. Consents of all participants were archived. All published data has been pseudonymized, grouped where possible, and minimized.

Consent for publication Participants or their legal guardians provided written informed consent for pseudonymized publication of their data. Consents of all participants were archived. All published data have been pseudonymized, grouped where possible, and minimized.

References

Abe KT, Rizzo IMPO, Coelho ALV et al (2018) 19q13.11 microdeletion: clinical features overlapping ectrodactyly ectodermal dysplasia-clefting syndrome phenotype. Clin Case Rep 6:1300–1307. https://doi.org/10.1002/ccr3.1600

Aerden M, Bauters M, Van Den Bogaert K et al (2020) Genotype-phenotype correlations of UBA2 mutations in patients with ectrodactyly. Eur J Med Genet 63:104009. https://doi.org/10.1016/j.ejmg.2020.104009

Albers CA, Paul DS, Schulze H et al (2012) Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit RBM8A causes TAR syndrome. Nat Genet 44(435–9):S1-2. https://doi.org/10.1038/ng.1083

Anggayasti WL, Ogino K, Yamamoto E et al (2020) The acidic tail of HMGB1 regulates its secondary structure and conformational flexibility: a circular dichroism and molecular dynamics simulation study. Comput Struct Biotechnol J 18:1160–1172

Bae B-I, Tietjen I, Atabay KD et al (2014) Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. Science 343:764–768. https://doi.org/10.1126/science.1244392

Belgrano FS, de Abreu da Silva IC, Bastos de Oliveira FM et al (2013) Role of the acidic tail of high mobility group protein B1 (HMGB1) in protein stability and DNA bending. PLoS ONE 8:e79572. https://doi.org/10.1371/journal.pone.0079572

Berisha SZ, Shetty S, Prior TW, Mitchell AL (2020) Cytogenetic and molecular diagnostic testing associated with prenatal and postnatal birth defects. Birth Defects Res 112:293–306. https://doi.org/10.1002/bdr2.1648

Bhatia S, Bengani H, Fish M et al (2013) Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. Am J Hum Genet 93:1126–1134. https://doi.org/10.1016/j.ajhg.2013.10.028

Brison N, Debeer P, Tylzanowski P (2014) Joining the fingers: a HOXD13 story. Dev Dyn 243:37–48. https://doi.org/10.1002/dvdy.24037

Cao J, Spielmann M, Qiu X et al (2019) The single-cell transcriptional landscape of mammalian organogenesis. Nature 566:496–502. https://doi.org/10.1038/s41586-019-0969-x

Chowdhury S, Bandholz AM, Parkash S et al (2014) Phenotypic and molecular characterization of 19q12q13.1 deletions: a report of five patients. Am J Med Genet A 164A:62–69. https://doi.org/10.1002/ajmg.a.36201

Cotney J, Leng J, Yin J et al (2013) The evolution of lineage-specific regulatory activities in the human embryonic limb. Cell 154:185–196. https://doi.org/10.1016/j.cell.2013.05.056

de Mollerat XJ, Gurrieri F, Morgan CT et al (2003) A genomic rearrangement resulting in a tandem duplication is associated with split hand-split foot malformation 3 (SHFM3) at 10q24. Hum Mol Genet 12:1959–1971

Deciphering Developmental Disorders Study (2017) Prevalence and architecture of de novo mutations in developmental disorders. Nature 542:433–438. https://doi.org/10.1038/nature21062

Dixon JR, Selvaraj S, Yue F et al (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485:376–380. https://doi.org/10.1038/nature11082

Flöttmann R, Knaus A, Zemojtel T et al (2015) FGFR2 mutation in a patient without typical features of Pfeiffer syndrome—the emerging role of combined NGS and phenotype based strategies. Eur J Hum Genet 58:376–380. https://doi.org/10.1016/j.ejmg.2015.05.007

Flöttmann R, Kragesteen BK, Geuer S et al (2018) Noncoding copy-number variations are associated with congenital limb malformation. Genet Med 20:599–607. https://doi.org/10.1038/gim.2017.154

Gilissen C, Hehir-Kwa JY, Thung DT et al (2014) Genome sequencing identifies major causes of severe intellectual disability. Nature 511:344–347. https://doi.org/10.1038/nature13394

Govindan J, Tun KM, Iovine MK (2016) Cx43-dependent skeletal phenotypes are mediated by interactions between the Hapln1a-ECM and Sema3d during Fin regeneration. PLoS ONE 11:e0148202. https://doi.org/10.1371/journal.pone.0148202

Holder-Espinasse M, Jamsheer A, Escande F et al (2019) Duplication of 10q24 locus: broadening the clinical and radiological spectrum. Eur J Hum Genet 27:525–534. https://doi.org/10.1038/s41431-018-0326-9

Holtgrewe M, Stolpe O, Nieminen M et al (2020) VarFish: comprehensive DNA variant analysis for diagnostics and research. Nucleic Acids Res 48:W162–W169. https://doi.org/10.1093/nar/gkaa241

Itou J, Taniguchi N, Oishi I et al (2011) HMGB factors are required for posterior digit development through integrating signaling pathway activities. Dev Dyn 240:1151–1162

Jeong Y, Leskow FC, El-Jaick K et al (2008) Regulation of a remote Shh forebrain enhancer by the Six3 homeoprotein. Nat Genet 40:1348–1353. https://doi.org/10.1038/ng.230

Klopocki E, Lohan S, Doelken SC et al (2012) Duplications of BHLHA9 are associated with ectrodactyly and tibia hemimelia inherited in non-Mendelian fashion. J Med Genet 49:119–125. https://doi.org/10.1136/jmedgenet-2011-100409

Kroeldrup L, Kjaergaard S, Kirchhoff M et al (2012) Duplication of 7q36.3 encompassing the Sonic Hedgehog (SHH) gene is associated with congenital muscular hypertrophy. Eur J Med Genet 55:557–560. https://doi.org/10.1016/j.ejmg.2012.04.009

Lettice LA, Heaney SJH, Purdie LA et al (2003) A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet 12:1725–1735. https://doi.org/10.1093/hmg/ddg180

Levy SE, Myers RM (2016) Advancements in next-generation sequencing. Annu Rev Genomics Hum Genet 17:95–115. https://doi.org/10.1146/annurev-genom-083115-022413

Lindstrand A, Eisfeldt J, Pettersson M et al (2019) From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. Genome Med 11:68. https://doi.org/10.1186/s13073-019-0675-1

Lohan S, Spielmann M, Doelken SC et al (2014) Microduplications encompassing the Sonic hedgehog limb enhancer ZRS are associated with Haas-type polysyndactyly and Laurin-Sandrow syndrome. Clin Genet 86:318–325. https://doi.org/10.1111/cge.12352

Muenke M, Schell U, Hehr A et al (2014) A common mutation in the fibroblast growth factor receptor 1 gene in Pfeiffer syndrome. Nat Genet 8:269–274. https://doi.org/10.1038/ng1194-269

Nieminen M, Stolpe O, Schumann F et al (2020) SODAR core: a Django-based framework for scientific data management and analysis web apps. J Open Source Softw JOSS 55:1584. https://doi.org/10.21105/joss.01584

Olsen SK, Capili AD, Lu X et al (2010) Active site remodelling accompanies thioester bond formation in the SUMO E1. Nature 463:906–912. https://doi.org/10.1038/nature08765

Sanchez-Castro M, Pichon O, Briand A et al (2015) Disruption of the SEMA3D gene in a patient with congenital heart defects. Hum Mutat 36:30–33. https://doi.org/10.1002/humu.22702

Shou S, Carlson HL, Perez WD, Stadler HS (2013) HOXA13 regulates Aldh1a2 expression in the autopod to facilitate interdigital programmed cell death. Dev Dyn 242:687–698. https://doi.org/10.1002/dvdy.23966

Singh A, Mia MM, Cibi DM et al (2019) Deficiency in the secreted protein Semaphorin3d causes abnormal parathyroid development in mice. J Biol Chem 294:8336–8347. https://doi.org/10.1074/jbc.RA118.007063

Turro E, Astle WJ, Megy K et al (2020) Whole-genome sequencing of patients with rare diseases in a national health system. Nature 583:96–102. https://doi.org/10.1038/s41586-020-2434-2

Visel A, Blow MJ, Li Z et al (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457:854–858. https://doi.org/10.1038/nature07730

Weedon MN, Cebola I, Patch A-M et al (2014) Recessive mutations in a distal PTF1A enhancer cause isolated pancreatic agenesis. Nat Genet 46:61–64. https://doi.org/10.1038/ng.2826

Xue Y, Ankala A, Wilcox WR, Hegde MR (2015) Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. Genet Med 17:444–451. https://doi.org/10.1038/gim.2014.122

Yamoto K, Saitsu H, Nishimura G et al (2019) Comprehensive clinical and molecular studies in split-hand/foot malformation: identification of two plausible candidate genes (LRP6 and UBA2). Eur J Hum Genet 27:1845–1857. https://doi.org/10.1038/s41431-019-0473-7

Zhao X, Brade T, Cunningham TJ, Duester G (2010) Retinoic acid controls expression of tissue remodeling genes Hmgn1 and Fgf18 at the digit-interdigit junction. Dev Dyn 239:665–671

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Jonas Elsner[1] · Martin A. Mensah[1,2] · Manuel Holtgrewe[3] · Jakob Hertzberg[4] · Stefania Bigoni[5] · Andreas Busche[6] · Marie Coutelier[1,7] · Deepthi C. de Silva[8] · Nursel Elçioglu[9,10] · Isabel Filges[11] · Erica Gerkes[12] · Katta M. Girisha[13] · Luitgard Graul-Neumann[1] · Aleksander Jamsheer[14] · Peter Krawitz[15] · Ingo Kurth[16] · Susanne Markus[17] · Andre Megarbane[18] · André Reis[19] · Miriam S. Reuter[19] · Daniel Svoboda[20] · Christopher Teller[21] · Beyhan Tuysuz[22] · Seval Türkmen[1,23] · Meredith Wilson[24] · Rixa Woitschach[25] · Inga Vater[26] · Almuth Caliebe[26] · Wiebke Hülsemann[27] · Denise Horn[1] · Stefan Mundlos[1,4] · Malte Spielmann[4,26,28]

[1] Institute of Medical Genetics and Human Genetics, Charité-Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

[2] Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany

[3] Core Unit Bioinformatics, Berlin Institute of Health (BIH), Berlin, Germany

[4] Max Planck Institute for Molecular Genetics, RG Development and Disease, Berlin, Germany

[5] Medical Genetics Unit, Department of Mother and Child, Ferrara Sant'Anna University Hospital, Ferrara, Italy

[6] Institut Für Humangenetik, Universitätsklinikum Münster, Münster, Germany

[7] Department of Human Genetics, Faculty of Medicine, Jewish General Hospital, McGill University, Montreal, QC, Canada

[8] Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

[9] Department of Pediatric Genetics, Marmara University Medical School, Istanbul, Turkey

[10] Eastern Mediterranean University Medical School, Cyprus, Mersin 10, Turkey

[11] Institut für Medizinische Genetik und Pathologie, Universitätsspital Basel, Basel, Switzerland

[12] Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

[13] Department of Medical Genetics, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, India

[14] Department of Medical Genetics, Poznan University of Medical Sciences, Poznan, Poland

[15] Institute for Genomic Statistics and Bioinformatics, University of Bonn, Bonn, Germany

[16] Institute of Human Genetics, Medical Faculty, RWTH Aachen University Hospital, Aachen, Germany

[17] Fachärztin Für Humangenetik, Bischof-von-Henle-Straße 2a, Regensburg, Germany

[18] Department of Human Genetics, Gilbert and Rose-Marie Chagoury School of Medicine, Lebanese American University, Byblos, Lebanon

[19] Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

[20] Kinderhandchirurgie, Medizinische Fakultät Mannheim der Universität Heidelberg, Heidelberg, Germany

[21] Synlab MVZ Bad Nauheim, Mondorfstr. 1761231, Bad Nauheim, Germany

[22] Department of Pediatric Genetics, Cerrahpasa Medical Faculty, Istanbul University-Cerrahpasa, Istanbul, Turkey

[23] National Center of Genetics (NCG), Laboratoire National de Santé 1, Rue Louis Rech, 3555 Dudelange, Luxembourg

[24] Genetic Medicine, Children's Hospital at Westmead, Paediatrics and Child Health, Sydney, Australia

[25] Institute of Human Genetics, University Medical Center Hamburg, Eppendorf, Germany

[26] Institute of Human Genetics, University of Kiel, Kiel, Germany

[27] Katholisches Kinderkrankenhaus Wilhelmstift, Hamburg, Germany

[28] Institute of Human Genetics, University of Lübeck, Lübeck, Germany

2.6 Aberrante Phasentrennung und nukleolare Dysfunktion bei seltenen genetischen Krankheiten

Mensah MA*, Niskanen H*, Magalhaes AP, Basu S, Kircher M, Sczakiel HL, Reiter AMV, Elsner J, Meinecke P, Biskup S, Chung BHY, Dombrowsky G, Eckmann-Scholz C, Hitz MP, Hoischen A, Holterhus PM, Hülsemann W, Kahrizi K, Kalscheuer VM, Kan A, Krumbiegel M, Kurth I, Leubner J, Longardt AC, Moritz JD, Najmabadi H, Skipalova K, Snijders Blok L, Tzschach A, Wiedersberg E, Zenker M, Garcia-Cabau C, Buschow R, Salvatella X, Kraushar ML, Mundlos S, Caliebe A, Spielmann M, Horn D, Hnisz D.

Aberrant phase separation and nucleolar dysfunction in rare genetic diseases.

Nature. 2023 Feb;614(7948):564-571. doi: 10.1038/s41586-022-05682-1

*contributed equally

Zwar gehören genetische Syndrome typischerweise zu den seltenen Erkrankungen, die Häufigkeiten der einzelnen genetischen Krankheitsentitäten unterscheiden sich jedoch erheblich. Die Hürden der genetischen Diagnostik sind besonders in der Gruppe der ultra-seltenen genetischen Syndrome - also solcher Syndrome, von denen nur wenige Dutzend Fallberichte in der Literatur existieren - hoch. Die phänotypischen Merkmale dieser Erkrankungen sind vielen Untersuchern unbekannt und deren molekulargenetische Ursachen sind häufig unklar. Aufgrund der extremen Seltenheit dieser Erkrankungen ist die Bildung von Studienkohorten nur eingeschränkt möglich.

Nachdem wir *HMGB1* aufgrund der *de novo* Frameshift-Mutation *HMGB1*(NM_002128.7):c.551_554delAGAA p.(Lys184Argfs*44) bei einem weiblichen Feten mit komplexer Extremitätenfehlbildung als Kandidatengen identifiziert hatten, wurde bei einem Kind mit der klinischen Diagnose BPTAS mittels WGS die nahezu identische Mutation *HMGB1*(NM_002128.7):c.556_559delGAAG p.(Glu186Argfs*42) detektiert. Lediglich neun Patienten mit der Diagnose BPTAS waren zu diesem Zeitpunkt in der Literatur beschrieben, die molekulare Ursache war unbekannt. Durch eine reverse Phänotypisierung konnten bei dem oben genannten Fetus und bei einem weiteren Patienten, bei dem ein WES eine identische *de novo* Mutation identifiziert hatte, die Diagnose BPTAS vergeben werden. Eine gezielte Sequenzierung von *HMGB1* bei zwei weiteren Patientinnen mit der klinischen Diagnose BPTAS zeigte erneut denselben Frameshift. Mittels fazialem NGP durch DeepGestalt konnte darüber hinaus eine hohe Übereinstimmung des fazialen Aspekts von diesen und von in der Literatur beschriebenen BPTAS-Patienten gezeigt werden. So war das

System zwar nicht für ein BPTAS trainiert, zeigte aber bei sechs von neun geeigneten Patientenbildern die Diagnose Blepharophimose-Ptose-und-Epikanthus-inversus-Syndrom, welche bei Kontrollen (syndromal und unauffällig) nur sehr selten genannt wird. Da die frameshift-Mutationen bei allen sequenzierten Patienten zu einem Austausch der hoch konservierten, sauren, C-terminalen Domäne von HMGB1 durch eine basische Domäne führten, und bei vier Patienten nachweislich *de novo* aufgetreten waren (bei einem Patienten standen die gesunden Eltern nicht zur Testung zur Verfügung), war eine Pathogenität anzunehmen.

Allerdings sind Loss-of-Function-Mutationen von HMGB1 (u.a. Deletionen des gesamten Gens) mit einer syndromalen geistigen Entwicklungsverzögerung assoziiert, deren skelettale Merkmale nicht mit einem BPTAS vergleichbar sind. Wir nahmen daher an, dass die BPTAS-assoziierten Frameshifts mit einem Gain-of-Function-Mechanismus einhergingen.

Bei der C-terminalen, sauren Domäne von HMGB1 handelt es sich um eine intrinsisch ungeordnete Region (*intrinsically disordered region, IDR*). IDRs spielen eine wichtige Rolle bei der Bildung von intrazellulären Kompartimenten, die nicht durch eine Lipidmembran voneinander abgegrenzt sind, sog. biomolekularen Kondensaten. Ein Beispiel für ein solches Kompartiment ist der Nukleolus. Wir konnten zeigen, dass die BPTAS-assoziierten HMGB1-frameshifts zu einer veränderten Phasentrennung im Zellkern führten und, dass sich entsprechend mutiertes HMGB1, dessen Wildtyp-Variante sich üblicherweise im Karyoplasma, aber nicht im Nukleolus findet, in die granuläre Komponente des Nukleolus abtrennte.

Eine Analyse von Mutationsdatenbanken und eine funktionelle Aufarbeitung ausgewählter Varianten zeigte, dass auch Frameshift-Mutationen anderer Gene zu einem pathogenen IDR-Austausch führten. Bei dem frameshift-induzierten IDR-Austausch handelt es sich folglich nicht um einen HMGB1-spezifischen Pathomechanismus, sondern um einen grundsätzlichen Effekt, der wahrscheinlich die funktionelle Bedeutung zahlreicher pathogener Varianten erklären kann.

Die Ergebnisse zeigen, dass ein WGS das Potenzial hat, in Kombination mit einem fazialen NGP eine reverse Phänotypisierung auch ultra-seltener Syndrome zu ermöglichen, deren molekulare Ursache aufzuklären und dabei nicht nur Gen-Syndrom-Assoziationen, sondern auch Varianten-Syndrom-Assoziationen zu identifizieren.

Aberrant phase separation and nucleolar dysfunction in rare genetic diseases

https://doi.org/10.1038/s41586-022-05682-1

Received: 9 November 2021

Accepted: 22 December 2022

Published online: 8 February 2023

Open access

Check for updates

Martin A. Mensah[1,2,3,32], Henri Niskanen[4,32], Alexandre P. Magalhaes[4], Shaon Basu[4], Martin Kircher[5,6], Henrike L. Sczakiel[1,2,3], Alisa M. V. Reiter[1], Jonas Elsner[1], Peter Meinecke[7], Saskia Biskup[8], Brian H. Y. Chung[9], Gregor Dombrowsky[10,11], Christel Eckmann-Scholz[12], Marc Phillip Hitz[10,11], Alexander Hoischen[13,14], Paul-Martin Holterhus[15], Wiebke Hülsemann[16], Kimia Kahrizi[17], Vera M. Kalscheuer[3], Anita Kan[18], Mandy Krumbiegel[19], Ingo Kurth[20], Jonas Leubner[21], Ann Carolin Longardt[22], Jörg D. Moritz[23], Hossein Najmabadi[17], Karolina Skipalova[1], Lot Snijders Blok[14], Andreas Tzschach[24], Eberhard Wiedersberg[25], Martin Zenker[26], Carla Garcia-Cabau[27], René Buschow[28], Xavier Salvatella[27,29], Matthew L. Kraushar[4], Stefan Mundlos[1,2,3,30], Almuth Caliebe[6], Malte Spielmann[3,6,31,33] ✉, Denise Horn[1,33] ✉ & Denes Hnisz[4,33] ✉

Thousands of genetic variants in protein-coding genes have been linked to disease. However, the functional impact of most variants is unknown as they occur within intrinsically disordered protein regions that have poorly defined functions[1–3]. Intrinsically disordered regions can mediate phase separation and the formation of biomolecular condensates, such as the nucleolus[4,5]. This suggests that mutations in disordered proteins may alter condensate properties and function[6–8]. Here we show that a subset of disease-associated variants in disordered regions alter phase separation, cause mispartitioning into the nucleolus and disrupt nucleolar function. We discover de novo frameshift variants in *HMGB1* that cause brachyphalangy, polydactyly and tibial aplasia syndrome, a rare complex malformation syndrome. The frameshifts replace the intrinsically disordered acidic tail of HMGB1 with an arginine-rich basic tail. The mutant tail alters HMGB1 phase separation, enhances its partitioning into the nucleolus and causes nucleolar dysfunction. We built a catalogue of more than 200,000 variants in disordered carboxy-terminal tails and identified more than 600 frameshifts that create arginine-rich basic tails in transcription factors and other proteins. For 12 out of the 13 disease-associated variants tested, the mutation enhanced partitioning into the nucleolus, and several variants altered rRNA biogenesis. These data identify the cause of a rare complex syndrome and suggest that a large number of genetic variants may dysregulate nucleoli and other biomolecular condensates in humans.

Monogenic and common diseases are frequently associated with mutations in transcriptional regulatory proteins, including DNA-binding transcription factors. However, the functional impact of the majority of such mutations is unknown, and many complex diseases still lack a clear underlying genetic component[1,9–12]. We initially set out to identify the molecular basis of brachyphalangy, polydactyly and tibial aplasia/hypoplasia syndrome (BPTAS; Online Mendelian Inheritance in Man database identifier: 609945), an extremely rare complex malformation syndrome with an as yet unknown molecular aetiology[13–19]. During the study, five individuals (I1–I5) were diagnosed with BPTAS. All five exhibited a distinct skeletal phenotype, including short and malformed lower limbs characterized by tibia aplasia or hypoplasia, preaxial polysyndactyly and contractures of large joints (Fig. 1a,b). In all five individuals, anomalies of the upper limbs were less severe compared with those of the lower limbs, and included brachydactyly or brachyphalangy of fingers with an irregular finger length (Fig. 1a,b). Short radius and ulna and contractures or pterygia of the elbow joints were present in four out of five individuals. All individuals with BPTAS diagnosed during our study or described in previous reports also presented with distinct craniofacial, neurological and genitourinary features. Phenotypic findings are summarized in Supplementary Table 1. Detailed clinical and family histories are provided in Supplementary Note and Extended Data Fig. 1.

De novo *HMGB1* frameshifts in BPTAS

We performed genome sequencing of I1 and detected a potentially pathogenic variant: the heterozygous frameshift NM_002128.7(*HMGB1*): c.556_559delGAAG;p.(Glu186Argfs*42) in the final exon of *HMGB1* (Extended Data. Fig. 2a). *HMGB1* encodes a highly conserved, low-specificity DNA binding factor associated with cell signalling[20], cell motility[21,22], base excision repair[23,24] and chromatin looping[25].

A list of affiliations appears at the end of the paper.

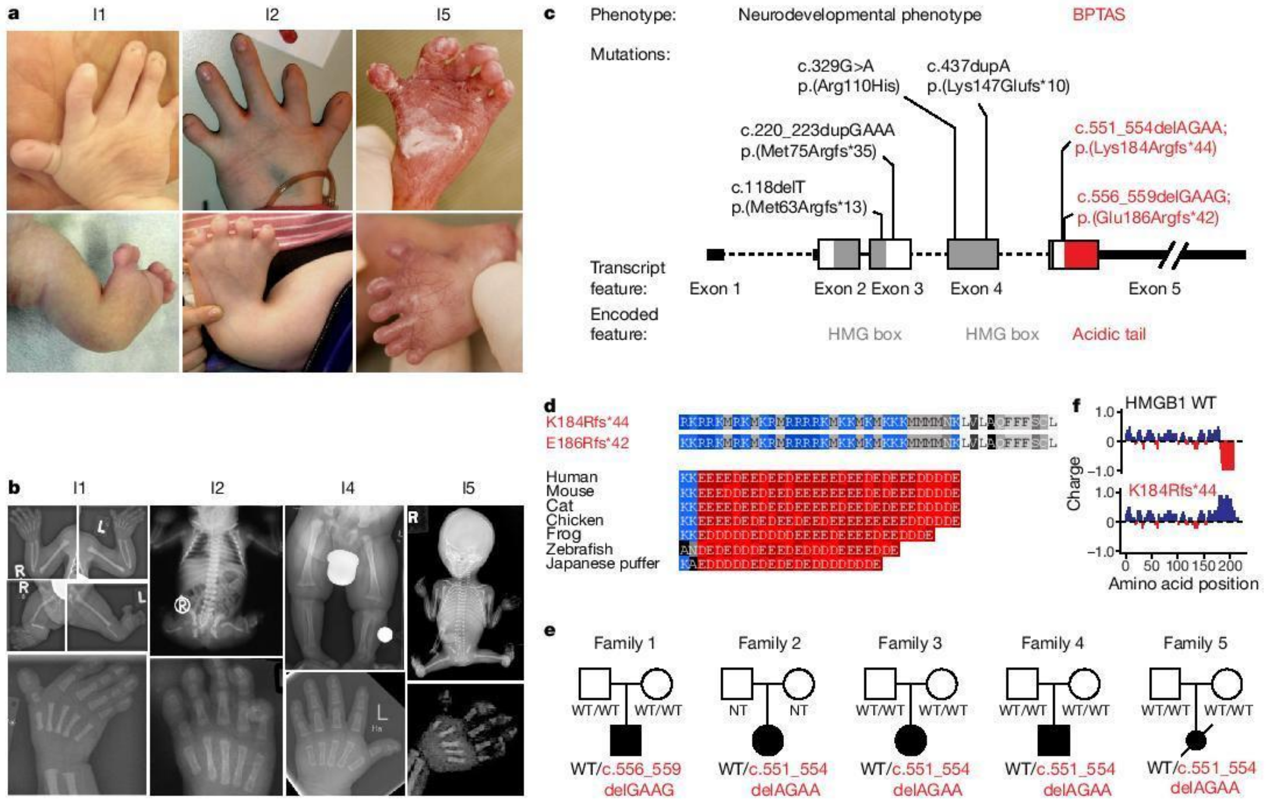


Fig. 1 | De novo frameshifts in *HMGB1* cause BPTAS. a, Photographs of individuals diagnosed with BPTAS. Top row, hands of I1, I2 and I5. Note brachydactyly, irregular finger length and hypoplasia of the nails. Bottom row, lower extremities of I1, I2 and I5, presenting with malformed legs, joint contractures, preaxial polysyndactyly and hypoplasia of the nails. **b**, Radiograms of I1, I2, I4 and I5. Top far left, limb radiograms (at newborn age) of I1 showing brachydactyly and brachyphalangy, tibial aplasia, hypoplastic fibulae and preaxial polysyndactyly. Top middle left, babygram of I2. Note tibial aplasia, hypoplastic and absent fibulae, hypoplastic pelvic bones and hypoplastic right femur. Top middle right, lower extremities of I4 (at 6 months) showing asymmetric shortness of tibiae and fibulae. Top far right, fetogram of I5 showing tibial aplasia, hypoplastic and absent fibulae, hypoplastic pelvic bones and contractures of joints. Bottom row, hand radiograms of I1, I2 (both at newborn age), I4 (at 6 months) and of I5 (at 21 weeks of gestation). Note the short middle phalanges and short proximal phalanges of the thumbs. **c**, Pathogenic frameshift variants in the acidic tail of HMGB1 in the individuals with BPTAS reported in this article are highlighted in red. Previously reported variants associated with developmental delay are in black. Note the genotype–phenotype correlation: C-terminal frameshifts result in BPTAS, whereas other variants lead to a neurodevelopmental phenotype. **d**, Amino acid sequence of the C terminus of HMGB1 in individuals with BPTAS and in selected vertebrates. Acidic residues glutamate and aspartate are shaded in red, basic residues arginine and lysine are shaded in blue. Note the replacement of the conserved acidic tail in individuals with BPTAS. **e**, Family pedigrees. Individuals with BPTAS are highlighted with black boxes, and the genotypes are below the boxes. **f**, Charge plots of WT and mutant HMGB1. I, individual; L, left; NT, not tested; R, right; WT, wild type.

Sanger sequencing of I1 and his parents confirmed the presence of the frameshift variant and revealed de novo occurrence (Fig. 1c–e and Extended Data Fig. 2b). Sanger sequencing of *HMGB1* in I2 and I3, and trio exome sequencing of I4 identified a similar de novo heterozygous frameshift: NM_002128.7(*HMGB1*):c.551_554delAGAA;p.(Lys184Argfs*44), a variant also detected in a previously described female fetus[26] (I5) (Fig. 1c–e and Extended Data Fig. 2c,d). Sequencing of cDNA from a lymphoblastoid cell line derived from peripheral blood cells from I3 confirmed the presence of both wild-type and mutant *HMGB1* transcripts (Extended Data Fig. 2e). The two frameshift mutations result in almost identical, positively charged sequences (Fig. 1d,f).

Altered HMGB1 phase separation in vitro

To investigate the potential pathogenic role of the frameshift variant in HMGB1, we first explored structural and sequence features of the wild-type and mutant proteins. HMGB1 is a low-specificity DNA-binding protein that contains two HMG boxes that are responsible for DNA binding[27] and a C-terminal acidic tail (Fig. 2a,b). The acidic tail is predicted to be intrinsically disordered and resides within an approximately 60-amino-acid long conserved intrinsically disordered region (IDR), as revealed by AlphaFold2 and PONDR analyses (Fig. 2a,b and Extended Data Fig. 3a). Both algorithms predicted a slight propensity of the C-terminal portion of the IDR to assume a helical conformation in the frameshift mutant HMGB1 (Fig. 2a,b and Extended Data Fig. 3b). This prediction was confirmed by circular dichroism experiments on synthetic peptides that corresponded to the C-terminal 80–90 amino acid region (Extended Data Fig. 3c–e). IDRs of numerous proteins, including transcription factors, co-activators (for example, Mediator) and RNA polymerase II (RNAPII), contribute to phase separation by mediating multivalent low-affinity interactions[6,28–31]. Therefore, we hypothesized that the potentially BPTAS-causing frameshift may alter the phase-separation capacity of HMGB1.

To test the phase-separation capacity of HMGB1, we purified recombinant HMGB1 proteins tagged with enhanced green fluorescent protein (eGFP) and examined their behaviour in vitro. Wild-type full-length HMGB1 formed droplets in the presence of a crowding agent (10% polyethylene glycol (PEG)), and the number and size of droplets scaled with

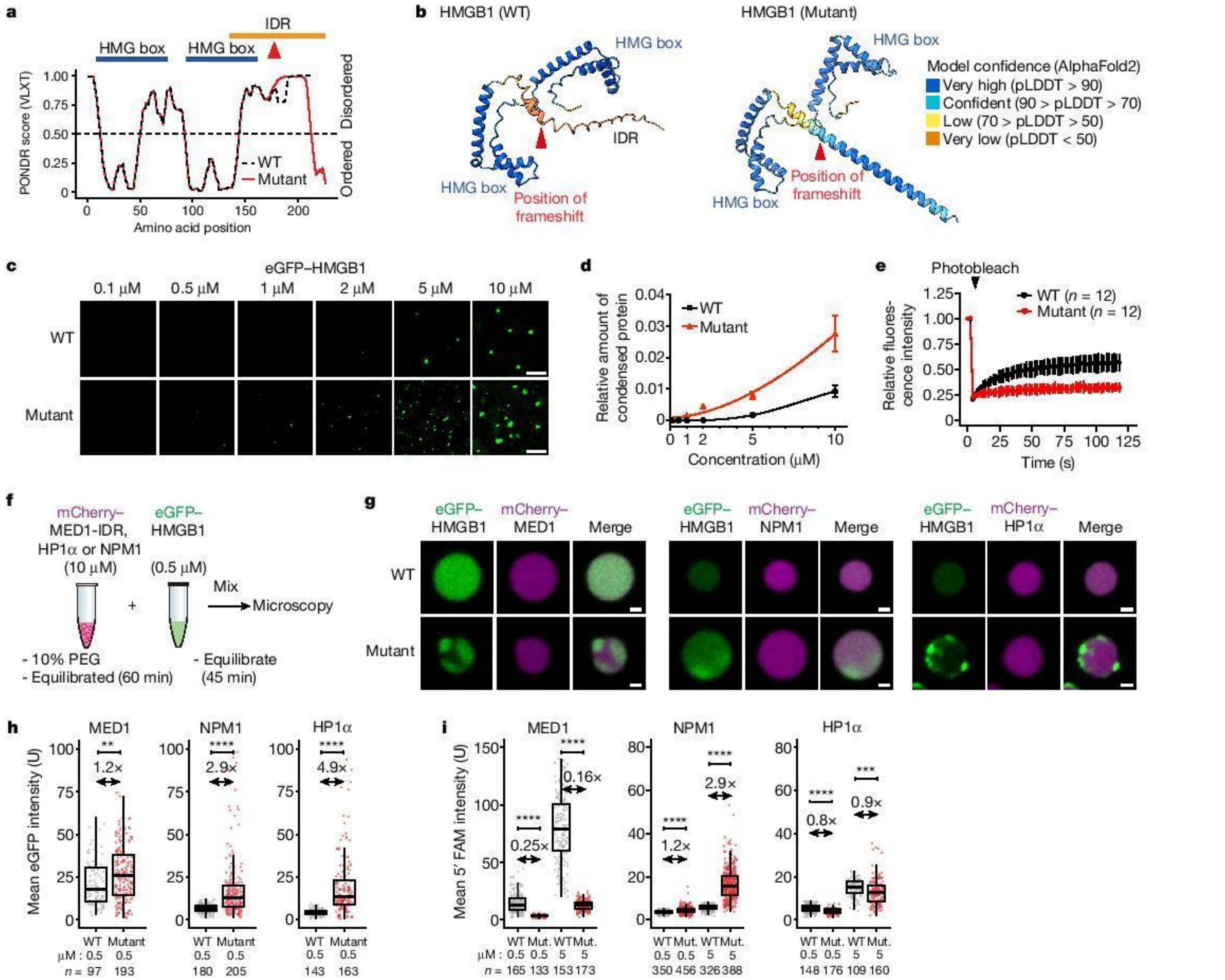


Fig. 2 | A BPTAS-causing frameshift alters HMGB1 phase separation in vitro. **a**, Graph plotting the intrinsic disorder of HMGB1. Red arrowhead shows the position of the BPTAS frameshift. The position of the IDR is highlighted with an orange bar and the position of HMG boxes with blue bars. **b**, Structures of WT and mutant HMGB1 predicted with AlphaFold2. Colours ranging from blue to orange depict the per-residue measure of local confidence (pLDDT) for the model. **c**, Representative images from droplet formation assays of eGFP–HMGB1 variants at the indicated concentrations. The experiment was repeated three times, with similar results obtained. **d**, Quantification of the relative amount of condensed protein at the indicated concentrations. Data displayed as the mean ± s.d. **e**, Relative fluorescence intensity of the bleached area from eGFP–HMGB1 condensates before and after photobleaching. Data displayed as the mean ± s.d. **f**, Scheme of co-droplet assays. **g**, Representative images of eGFP–HMGB1 proteins mixed with preassembled mCherry-labelled MED1-IDR, HP1α or NPM1 droplets. **h**,**i**, Quantification of eGFP (**h**) and 5′ FAM (**i**) fluorescence intensity in mCherry-labelled MED1-IDR, HP1α and NPM1 droplets mixed with full-length mEGFP–HMGB1 proteins (**h**) or 5′ FAM–HMGB1-IDR peptides (**i**). Fold change values between the mean intensities of WT and mutants (Mut.) are indicated above the plot. Median is shown as a line within the boxplot, which spans from the 25th to 75th percentiles. Whiskers depict a 1.5× interquartile range. *P* values are from two-tailed Welch's *t*-test. $^{**}P < 1\times10^{-2}$, $^{***}P < 1\times10^{-3}$, $^{****}P < 1\times10^{-4}$. Scale bars, 5 μm (**c**) and 10 μm (**g**).

the concentration of the protein (Fig. 2c,d and Extended Data Fig. 4a–c). The droplets were spherical, settled on the surface and occasionally underwent fusion (Supplementary Video 1), which are hallmarks of phase separation[32]. By contrast, the frameshift mutant HMGB1 formed amorphous condensates that appeared at a lower saturation concentration (Fig. 2c,d) and, after photobleaching, recovered fluorescence slower than wild-type HMGB1 droplets (Fig. 2e). Similar results were observed using synthetic peptides that corresponded to the C-terminal 80–90 amino acid region of wild-type and frameshift mutant HMGB1 (Extended Data Fig. 4d–g). These results indicate that the frameshift in HMGB1 enhances condensate formation and alters condensate properties in vitro.

Mammalian nuclei contain numerous biomolecular condensates, for example, the nucleolus, heterochromatin, co-activator and RNAPII condensates[4,5]. IDRs play important roles in the partitioning of proteins into nuclear condensates[4,5]. We therefore tested whether the frameshift mutation alters the partitioning of HMGB1 into nuclear condensates. Using purified marker proteins, we assembled the following model condensates: recombinant mCherry-tagged MED1 IDR droplets as an in vitro model for Mediator co-activator condensates[29,31,33]; mCherry-tagged HP1α droplets as an in vitro model for heterochromatin[34]; and mCherry-tagged NPM1 droplets as an in vitro model for the granular component of the nucleolus[35]. Wild-type and

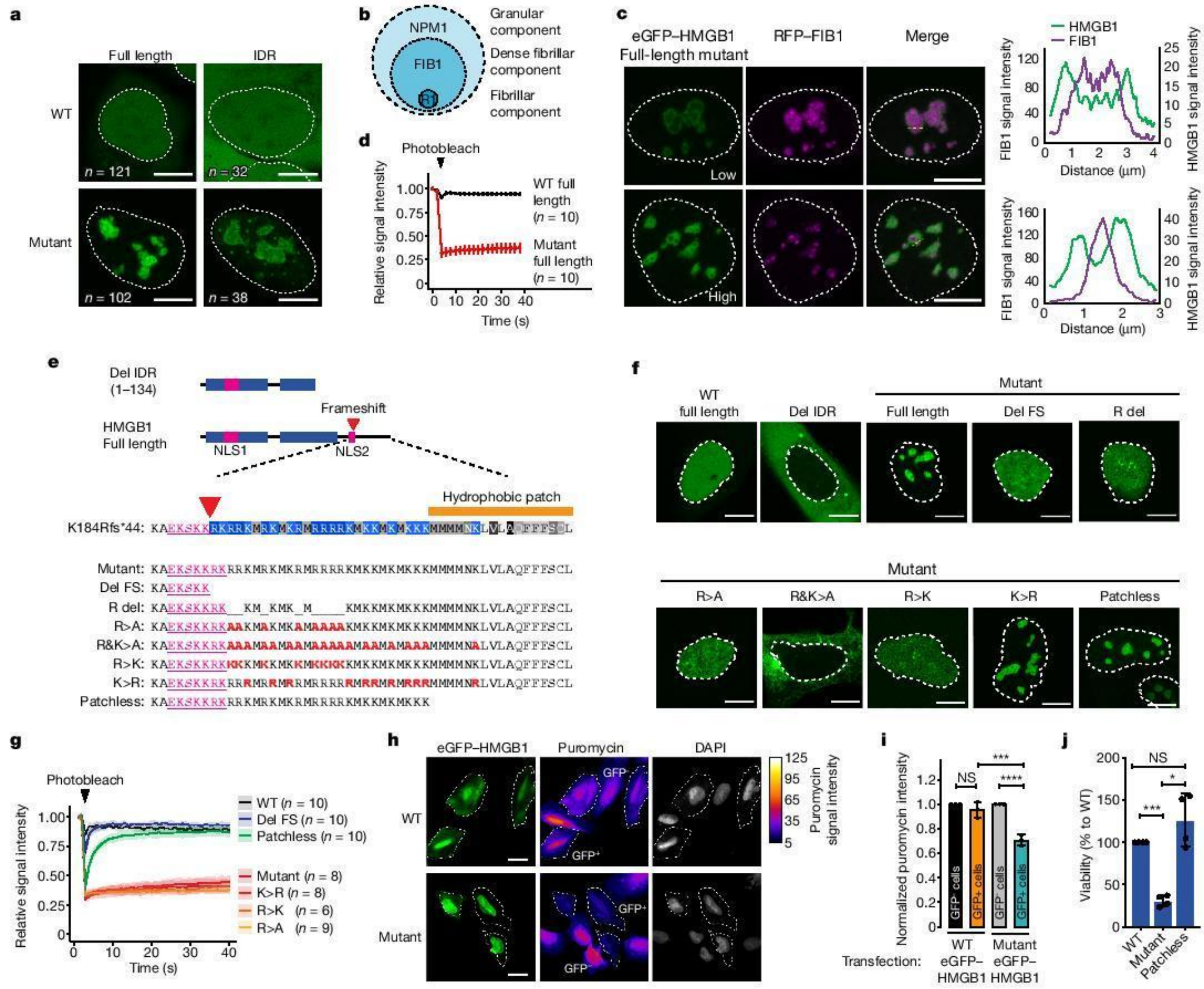


Fig. 3 | Mutant HMGB1 replaces the granular component of the nucleolus in vivo. a, Representative images of live U2OS cells expressing eGFP–HMGB1 proteins. Nuclear area are revealed by Hoechst staining is shown as dashed white lines in **a**, **c** and **f**. **b**, Model of the nucleolus. R1, RNA polymerase I. **c**, Left, representative images of U2OS cells expressing RFP–FIB1 and mutant eGFP–HMGB1. Right, fluorescence intensity profiles from the region highlighted by the dashed yellow line. Low and high indicate nuclei with a relatively low or high amount, respectively, of the mutant protein. **d**, Relative fluorescence intensity of eGFP–HMGB1 before and after photobleaching. Data displayed as the mean ± s.d. **e**, Schematic and sequence representation of HMGB1 variants. Blue bars, HMG boxes. NLS, nuclear localization signal. Red arrow marks the position of the frameshift mutation (K184Rfs*44) and red letters highlight mutagenized amino acids. **f**, Representative images of live U2OS cells expressing the indicated eGFP–HMGB1 variants. **g**, Relative fluorescence intensity of eGFP–HMGB1 variants before and after photobleaching. Data are displayed as a line for the mean signal, with the shaded region representing ± s.d., *n* = number of cells examined. **h**, Representative images from puromycin-staining experiments with U2OS cells ectopically expressing eGFP–HMGB1 proteins. The puromycin signal was used to trace the cell area to highlight GFP$^+$ cells with a dashed line. **i**, Normalized puromycin intensities displayed as the mean ± s.d. from three independent biological replicate experiments. ***$P < 0.0002$, ****$P < 0.0001$ by one-way ANOVA. **j**, Quantification of the viability of cells expressing the indicated HMGB1 proteins. Data displayed as individual points from independent biological replicates (*n* = 4). Bar charts show mean ± s.d. ***$P = 0.0005$, *$P = 0.0177$ by one-way ANOVA. Scale bars, 10 μm (**a**,**c**,**f**) or 20 μm (**h**).

mutant HMGB1 proteins were then added to the droplets (Fig. 2f). Wild-type eGFP-tagged HMGB1 partitioned into all three model condensates, with the highest partitioning observed in MED1-IDR droplets (Fig. 2g,h). The mutant HMGB1 protein displayed enhanced partitioning into NPM1 droplets (threefold compared with wild type, $P < 1 \times 10^{-5}$, Welch's *t*-test) and to some extent in HP1α droplets (Fig. 2g,h). Mutant HMGB1 also tended to form dense foci within the MED1-IDR, HP1α and NPM1 droplets over time that appeared sequestered to the surface of the droplets (Fig. 2g). Enhanced partitioning into NPM1 condensates and foci formation were also observed using a 5′-carboxyfluorescein (5′ FAM)-labelled synthetic HMGB1 IDR mutant peptide, tested at multiple concentrations (Fig. 2i and Extended Data Fig. 4h–j). These results reveal that mutant HMGB1 exhibits enhanced partitioning into NPM1 condensates in vitro.

Nucleolar HMGB1 mispartitioning in vivo

We next sought to investigate the condensate behaviour of mutant HMGB1 in human cells. As primary culturable cells from individuals with BPTAS were not available, we ectopically expressed eGFP-tagged HMGB1 in U2OS cells. Wild-type HMGB1 displayed diffuse nuclear localization in live cells (Fig. 3a and Extended Data Fig. 5a,b). By contrast, mutant HMGB1 localized to discrete nuclear inclusions (Fig. 3a, Extended Data Fig. 5a,b and Supplementary Video 2). Ectopic expression of the mutant HMGB1 IDR also led to the formation of nuclear inclusions in live U2OS cells (Fig. 3a and Extended Data Fig. 5a–c), which indicated that the replaced IDR of the mutant HMGB1 is responsible for its altered subnuclear localization. Nuclear inclusions were observed in several other human cell types expressing mutant HMGB1 (Extended Data Fig. 5d).

Mutant HMGB1 nuclear inclusions frequently contained cavities and resembled nucleoli. Nucleoli are phase-separated multiphasic condensates that contain an outer granular component enriched in NPM1 and an inner dense fibrillar component enriched in FIB1 (ref.[35]) (Fig. 3b). To gain initial insights into the nature of the mutant HMGB1 nuclear inclusions, we expressed FIB1 tagged with red fluorescent protein (RFP–FIB1) and eGFP–HMGB1 in live U2OS cells. The cavities in the mutant HMGB1 inclusions tended to encapsulate FIB1 (Fig. 3c). Fluorescence recovery after photobleaching (FRAP) experiments revealed that the HMGB1 shell displayed arrested dynamics around the FIB1 cores (Fig. 3d and Extended Data Fig. 5e). These results suggest that the mutant HMGB1 inclusions may be abnormal, arrested nucleoli.

To further probe the identity of mutant HMGB1 nuclear inclusions, we performed immunofluorescence against various nuclear proteins known to form condensates. The immunofluorescence analyses revealed that mutant HMGB1 inclusions were distinct from RNAPII and MED1 puncta, nuclear speckles and heterochromatin (Extended Data Fig. 5f–h). However, they overlapped with NPM1 and FIB1 (Extended Data Fig. 5h). The NPM1 signal within the HMGB1 inclusions inversely correlated with the HMGB1 signal (Pearson's $r = -0.70$) (Extended Data Fig. 5i). Moreover, the amount of diffuse NPM1 outside nucleoli correlated with the amount of HMGB1 in the inclusions (Pearson's $r = 0.50$) (Extended Data Fig. 5j). These results indicate that the mutant HMGB1 inclusions replace the NPM1-enriched granular component of nucleoli.

Targeted mutagenesis experiments revealed that arginine residues in the mutant HMGB1 tail drive nucleolar mispartitioning, and a hydrophobic patch drives nucleolar arrest. Various mutant HMGB1 sequences were expressed in live U2OS cells. A HMGB1 protein lacking the entire IDR (Del IDR) or the sequence after the frameshift position (Del FS) was not enriched in the nucleolus (Fig. 3e,f). Deletion of arginine residues (R del), substitution of arginine residues with alanine residues (R>A), substitution of arginine and lysine residues with alanine residues (R&K>A), and substitution of arginine residues with lysine residues (R>K) within the sequence created by the frameshift led to failure of the mutant protein to partition into the nucleolus (Fig. 3e,f). Furthermore, deletion of the short hydrophobic patch at the C terminus of the frameshifted sequence (Patchless) did not alter nucleolar mispartitioning (Fig. 3e,f and Extended Data Fig. 5k), but it did rescue the arrested dynamics of the mutant HMGB1 nucleoli assessed by FRAP (Fig. 3e,g). These results demonstrate that nucleolar mispartitioning of the frameshift mutant HMGB1 depends on arginine residues within the sequence created by the frameshift and that the hydrophobic patch contributes to nucleolar arrest.

Mutant HMGB1 and nucleolar dysfunction

To test whether the nucleolar mispartitioning of HMGB1 affects nucleolar function, we investigated ribosomal RNA (rRNA) production using quantitative PCR with reverse transcription (RT–qPCR)[36]. The level of 28S rRNA in U2OS cells expressing the frameshift mutant HMGB1 was significantly reduced by about 1.5-fold ($P < 0.05$, Student's t-test) (Extended Data Fig. 6a). Ribosomal dysfunction was subsequently probed using an assay of nascent translation that measures puromycin incorporation[37]. U2OS cells expressing mutant eGFP–HMGB1 consistently displayed lower levels of puromycin intensity than non-transfected (that is, GFP$^-$) cells and cells transfected with wild-type eGFP–HMGB1 (Fig. 3h,i and Extended Data Fig. 6b,c). Furthermore, U2OS cells expressing the mutant HMGB1 exhibited substantially reduced viability after several days of culture compared with cells expressing wild-type HMGB1 ($P < 5 \times 10^{-4}$, one-way analysis of variance (ANOVA)) (Fig. 3j and Extended Data Fig. 6d). The reduced viability was associated with nucleolar arrest, as transfection of cells with the Patchless mutant did not compromise viability (Fig. 3j and Extended Data Fig. 6d). The findings of nucleolar mispartitioning, nucleolar arrest and viability were corroborated using cell lines expressing stably integrated eGFP–HMGB1 transgenes from a PiggyBac transposon (Extended Data Fig. 6e–j). These results indicate that the presence of the HMGB1 frameshift mutant in cells disrupts nucleolar function and is cytotoxic.

ACMG classification of HMGB1 variants

The clinical and genetic information of the five individuals with BPTAS and the functional data were used to classify the *HMGB1* frameshift variants as pathogenic. This classification was made in accordance with the criteria of the American College of Medical Genetics and Genomics (ACMG)[38]. Both frameshifts observed in individuals with BPTAS result in the replacement of the highly conserved acidic tail of the protein (ACMG criterion PM1), and were classified as pathogenic by MutationTaster (ACMG criterion PP3). Notably, of the 43 nonsynonymous variants in the HMGB1 tail (1,123 alleles) listed in the gnomAD database (v.2.1.1)[39], only 4 variants (5 alleles) introduce amino acids other than aspartate and glutamate (ACMG criterion PM2) (Extended Data Fig. 2f,g). All previously described pathogenic *HMGB1* variants are associated with neurodevelopmental phenotypes without severe skeletal anomalies, which are therefore distinct from BPTAS (Fig. 1c and Supplementary Table 2), including a chromosomal microdeletion encompassing the *HMGB1* locus in an individual (I6) diagnosed in our study (Extended Data Fig. 2h–j, Supplementary Note and Supplementary Table 2). The functional data presented in this study suggest a deleterious effect (ACMG criterion PS3). In summary, the identification of almost the same (ACMG criterion PS4) *HMGB1* frameshift variants in five (shown to be de novo in four; ACMG criteria PS2 and PM6) unrelated individuals with the same ultrarare diagnosis of BPTAS (ACMG criterion PP4) argues for the classification of these variants as pathogenic (ACMG evidence level: 3S+2M+2P; Supplementary Note).

Catalogue of variants in C-terminal IDRs

We then sought to investigate whether replacement of a disordered C-terminal tail with an arginine-rich basic tail and the consequent nucleolar mispartitioning and dysfunction could occur in other diseases. To this end, we generated a catalogue of genetic variants in intrinsically disordered tails of cellular proteins. First, we annotated 9,303 isoforms of 5,618 genes that have a C-terminal IDR consisting of at least 20 amino acids (Supplementary Table 3). We then identified genetic variants that occur in the disordered tails of the 5,618 genes annotated in the 1000 Genomes Project, ClinVar, COSMIC and dbSNP databases. These analyses revealed 249,464 genetic variants in C-terminal IDRs, including 10,023 truncating variants and 3,888 frameshifts that replace the C-terminal sequence with ≥20 amino acids (Fig. 4a, Extended Data Fig. 7a–c and Supplementary Tables 4 and 5). Of the 3,888 frameshifts, 426 were annotated as pathogenic in ClinVar, 763 were common variants curated in the 1000 Genomes Project and 189 in the dbSNP databases (Fig. 4b, Extended Data Fig. 7b,c and Supplementary Table 5). The frameshifts were associated with higher-than-average pathogenicity (Fig. 4a), and frameshifts were enriched for pathogenic variants (Fig. 4b and Extended Data Fig. 7c). Genes encoding transcription factors were highly enriched among those that contained C-terminal IDR mutations (Extended Data Fig. 7d). Among the 3,888 frameshift variants, 624 were predicted to result in a sequence consisting of at least 15% arginine residues, of which 101 were classified as pathogenic in ClinVar (Fig. 4b,c

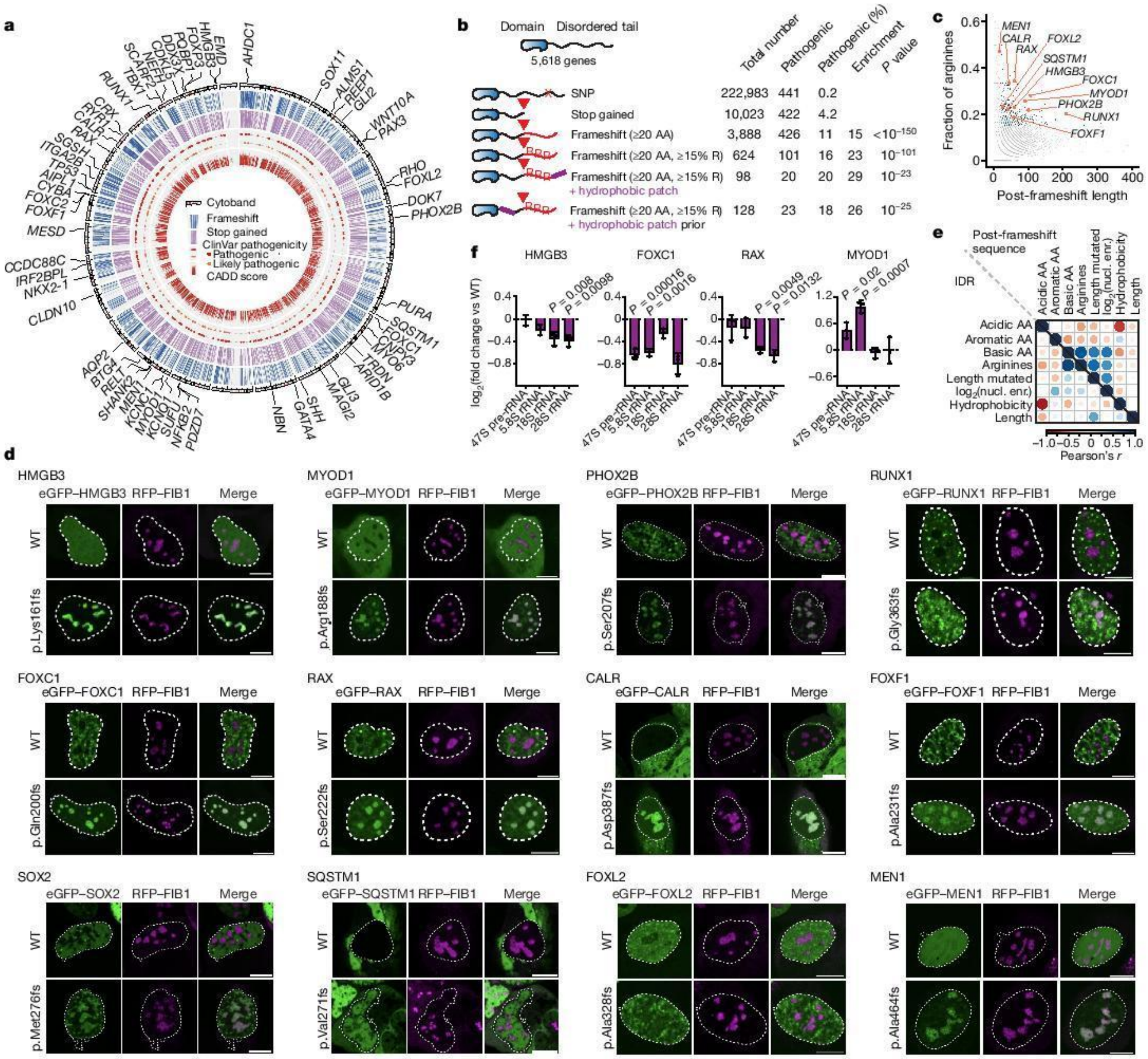


Fig. 4 | A catalogue of variants in C-terminal IDRs reveals frameshifts associated with nucleolar mispartitioning and dysfunction. a, Circos plot of the IDR variant catalogue. The circles indicate the location of genes that contain a truncation (stop gained) or frameshift variant in the dbSNP, 1000 Genomes Project, COSMIC and ClinVar databases. The highlighted genes contain a pathogenic frameshift that creates a sequence of ≥20 amino acids comprising ≥15% arginine residues. **b**, Summary statistics and features of variant types in C-terminal IDRs. *P* values are from hypergeometric tests. **c**, Identification of frameshifts creating a sequence of ≥20 amino acids that consist of ≥15% arginine residues. Plotted is the fraction of arginine residues against the length of the sequence created by the frameshift. The genes containing the variants selected for further validation are highlighted orange. Pathogenic gene variants are in blue. **d**, Representative images of U2OS cells co-expressing RFP–FIB1 and the indicated eGFP-tagged proteins. Nuclear area revealed by Hoechst staining is shown as dashed white lines. Mutations in the following genes are associated with the indicated conditions: microphthalmia (*HMGB3* and *RAX*); myopathy (*MYOD1*); congenital central hypoventilation (*PHOX2B*); myelodysplasia (*RUNX1*); Axenfeld–Rieger syndrome type 3 (*FOXC1*); myelofibrosis (*CALR*); alveolar capillary dysplasia (*FOXF1*); anophthalmia/microphthalmia-oesophaegal atresia syndrome (*SOX2*); Paget disease of bone 2, early-onset frontotemporal dementia and amyotrophic lateral sclerosis (*SQSTM1*); blepharophimosis, ptosis and epicanthus inversus (*FOXL2*); and hereditary cancer predisposing syndrome (*MEN1*). Scale bar, 10 μm. **e**, Nucleolar mispartitioning strongly correlates with the fraction of arginine residues in the frameshift sequence. Plotted are Pearson's correlation coefficients of the extent of nucleolar mispartitioning of mutant proteins with protein features of their IDRs (left triangle) and features of the sequences created by the frameshifts (right triangle). The colour corresponds to the value of Pearson's correlation coefficients, and the size of the circles is proportional to the *P* value of the Pearson's *r*. **f**, RT–qPCR analysis of rRNA species in U2OS cells expressing the indicated WT and mutant proteins. rRNA levels are normalized against an RNAPII transcript (*GAPDH*), and fold changes are calculated against the rRNA/*GAPDH* level measured in the cells expressing WT protein. Data are shown as mean ± s.d., *P* values are from two-tailed Welch's *t*-test. AA, amino acid; SNP, single nucleotide polymorphism; nucl. enr., nucleolar enrichment.

and Supplementary Fig. 2a–g). Overall, 29 out of 66 genes containing arginine-rich frameshift variants had a probability of loss-of-function intolerance (pLI) score of <0.05, which is consistent with a potential gain-of-function effect of the variants (Extended Data Fig. 7e). The variants were associated with various pathogenic conditions, including neurodevelopmental diseases and cancer predisposition (Extended Data Fig. 7f–h). Moreover, 98 of the frameshifts also created a sequence resembling the short hydrophobic patch encoded by the HMGB1 frameshift (Fig. 4b), and 128 of the frameshifts occurred in genes that contained at least one hydrophobic patch in their IDR (Fig. 4b). Overall, the catalogue revealed >200,000 variants in C-terminal IDRs, including 624 frameshifts that replace a C-terminal tail with an arginine-rich basic tail, of which 101 frameshifts were classified as pathogenic.

Genes containing pathogenic frameshift variants that create an arginine-rich basic tail were expressed in U2OS cells. As such frameshifts are highly enriched in genes that encode transcription factors, we selected nine transcription factors (HMGB3, FOXC1, FOXF1, MYOD1, RAX, RUNX1, SOX2, PHOX2B and FOXL2) and four additional proteins (MEN1, SQSTM1, CALR and DVL1) for functional testing (Fig. 4d, Extended Data Figs. 8 and 9a–d and Supplementary Fig. 3a–d). The frameshift mutants of 12 out of the 13 proteins formed nuclear inclusions that overlapped the FIB1–RFP-labelled dense fibrillar component of the nucleolus in live cells (Fig. 4d and Extended Data Fig. 9e–i). The extent of mispartitioning into the nucleolus strongly correlated with the length of the IDR sequence replaced by the frameshift and the fraction of arginine residues in the sequence created by the frameshifts (Fig. 4e and Supplementary Fig. 4a,b). For six variant proteins, cavities enriched in FIB1–RFP were apparent (Fig. 4d and Extended Data Fig. 10a). FRAP experiments showed that condensate properties for 7 out of the 13 variants were affected (Extended Data Figs. 9g and 10b). Six of the mutant proteins that showed significant nucleolar enrichment were further analysed. For four out of six, changes in the level of rRNA species in cells expressing the frameshift mutants were detectable (Fig. 4f and Extended Data Fig. 10c,d). These results indicate that disease-associated frameshifts that generate an arginine-rich basic tail in C-terminal IDRs can cause nucleolar mispartitioning and dysfunction.

Discussion

We propose that disease-associated and common variants in disordered regions may alter phase separation and partitioning of proteins into biomolecular condensates. In particular, the results presented here indicate that frameshift variants that substantially increase the arginine content of various proteins lead to mispartitioning into the nucleolus and disruption of nucleolar function. Our data identified the replacement of the disordered tail with an arginine-rich basic tail in HMGB1 as the pathomechanism underlying BPTAS, a rare complex malformation syndrome[13]. The HMGB1 variant appears to encode a sequence that combines high arginine content, reminiscent of the phase-separation grammar of native nucleolar proteins[40], and a hydrophobic patch that predominantly contributes to nucleolar arrest and dysfunction (Fig. 3e–j). The frameshift therefore interferes with the 'molecular grammar' of phase separation encoded in *HMGB1*, and the resulting mutant protein disrupts condensate features and function of the nucleolus where it accumulates. The extent to which the minimal propensity of the HMGB1 mutant sequence to form a helix contributes to these effects remains to be tested.

We provided evidence that arginine-rich frameshifts occur in hundreds of proteins, which implies that there is a common mechanism for hundreds of disease-associated and common genetic variants with previously unknown functions. The organismal effects of such frameshifts are probably influenced by tissue-specific expression and haplosufficiency (or haploinsufficiency) of the genes in which they occur. For example, BPTAS is associated with a frameshift in *HMGB1* that is broadly expressed and haploinsufficient (pLI score of 0.83), which is consistent with phenotypic features presenting in multiple organ systems and partially overlapping with those seen when the locus is deleted (Supplementary Note). Of note, mispartitioning into the nucleolus and nucleolar dysfunction have been reported for poly-(proline:arginine)-dipeptides produced by repeat-expanded variants of *C9orf72* linked to amyotrophic lateral sclerosis[36,41,42]. Aberrant phase separation and nucleolar dysfunction may therefore occur in a wide range of genetic conditions as a shared underlying molecular pathomechanism.

Finally, the IDR variant catalogue provides a resource for exploring further models of how disease-associated variants may alter biomolecular condensates. For example, the >10,000 variants that truncate a C-terminal IDR may inhibit biogenesis of condensates, and several such variants have been associated with condensate dissolution in cultured cells[43]. Disease-associated alanine repeat expansions in a few transcription factors have been shown to alter the composition of their condensates[6], and our catalogue contains >200 frameshift sequences consisting of at least 25% alanine residues. In summary, we propose that disruption of phase separation may frequently occur in genetic diseases. Further investigation of the underlying molecular basis may lead to future strategies that alter phase separation with therapeutic intent.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05682-1.

1. Claussnitzer, M. et al. A brief history of human disease genetics. *Nature* **577**, 179–189 (2020).
2. Tsang, B., Pritisanac, I., Scherer, S. W., Moses, A. M. & Forman-Kay, J. D. Phase separation as a missing mechanism for interpretation of disease mutations. *Cell* **183**, 1742–1756 (2020).
3. Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol*. **6**, 197–208 (2005).
4. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
5. Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
6. Basu, S. et al. Unblending of transcriptional condensates in human repeat expansion disease. *Cell* **181**, 1062–1079.e30 (2020).
7. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
8. Patel, A. et al. A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 (2015).
9. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **519**, 223–228 (2015).
10. Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. *Nucleic Acids Res.* **43**, D789–D798 (2015).
11. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
12. Kandoth, C. et al. Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
13. Baraitser, M. et al. A syndrome of brachyphalangy, polydactyly and absent tibiae. *Clin. Dysmorphol.* **6**, 111–121 (1997).
14. Faravelli, F., Di Rocco, M., Stella, G., Selicorni, A. & Camera, G. Brachyphalangy, feet polydactyly, absent/hypoplastic tibiae: a further case and review of main diagnostic findings. *Clin. Dysmorphol.* **10**, 101–103 (2001).
15. Pierson, D. M. et al. Total anomalous pulmonary venous connection and a constellation of craniofacial, skeletal, and urogenital anomalies in a newborn and similar features in his 36-year-old father. *Clin. Dysmorphol.* **10**, 95–99 (2001).
16. Olney, R. S. et al. Limb/pelvis hypoplasia/aplasia with skull defect (Schinzel phocomelia): distinctive features and prenatal detection. *Am. J. Med. Genet.* **103**, 295–301 (2001).
17. Wechsler, S. B., Lehoczky, J. A., Hall, J. G. & Innis, J. W. Tibial aplasia, lower extremity mirror image polydactyly, brachyphalangy, craniofacial dysmorphism and genital hypoplasia: further delineation and mutational analysis. *Clin. Dysmorphol*. **13**, 63–69 (2004).
18. Bernardi, P. et al. Additional features in a new case of a girl presenting brachyphalangy, polydactyly and tibial aplasia/hypoplasia. *Am. J. Med. Genet. A* **149A**, 1532–1538 (2009).
19. Shafeghati, Y. et al. Brachyphalangy, polydactyly and tibial aplasia/hypoplasia syndrome (OMIM 609945): case report and review of the literature. *Eur. J. Pediatr.* **169**, 1535–1539 (2010).
20. Itou, J. et al. HMGB factors are required for posterior digit development through integrating signaling pathway activities. *Dev. Dyn.* **240**, 1151–1162 (2011).
21. Yanai, H. et al. HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. *Nature* **462**, 99–103 (2009).

22. Scaffidi, P., Misteli, T. & Bianchi, M. E. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature* **418**, 191–195 (2002).
23. Bianchi, M. E., Beltrame, M. & Paonessa, G. Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**, 1056–1059 (1989).
24. Prasad, R. et al. HMGB1 is a cofactor in mammalian base excision repair. *Mol. Cell* **27**, 829–841 (2007).
25. Sofiadis, K. et al. HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence. *Mol. Syst. Biol.* **17**, e9760 (2021).
26. Elsner, J. et al. Genome sequencing in families with congenital limb malformations. *Hum. Genet.* **140**, 1229–1239 (2021).
27. Lambert, S. A. et al. The human transcription factors. *Cell* **175**, 598–599 (2018).
28. Boehning, M. et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
29. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).
30. Cho, W. K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
31. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
32. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid–liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
33. Asimi, V. et al. Hijacking of transcriptional condensates by endogenous retroviruses. *Nat. Genet.* **54**, 1238–1247 (2022).
34. Larson, A. G. et al. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
35. Feric, M. et al. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016).
36. Kwon, I. et al. Poly-dipeptides encoded by the *C9orf72* repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139–1145 (2014).
37. Aviner, R. The science of puromycin: from studies of ribosome function to applications in biotechnology. *Comput. Struct. Biotechnol. J.* **18**, 1074–1083 (2020).
38. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
39. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
40. Mitrea, D. M. et al. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife* **5**, e13571 (2016).
41. White, M. R. et al. *C9orf72* poly(PR) dipeptide repeats disturb biomolecular phase separation and disrupt nucleolar function. *Mol. Cell* **74**, 713–728.e6 (2019).
42. Lee, K. H. et al. C9orf72 dipeptide repeats impair the assembly, dynamics, and function of membrane-less organelles. *Cell* **167**, 774–788.e17 (2016).
43. Banani, S. F. et al. Genetic variation associated with condensate dysregulation in disease. *Dev. Cell* **57**, 1776–1788.e8 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Institute of Medical Genetics and Human Genetics, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. [2]BIH Biomedical Innovation Academy, Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Berlin, Germany. [3]RG Development and Disease, Max Planck Institute for Molecular Genetics, Berlin, Germany. [4]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany. [5]Exploratory Diagnostic Sciences, Berlin Institute of Health at Charité–Universitätsmedizin Berlin, Berlin, Germany. [6]Institute of Human Genetics, University Hospitals Schleswig-Holstein, University of Lübeck and Kiel University, Lübeck, Kiel, Germany. [7]Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [8]Center for Genomics and Transcriptomics (CeGaT), Tübingen, Germany. [9]Department of Pediatrics and Adolescent Medicine, School of Clinical Medicine, LKS Faculty of Medicine, The University of Hong Kong, Pok Fu Lam, Hong Kong. [10]Department of Congenital Heart Disease and Pediatric Cardiology, University Hospital Schleswig-Holstein, Kiel, Germany. [11]Department of Medical Genetics, Carl von Ossietzky University, Oldenburg, Germany. [12]Department of Obstetrics and Gynecology, University Hospital Schleswig-Holstein, Kiel, Germany. [13]Department of Internal Medicine, Radboud Institute for Molecular Life Sciences, Radboud Expertise Center for Immunodeficiency and Autoinflammation and Radboud Center for Infectious Disease (RCI), Radboud University Medical Center, Nijmegen, The Netherlands. [14]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands. [15]Department of Pediatrics, Pediatric Endocrinology and Diabetes, University Hospital Schleswig-Holstein, Schleswig-Holstein, Germany. [16]Handchirurgie, Katholisches Kinderkrankenhaus Wilhelmstift, Hamburg, Germany. [17]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. [18]Department of Obstetrics and Gynaecology, Queen Mary Hospital, Pok Fu Lam, Hong Kong. [19]Institute of Human Genetics, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany. [20]Institute for Human Genetics and Genomic Medicine, Medical Faculty, RWTH Aachen University Hospital, Aachen, Germany. [21]Department of Pediatric Neurology, Charité–Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. [22]Department of Pediatrics, University Hospital Center Schleswig-Holstein, Kiel, Germany. [23]Department of Radiology and Neuroradiology, Pediatric Radiology, University Hospital Schleswig-Holstein, Kiel, Germany. [24]Institute of Human Genetics, Medical Center, University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [25]Zentrum für Kinder-und Jugendmedizin, Helios Kliniken Schwerin, Schwerin, Germany. [26]Institute of Human Genetics, University Hospital, Otto-von-Guericke University, Magdeburg, Germany. [27]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Spain. [28]Microscopy Core Facility, Max Planck Institute for Molecular Genetics, Berlin, Germany. [29]ICREA, Passeig Lluís Companys 23, Barcelona, Spain. [30]BCRT-Berlin Institute of Health Center for Regenerative Therapies, Berlin, Germany. [31]DZHK (German Centre for Cardiovascular Research), partner site Hamburg, Lübeck, Kiel, Lübeck, Germany. [32]These authors contributed equally: Martin A. Mensah, Henri Niskanen. [33]These authors jointly supervised this work: Malte Spielmann, Denise Horn, Denes Hnisz. ✉e-mail: Malte.Spielmann@uksh.de; Denise.Horn@charite.de; hnisz@molgen.mpg.de

Methods

DNA sequencing, array comparative genomic hybridization and qPCR

Genome sequencing and exome sequencing were performed using Illumina technology with a paired-end sequencing approach[26]. Genome sequencing data were filtered using VarFish. Information on excluded variants and filtering strategy are displayed in Extended Data Fig. 2a. Sanger sequencing and real-time qPCR were performed on a 3730 DNA analyzer (Thermo Fisher Scientific). Sanger sequencing of *HMGB1* from gDNA from individuals included in this study was performed using primers listed in Supplementary Table 6. For cDNA Sanger sequencing and RT–qPCR of I3, RNA was extracted from a patient and a control lymphoblastoid cell line using a Direct-zol RNA Miniprep kit (Zymo Research Europe). RNA was measured on a Nanodrop instrument (Thermo Fisher Scientific), and 1 µg of RNA was transcribed to cDNA using a RevertAid H Minus First Strand cDNA Synthesis kit (Thermo Fisher Scientific). Raw data of RT–qPCRs were analysed using the $2^{(-\Delta\Delta CT)}$ method normalized to *GAPDH*. For cDNA Sanger sequencing, the primers used for amplification and sequencing are listed in Supplementary Table 6. For RT–qPCR of cDNA from individuals included in this study, *HMGB1* and *GAPDH* primers are listed in Supplementary Table 6. Chromosomal microarray analysis was performed using a 4 × 180 k oligonucleotide slide from Agilent on a DNA microarray scanner (Agilent). Chromosomal microarray analysis results were confirmed by RT–qPCR. All procedures were performed using the manufacturers' protocols. All variants were annotated according to genome build hg19 and the *HMGB1* transcript NM_002128.7.

Patient consent

Parental consent was obtained for all clinical and molecular studies of this article and for the publication of the relevant causative variants and of clinical photographs. Patient consent did not cover the release of personal sequence information other than the causative pathogenic variants. Therefore, whole-genome sequencing and exome sequencing data cannot be made publicly available. All studies and investigations were performed according to the declaration of Helsinki principles of medical research involving human participants, and the study was approved by the ethics committee of the Charité–Universitätsmedizin Berlin (EA2/087/15).

Patient recruitment and clinical protocol

Individuals were recruited during routine patient care at five departments of genetics (Berlin, Kiel, Nuremberg, Schwerin, Hong Kong). Fetuses from spontaneous abortions were not systematically screened for BPTAS. No statistical methods were used to predetermine sample sizes. Investigators were not blinded and no randomization was used.

Computer-aided facial phenotyping

Facial frontal images were analysed using the Face2Gene suite (v.20.1.4, https://www.face2gene.com). Face2Gene Clinic was used for computer-aided facial phenotyping[44]. We created a composite mask using Face2Gene Research. If several images of the same patient were available, the image depicting the individual at the oldest age was used for facial analysis by Face2Gene Clinic. Seven images of unrelated individuals diagnosed with BPTAS were taken from the literature (of those reported in ref. [15], only the father was included)[13–19]. In addition, I1 and I2 of the current study were included in the analysis. Each selected BPTAS image was used twice for Face2Gene Research analysis to reach more than the ten images necessary for composite mask creation (Extended Data Fig. 1s).

AlphaFold predictions for protein structures

AlphaFold predictions were computed using an in-house implementation of AlphaFold[45] using v.2.0.0 from 16 July 2021. The preset parameter was set to --preset=casp14 to use all genetic databases and eight ensembles, matching the CASP14 prediction pipeline. Templates were restricted to those available before the CASP14 predictions using the parameter --max_template_date=2020-05-14. Models were rendered using UCSF ChimeraX (v.1.5)[46,47], colouring the structure with the pLDDT score. Multiple sequence analysis depth plots and per-model pLDDT sequence plots were made using custom scripts based on ColabFold notebook AlphaFold2 with MMseqs2 (ref. [48]). Predictions of *Mus musculus*, *Rattus norvegicus* and *Danio rerio* HMGB1(A) protein structures, shown in Extended Data Fig. 4a, are from the AlphaFold Protein Structure Database[45].

Generation of DNA constructs for protein purification and expression in human cells

To generate plasmids for recombinant protein expression, *HMGB1* cDNA sequences containing the wild-type or NM_002128.7(*HMGB1*): c.551_554delAGAA;p.(Lys184Argfs*44) variant were ordered from Twist Bioscience. Full-length cDNAs and the regions encoding IDR sequences were cloned into a monomeric eGFP (meGFP)-pET45 backbone by Gibson assembly using NEBuilder HiFi DNA Assembly MasterMix (NEB); primers are listed in Supplementary Table 6. For the generation of pET45-mCherry–NPM1 and pET45-mCherry–HP1a, *NPM1* and *HP1A* open-reading frames were amplified from mouse cDNA using primers flanked with Gibson overhangs (sequences listed in Supplementary Table 6). The resulting amplicons were gel purified and cloned into pET45-mCherry (Addgene, 145279) linearized with AscI and HindIII restriction enzymes. For the generation of pET28-mCherry–MED1-IDR, mCherry was subcloned into the pET28-meGFP–MED1-IDR vector as previously described[6,31] using NcoI and BsrGI restriction sites.

To express monomeric eGFP–HMGB1 variants in mammalian cells, eGFP–HMGB1 sequences were subcloned from pET45-meGFP vectors into a pRK5-meGFP vector digested with AgeI and XbaI (Addgene, 18696); primers used are listed below. To express wild-type and frameshift variants of FOXC1, FOXF1, HMGB3, MYOD1, RAX, RUNX1, PHOX2B, CALR, SOX2, SQSTM1, FOXL2, MEN1 and DVL1, the following cDNA sequences were ordered from Twist Bioscience: NM_001453.3(*FOXC1*):c.599_617del;p.(Gln200Argfs*109), variant rs1057519478; NM_001451.3(*FOXF1*):c.691_698del;p.(Ala231Argfs*61), variant 692054; NM_005342.4(*HMGB3*):c.480_481dup;p.(Lys161Ilefs*55), variant rs431825172; NM_002478.5(*MYOD1*):c.557dup;p.(Arg188Profs*90), variant rs1179926739; NM_013435.3(*RAX*):c.664del; p(Ser222Argfs*63), variant rs1603388837; NM_001754.5(*RUNX1*): c.1088_1094del;p.(Gly363Alafs*229), variant 1013621; NM_004343.4 (*CALR*):c.1157_1158dup;p.(Asp387Argfs*44), variant COSV104394382; NM_003924.4(*PHOX2B*):c.618del;p.(Ser207Alafs*102), variant 658418; NM_023067.4(*FOXL2*):c.982del;p.(Ala328Profs*28), variant 369937; NM_003106.4(*SOX2*):c.828del;p(Met276Ilefs*95) variant 986766; NM_003900.5(*SQSTM1*):c.810del;p.(Val271Serfs*41) variant 967349; NM_001370259.2(*MEN1*):c.1382_1389dup;p.(Ala464Argfs*98) variant 428075; NM_004421.2(*DVL1*):c.1505_1517del;p.(His502Profs*143). For genotype–phenotype correlations see Supplementary Note.

cDNAs were amplified with primers listed in Supplementary Table 6 and cloned into a pRK5-meGFP–HMGB1 vector using Gibson assembly after removing the *HMGB1* sequence with BsrGI and XbaI restriction enzymes. To test the contribution of arginine and lysine residues of the mutant HMGB1 sequence, cDNA sequences were ordered from Twist Bioscience, in which all arginine and lysine residues after Lys185 were replaced with alanine (R&K>A variant), all arginine residues after Lys185 were deleted (R del variant) or replaced with alanine or lysine (R>A and R>K, respectively, variants). cDNAs were amplified using the primers listed below and cloned into a pRK5-meGFP–HMGB1 vector as described above. To create truncated versions of HMGB1, in which the IDR (amino acids after Asn134), or the sequence after the frameshift position (del FS) or the hydrophobic patch of the mutant sequence (amino acids after Lys209) is deleted, cDNA was amplified from pRK5-meGFP-HMGB1 using the primers listed in Supplementary Table 6 and cloned back to a vector digested with BsrGI and XbaI as described above. All constructs were sequence-verified. Plasmids are available from Addgene (https://www.addgene.org/Denes_Hnisz/).

Protein purification and peptide synthesis

Protein expression of mCherry constructs was performed as previously described[6,33], but with modifications to mCherry–MED1-IDR expression, which was performed in the presence of 400 μg ml^{-1} kanamycin. Protein expression of meGFP–HMGB1 constructs was performed in Rosetta (DE3)pLysS cells (Sigma-Aldrich) in the presence of 25 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} ampicillin. All bacterial pellets were stored at −80 °C. Pellets were resuspended in 20 ml of ice-cold buffer A (50 mM Tris pH 7.5, 500 mM NaCl, 20 mM imidazole and complete protease inhibitors (Sigma-Aldrich, 11697498001)), and cells were lysed using a Qsonica Q700 sonicator. Lysate was cleared by centrifugation at 15,500*g* for 30 min at 4 °C, and proteins were purified using an Äkta avant 25 chromatography system and a complete His-Tag purification column (Merck, 6781543001). Columns were pre-equilibrated in buffer A, loaded with cleared lysate and washed with 15 column volumes of buffer A. Fusion proteins were eluted in 10 column volumes of elution buffer (50 mM Tris pH 7.5, 500 mM NaCl and 250 mM imidazole). Protein preparations were diluted in storage buffer (50 mM Tris pH 7.5, 125 mM NaCl, 1 mM DTT and 10% glycerol) and concentrated using 3000 MWCO Amicon Ultra centrifugal filters (Merck, UFC803024) and stored at −80 °C. After His-Tag column purification, meGFP–HMGB1 protein preparations were further purified using Superdex 200 10/300 GL columns (GE28-9909-44) and concentrated and stored as noted above. Elution profiles are shown in Extended Data Fig. 4a. We note that the mutant protein elutes at lower elution volumes, which indicates that it may form soluble oligomers and that the potential to form soluble oligomers may be associated with the slight propensity of the mutant IDR to form a helix (Extended Data Fig. 3c,d). Immunoreactivity of purified meGFP–HMGB1 proteins were evaluated by western blotting. Equal amounts of protein were diluted in NuPAGE LDS buffer (Thermo Fisher Scientific, NP0007) with NuPAGE sample-reducing agent and heated at 70 °C for 10 min. Samples were run using NuPAGE 4–12% Bis-Tris protein gels (Invitrogen, NP0321PK2) and transferred to a nitrocellulose membrane with an iBlot2 device. The membrane was blocked with 5% non-fat milk TBST for 1 h and incubated 1 h with anti-HMGB1 (Sigma-Aldrich, H9664) or anti-eGFP (Invitrogen, A-11122) antibodies diluted 1:1,000 in 5% non-fat milk TBST. Membranes were washed five times with TBST, incubated with HRP-conjugated donkey anti-rabbit antibody (1:2,000, Jackson Immuno Research, 711-035-152) for 1 h, washed five times in TBST and visualized using SuperSignal West Dura Extended Duration substrate (Thermo Scientific, 34075). The identity of the fusion protein products was confirmed by mass spectrometry.

Synthetic peptides with amino-terminal 5′ FAM-labelling for in vitro droplet formation assays (Fig. 2i and Extended Data Fig. 4d–i) and circular dichroism (CD) spectroscopy experiments (Extended Data Fig. 3c,d) were ordered for wild-type and mutant HMGB1 C-terminal sequences (Asp135 onwards) from ProteoGenix. The synthetic peptides had >90% purity.

CD experiments

The synthetic peptides were dissolved in 20 mM sodium phosphate buffer, pH 7.4. The samples were centrifuged for 10 min at 15,000 r.p.m. to remove undissolved solid. The supernatant was extensively dialysed against 20 mM sodium phosphate buffer, pH 7.4, to remove traces of impurities from peptide synthesis. The protein concentration was determined by amino acid analysis. CD spectra were acquired on 10.6 μM samples in a Jasco 815 UV spectrophotopolarimeter at 278 K with a 1 mm optical path cuvette. Each spectrum is the result of 20 cumulative scans acquired at a scanning speed of 50 nm min^{-1} with a data pitch of 0.2 nm (Extended Data Fig. 4c,d).

Reference CD spectra in Extended Data Fig. 4e are included from the Protein Circular Dichroism Data Bank[49]. The following reference proteins were used: myoglobin (blue)[50], with a DSSP α-helix of 73.9%; outer membrane protein g (OmpG, purple)[51], with a DSSP β-strand of 67.6%; and translocated actin recruiting phosphoprotein (Tarp, green)[52], with a DSSP loop of 71.0%.

In vitro droplet formation experiments

For droplet formation experiments in Fig. 2c–e, proteins were diluted to desired concentrations in storage buffer, further diluted 1:1 in 20% PEG-8000 and mixed well with pipetting. Next, 10 μl of solution was immediately transferred on a chambered coverslip (Ibidi, 80826-96). Droplets were imaged using a LSM880 confocal microscope (Zeiss) with a ×63, 1.40 oil DIC objective. Images were acquired slightly above the solution interface; for FRAP experiments, images were acquired directly on the solution interface. Time series for FRAP experiments were acquired using 60 cycles of 2 s intervals, during which the eGFP signal was bleached using a 488 nm laser with 95% intensity after the second interval. FRAP was performed for at least ten droplets for both wild-type and mutant HMGB1 using 10 μM concentration. Recovery curves were fitted to a power-law model. For droplet assays using preassembled mCherry–HP1α, mCherry–MED1-IDR and mCherry–NPM1 condensates (Fig. 2g–i), mCherry-labelled proteins were diluted to 20 μM concentration in storage buffer, diluted 1:1 in 20% PEG-8000 and droplets were allowed to form for 1 h at room temperature, shielded from light. Next, eGFP–HMGB1 proteins or 5′ FAM-labelled synthetic IDR peptides were added to the desired concentration, thoroughly mixed and solutions were left to equilibrate for 45 min at room temperature, shielded from light. Droplets were imaged as described above. To test the contribution of RNA for the condensation propensity of HMGB1 IDR peptides, total RNA from V6.5 mouse embryonic stem cells was isolated using a Direct-zol RNA Miniprep kit and added in indicated concentrations into peptide dilutions. RNA–peptide dilutions were thoroughly mixed with pipetting, crowding agent was added and imaging was performed as described above.

Cell culture

U2OS, HCT116 and HEK293T cells were cultured in DMEM with GlutaMAX (Thermo Fisher Scientific, 31966-021) supplemented with 10% FBS and 100 U ml^{-1} penicillin–streptomycin (Gibco). MCF7 cells were cultured in RPMI-1640 supplemented with 20% FBS and 100 U ml^{-1} penicillin–streptomycin (Gibco). Human induced pluripotent stem (iPS) cells ZIP13K2 (ref. [53]), were grown in mTeSR Plus (Stem Cell Technologies, 100-0276) on plates coated with 1:100 diluted Matrigel (Corning, 354234) in KnockOut DMEM (Thermo Fisher Scientific, 10829-018) and supplemented with 10 μM of the Rho kinase inhibitor Y-27632 (Abcam, ab120129) once detached during passaging. Cells were cultured at 37 °C with 5% CO_2 in a humidified incubator. All cell lines were tested negative for mycoplasma contamination. For live-cell imaging and immunofluorescence, cells were seeded on chambered coverslips (Ibidi, 80826-96). On the next day, cells were transfected using FuGENE HD (Promega) according to the manufacturer's instructions. Human iPS cells were transfected using Lipofectamine 3000 according to the manufacturer's instructions. For viability experiments, cells were cultured on 6-well plates. Transfection series were repeated at least twice for each experiment.

RT–qPCR after expression of frameshift variants in U2OS cells

Cells were grown on 6-well plates, transfected with FuGENE HD according to manufacturer's instructions, and eGFP$^+$ cells were sorted by FACS 48 h after transfections and lysed in TRIzol reagent (Thermo Fisher Scientific). Experiments were performed in at least three biological replicates. RNA was extracted and cDNA synthesis was performed as described above, except that 125 ng of RNA was used. Primers are listed in Supplementary Table 6.

Live-cell imaging

Cells were imaged 24 h after transfections using a LSM880 confocal microscope (Zeiss) equipped with an incubation chamber with 5% CO_2 and a heated stage at 37 °C. Images were acquired using a ×63, 1.40 oil DIC objective. To visualize cell nuclei, cells were incubated with 0.2 μg ml^{-1} Hoechst (Thermo Scientific, 33342) at least 10 min before imaging. To visualize nucleoli in living cells, we expressed RFP–fibrillarin fusion proteins by transfecting cells with pTagRFP-C1-fibrillarin plasmid (Addgene, 70649) together with plasmids for eGFP–HMGB1 and other transcription factor variants.

FRAP experiments were performed for nucleolar regions in cells expressing wild-type or mutant eGFP–HMGB1, guided by the RFP–fibrillarin fluorescence channel. Time series for FRAP experiments were acquired using 20 cycles of 2 s intervals, during which the eGFP signal was bleached using a 488 nm laser with 85% intensity after the second interval. FRAP experiments with designed variants of HMGB1 and other frameshift variants were performed as described above, but using 85–100% laser intensities for bleaching with identical settings for each wild type–mutant comparison. Fluorescence intensities were acquired from around ten regions of interest from separate nuclei, quantified using ZEN Black 2.3 software and reported as relative values to the pre-bleaching time point.

Time-lapse imaging of mutant HMGB1 expressing U2OS cells was performed on a Screenstar microplate (Greiner bio-one, 655866) with Zeiss Celldiscoverer 7. Images were acquired fully automated with a Plan-ApoChromat ×20 objective, NA = 0.7 and 1× tubelense (Optovar) using 15 min intervals and a camera binning of 1 × 1 pixel in 8-bit mode (Supplementary Video 2).

Immunofluorescence

For fixed-cell immunofluorescence, cells were fixed 24 h after transfections with 4% PFA in PBS for 10 min. After two washes with PBS, cells were permeabilized by incubating 30 min with 0.5% Triton X-100 at room temperature, washed three times with PBS and blocked for 1 h with blocking buffer (1% BSA, 0.1% Triton X-100 in PBS) at room temperature. Samples were incubated with primary antibodies diluted in blocking buffer (1:500 rabbit anti-HP1α, Cell Signaling, 2616S; 1:500 rabbit anti-MED1, Abcam, ab64965; 1:500 rabbit anti-RNAPII, ab26721; 1:250 mouse anti-NPM1, Thermo Fisher Scientific, 32–5200; 1:100 mouse anti-FIB1, Santa Cruz, sc-374022; 1:200 mouse anti-SC35, Sigma-Aldrich, S4045) overnight in 4 °C with gentle agitation. After four washes with blocking buffer, samples were incubated with secondary antibodies (1:1,000 dilutions of Alexa Fluor 647 donkey anti-mouse or anti-rabbit antibodies, Jackson Immuno Research, 715-605-150 and 711-605-152) for 1 h at room temperature. Samples were washed two times with blocking buffer, incubated for 3 min with 0.25 μg ml^{-1} DAPI (Invitrogen, D1306) in PBS and washed five times with PBS.

Protein synthesis labelling by puromycylation

U2OS cells were seeded on 24-well plates (15,000 cells per well) on sterilized 13 mm glass coverslips pretreated with 0.2% gelatin. The next day, cells were transfected with meGFP–HMGB1 full-length wild-type or mutant constructs using FuGENE HD according to the manufacturer's instructions. After 24 h, pulse labelling of nascent peptide chains actively translated by the ribosome was performed by replacing the medium supplemented with 20 μM puromycin (Sigma Aldrich, P8833) for 15 min at 37 °C, 5% CO_2. Cells were then washed three times with cold PBS, followed by fixation with 4% formaldehyde (Roth, P087.5) at room temperature, with shaking, for 20 min. Fixative was removed, and cells were washed two times with PBS, followed by incubation in blocking solution (1× PBS, 5% v/v normal donkey serum, 1% w/v BSA, 0.1% w/v glycine and lysine) with shaking for 45 min at room temperature. Anti-puromycin (1:1,000, mouse, Sigma Aldrich, MABE343, RRID:AB_2566826) and anti-GFP (1:2,000, chicken, Abcam, ab13970, RRID:AB_300798) primary antibodies were applied in blocking solution supplemented with 0.4% Triton-X-100 and incubated overnight with shaking at 4 °C. Cells were then washed three times with PBS for 5 min at room temperature, followed by secondary antibodies (1:250, Jackson ImmunoResearch, 488-anti-chicken, 703-545-155, RRID:AB_2340375; 647-anti-mouse, 715-605-151, RRID:AB_2340863) incubated in blocking solution with 0.4% Triton-X-100 shaking for 2 h at room temperature. After three PBS washes, cells were incubated in DAPI (1:2,500) in PBS for 30 min with shaking at room temperature, and washed with PBS an additional two times. Coverslips were removed from wells and sealed on poly-L-lysine slides (Thermo, J2800AMNZ) with ProLong Gold Antifade Mountant (Invitrogen, P36930). The experiment was performed in independent biological triplicates, with two to four technical replicate coverslips per conditions per experiment.

Coverslips were imaged using a Zeiss Celldiscoverer 7 running Zen Blue v.3.2 (Zeiss). All images were acquired in a fully automated fashion with a Plan-ApoChromat ×20 objective, NA = 0.95 and a ×2 tube lens (Optovar), and camera binning 2 × 2 pixels in 8-bit mode. The resulting lateral resolution (xy) is 0.227 μm $pixel^{-1}$. All images were acquired in tile regions of typically 20 × 20 individual tiles, resulting in 400 individual images per coverslip. Focus stabilization was achieved with an automated combined hardware and software focusing strategy at each second position (Fig. 3h,i and Extended Data Fig. 6b,c).

Viability experiments

For viability experiments, cells were collected 24 h after transfections or doxycycline inductions and sorted for $eGFP^+$ cells using a FACS Aria II flow cytometer (BD Biosciences) with BD FACS Diva v.6.1.3. software. The FACS gating strategy is shown in Supplementary Fig. 6. One thousand cells per well were seeded on white microwell plates and were cultured for an additional 48 h. Viability was measured using a CellTiter-Glo 2.0 Cell Viability assay (Promega, G9242) according to the manufacturer's instructions. Measurements were done in three to five technical replicate wells and performed in four to five independent biological replicates. For imaging cells at the end of viability assay, 40,000 sorted cells were seeded per well on 24-well plates and imaged 48 h later with a Nikon Eclipse Ti2 microscope with a ×10 objective.

Generation of doxycycline-inducible meGFP–HMGB1 transgenic cell lines

A PiggyBAC transposon system was used to integrate meGFP–HMGB1 wild-type and mutant sequences into U2OS cells. To generate the doxycycline-inducible expression cassette, *meGFP–HMGB1* cDNA was amplified from pRK5-meGFP–HMGB1 plasmids (primers listed in Supplementary Table 6), and Gibson assembly cloned into the backbone of a Caspex expression vector (Addgene, 97421) digested with NcoI and BsrGI restriction enzymes. Generated plasmids were transfected with a PiggyBAC transposase expression vector (SBI, PB210PA-1) into U2OS cells with FuGENE HD reagent according to the manufacturer's instructions using a molar ratio of 6:1 with meGFP–HMGB1 and transposase expression plasmids. Transfected cells were kept under puromycin (2 μg ml^{-1}) selection for 4 days, after which all untransfected control cells had died. Bulk populations of surviving cells were induced by adding 2 μg ml^{-1} doxycycline (Sigma) and imaged 24 h after doxycycline treatments (Extended Data Fig. 6e–j). GFP^+ cells were sorted by FACS for viability experiments, which were performed as described above. Single-cell clones of meGFP–HMGB1 mutant-expressing U2OS cells was used for time-lapse imaging (Supplementary Video 2).

Image analysis

For the detection of droplet regions for phase diagrams, we used the ZEN blue 3.2 Image Analysis and Intellesis software packages to analyse at least five images for each experimental condition. Image segmentation was performed using the Intellesis Trainable segmentation algorithm, which was trained on five representative images from the image series to classify each pixel into the droplet area and image background. Regions of interest were automatically detected for the entire image series, and mean signal intensities for the eGFP or 5′ FAM channel and object areas for droplets and background are reported. In Fig. 2d, the phase-shifted fraction was calculated as the total area of detected droplets divided by the total area.

Data for dual-colour in vitro condensation experiments were acquired from 15–20 image fields for each condition (corresponding to Fig. 2g–i and Extended Data Fig. 4i,j) using ZEN Blue 3.2. For Extended Data Fig. 4j,k, droplets were first detected using triangle thresholding for light regions in the meGFP or 5′ FAM channel. For data analyses in Fig. 2h,i, droplets were detected using Otsu thresholding for light regions in the mCherry channel. Mean fluorescence intensity within droplet regions, area and diameter were then measured on both channels and plotted as described.

To quantify nuclear enrichment of eGFP–HMGB1, Hoechst stain was used to identify nuclei as the regions of interest using the ZEN Blue 3.2 zones of influence method. Images were automatically segmented with Otsu thresholding, parameters of which were adjusted on the basis of five representative images from the image series. The cytoplasmic region was defined as a ring surrounding the nucleus with a distance of 9 and a width of 29 pixels. Mean and standard deviation values for eGFP fluorescence intensity were recorded for nuclear and cytoplasmic regions, and nuclear enrichment, calculated as a ratio between the two, was plotted in Extended Data Fig. 5a. Cells with no expression (eGFP fluorescence intensity below 5) were excluded from the analysis.

To quantify the correlation between eGFP–HMGB1 fluorescence and NPM1 staining intensities inside and outside nucleoli, images from around 120 cells per condition were analysed using ZEN Blue 3.2 software. Images were first segmented to nuclear regions of interest with Otsu thresholding on the basis of DAPI channel intensity. Nuclei were further segmented to nucleolar regions of interest and regions outside the nucleoli, based on NPM1 staining intensity, using fixed thresholds that detected nucleoli in cells with high and low NPM1 intensities. Parameters were empirically set with ten representative images for each experimental set. Mean signal intensities for eGFP and NPM1 staining were recorded for each region of interest and reported as an average for each detected nucleus.

To quantify nucleolar enrichment of wild-type and frameshift variant proteins (Extended Data Fig. 9e), nuclear regions of interest were defined with Hoechst staining as outlined above and nucleolar regions with RFP–FIB1 intensity using two fixed thresholds that detect nucleoli in cells with high and low RFP–FIB1 expression. Mean signal intensities for eGFP were recorded, and nucleolar enrichment was plotted as $\log_2$(mean signal intensity for regions within nucleoli/mean intensity outside nucleoli). When imaging human iPS cells, nuclear regions of interest were eroded by 8 pixels to avoid signals at the nuclear periphery.

Data wrangling was performed in base R, and plots were generated using the ggplot2 package.

Image analysis for puromycylation experiments was performed using Zen Blue software v.3.4. DAPI was used to localize each cell. In brief, DAPI images were smoothed, an Otsu threshold was applied to binarize images and watershedding was used to separate neighbouring objects. The resulting nuclei masks were filtered to fit an area of 75–900 μm^2 and a circularity (sqrt(4 × area/π × FeretMax^2)) of 0.6–1. The resulting primary objects were dilated with a total of 17 pixels, 3.9 μm. Puromycin and GFP signal intensities were quantified per cell. Puromycin intensity in each GFP$^+$ cell was normalized by the mean puromycin intensity in GFP$^-$ cells in the same image, for wild-type and mutant conditions, and plotted using R and GraphPad Prism, followed by comparisons for significant differences (one-way ANOVA) between condition means from biological replicates. A total of 37,979 single cells for mutant and 39,528 for wild-type conditions were identified and analysed (Fig. 3h,i and Extended Data Fig. 6b,c).

C-terminal IDR identification

Prediction of IDRs was performed using metapredict (v.1.51)[54], a deep-learning-based predictor for consensus disordered sequences. The threshold score was set to 0.5, the minimum IDR length was set to 20 amino acids and the analysis was restricted to only GENCODE canonical or GENCODE basic isoforms. To complete the IDR catalogue, sequences from MobiDB[55] were added to the database. Protein coordinates for each IDR and Interpro domain were used to define the C-terminal IDR. Using a combination of custom scripts, the C-terminal IDR of each isoform was defined as any IDR that started 20 amino acids downstream of the start of the protein, to filter all disordered proteins. The region where the start of the IDR was downstream of the start of the most C-terminal domain was mapped.

Variant identification and characterization

The resulting C-terminal IDR coordinates were then converted to genomic coordinates using the R package ensembldb[56] and the ensembl v.104 human annotation (v.2.22.0). The annotation version can affect the canonical isoforms that are selected for analysis, so the downstream analysis was locked to this version on Ensembl annotation. The resulting BED file was then used to filter ClinVar[57], COSMIC[58], dbSNP[59] and 1000 Genomes[60] to the designated genomic coordinates of the C-terminal IDR regions using BEDtools (v.2.30.0.)[61]. The resulting VCF file was filtered for protein-coding variant consequences using Ensembl Variant Effect Predictor (VEP, v.104)). The filtered VCF was then used to conduct downstream analysis using OpenCRAVAT[62] to annotate the variants for ClinVar annotation using the ClinVar and ClinGen[63] plugins, genomic frequencies using the 1000 Genome plugin, and CADD score[64,65] using the CADD plugin (v.1.6). The CADD score is a metric for the predicted effect of the variant on protein function (Fig. 4a). The same VCF file was also used to retrieve frameshift variant sequences using the Frameshift VEP plugin from pVACtools (v.3.1.0.)[66] and Downstream plugin for the stop gained sequences.

Sequences were then characterized using a combination of custom scripts to obtain protein sequence feature parameters based on localCIDER (v.0.1.18.)[67] and biopython (v.1.79.)[68] packages. All scatter and violin plots were made using the R package ggplot2. The fraction of amino acids was defined as the sum of the count of amino acids over the sequence length. The acidic fraction was defined as the sum of aspartic acid and glutamic acid. The basic fraction was defined as the sum of arginine, lysine and histidine. The RK fraction was defined as the sum of arginine, lysine, and the aromatic fraction as phenylalanine, tyrosine and tryptophan. Hydrophobic patches were identified using custom regex expression (r'([CAVILMFYW]..?)<6,>') using hydrophobic amino acids as the dictionary, allowing 1 or 2 amino acid gap and 6 residue minimum match. Nucleolar signal prediction was caried using NoD program (v.1.0.0.) with the command line with default settings[69]. Characterization of nonsense-mediated mRNA decay of variants was done using a custom script. In brief, wild-type exon boundaries were retrieved from GENECODE and mapped to the wild-type coding sequence. An NMD sensitive zone was established for each wild-type sequence with the following rules: >100 bp downstream of starting codon and <51 bp of the second to last exon boundary. Variants with only one exon were marked 'NMD_escaping', then the stop codon coordinate of the variant was compared with the NMD sensitive zone coordinates and variants of which the stop codon did not overlap with the NMD sensitive zone were also marked as 'NMD_escaping'. All other variants were left empty.

Combined disordered and pLDDT score plots were plotted with the metapredict meta.graph_disorder function and pLDDT_scores parameter set to 'true', using v.2 of the metapredict network and v.7 of the pLDDT score prediction network.

Circos visualization of the variant catalogue was done using Circos implementation in R, and Granges package in R (Fig. 4a).

Enrichment analysis of pathogenic variants was done using hypergeometric nonaccumulative test with *N* set as the full number of variants in the catalogue and *M* set as the full set of pathogenic variants ($N = 249{,}468$ and $M = 1{,}805$). Reported *P* values correspond to the calculated hypergeometric *P* value and fold change as the number of pathogenic variants/expected number of pathogenic variants (Fig. 4b).

Sequence feature correlation matrices in Fig. 4e and Supplementary Fig. 4 were calculated using the cor package in R using Pearson parametric correlation test and plotted using the corrplot package in R. The *P* value cut-off was set to 0.01. The fraction of mutated IDRs was defined as 1 − (frameshift position − IDR start)/IDR length. The *SQSTM1* wild-type sequence was excluded from correlation analysis because the wild-type isoform ENST00000510187.5 in our catalogue was replaced with isoform ENST00000389805.9 (NM_003900.5) in the imaging experiments owing to low transcript support level (TSL:5) for ENST00000510187.

Gene Ontology enrichment analysis (Extended Data Fig. 7d) for the variant type 'stop gained', 'frameshift' and 'ARG-rich FS' was done using gProfiler[70]. Multiple testing correction for *P* values was done using the g:SCS method from g:Profiler.

Scores for the predicted disorder plotted in Fig. 2a and Extended Data Fig. 9a,c were obtained using PONDR (http://www.pondr.com). Charge plots in Fig. 1f and Extended Data Fig. 9b,d were prepared using EMBOSS Charge tool (https://www.bioinformatics.nl/cgi-bin/emboss/charge) with a window size of 8. Isoelectric points (pI) for post-frameshift sequences were calculated using Expasy compute pI tool (https://web.expasy.org/compute_pi/).

The DVL1 variant NM_004421.2(*DVL1*):c.1505_1517del was not part of the catalogue because the frameshift sequence from the canonical isoform used in ensembl v.104 did not fulfil all selection criteria. Instead, this variant was identified through a literature search that revealed Robinow syndrome-associated frameshift variants in the *DVL1* gene that occur in a C-terminal IDR that generates arginine-rich sequences[71,72].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

CD spectra have been deposited at the Protein Circular Dichroism Data Bank under the accession identifiers CD0006401000, CD0006401001, CD0006404000, CD0006404001. Genome and exome-wide summary statistics of I1, I4 and I5, and direct sequencing results of HMGB1 of I1–I4 and array comparative genomic hybridization and qPCR results of the *HMGB1* locus of I6 are made available in this manuscript (Extended Data Fig. 2). Patient consent did not cover the public release of personal sequence information other than the causative pathogenic variants. Therefore, the pathogenic variants are disclosed in this article, but individual-level whole-genome sequencing (WGS) and exome sequencing data cannot be made publicly available for reasons of data protection and patient privacy and are available only upon reasonable request from the corresponding authors. Access to individual-level sequencing data is subject to the policies and approval of the data protection officer of the institution that stores the patient data. WGS and Sanger sequencing data of I1 are stored at the Institute of Human Genetics, University Hospitals Schleswig-Holstein. WGS data of I4 is stored at the Center for Genomics and Transcriptomics (CeGaT) Tübingen. WGS data of I5, and Sanger Sequencing data of I2–I5, and qPCR and array comparative genomic hybridization data of I6 are stored at the Institute of Medical Genetics and Human Genetics, Charité–Universitätsmedizin Berlin. The respective servers are physically located in Germany.

Code availability

Custom code is available at GitHub: https://github.com/hniszlab/HMGB1_2022; https://github.com/alexpmagalhaes/IDR-variant-catalog. Custom code and raw data for this study have been deposited at Zenodo (https://doi.org/10.5281/zenodo.7311150).

44. Gurovich, Y. et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nat. Med.* **25**, 60–64 (2019).
45. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
46. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
47. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
48. Mirdita, M. et al. ColabFold—making protein folding accessible to all. *Nat Methods* **19**, 679–682 (2022).
49. Whitmore, L., Miles, A. J., Mavridis, L., Janes, R. W. & Wallace, B. A. PCDDB: new developments at the Protein Circular Dichroism Data Bank. *Nucleic Acids Res.* **45**, D303–D307 (2017).
50. Lees, J. G., Miles, A. J., Wien, F. & Wallace, B. A. A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics* **22**, 1955–1962 (2006).
51. Abdul-Gader, A., Miles, A. J. & Wallace, B. A. A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics* **27**, 1630–1636 (2011).
52. Tolchard, J. et al. The intrinsically disordered Tarp protein from chlamydia binds actin with a partially preformed helix. *Sci. Rep.* **8**, 1960 (2018).
53. Tandon, R. et al. Generation of two human isogenic iPSC lines from fetal dermal fibroblasts. *Stem Cell Res.* **33**, 120–124 (2018).
54. Emenecker, R. J., Griffith, D. & Holehouse, A. S. Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. *Biophys. J.* **120**, 4312–4319 (2021).
55. Piovesan, D. et al. MobiDB: intrinsically disordered proteins in 2021. *Nucleic Acids Res.* **49**, D361–d367 (2021).
56. Rainer, J., Gatto, L. & Weichenberger, C. X. ensembldb: an R package to create and use Ensembl-based annotation resources. *Bioinformatics* **35**, 3151–3153 (2019).
57. Landrum, M. J. et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* **46**, D1062–D1067 (2018).
58. Tate, J. G. et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47**, D941–D947 (2018).
59. Sherry, S. T. et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
60. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
62. Pagel, K. A. et al. Integrated informatics analysis of cancer-related variants. *JCO Clin. Cancer Inform.* **4**, 310–317 (2020).
63. Rehm, H. L. et al. ClinGen—the clinical genome resource. *N. Engl. J. Med.* **372**, 2235–2242 (2015).
64. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
65. Rentzsch, P., Schubach, M., Shendure, J. & Kircher, M. CADD-Splice—improving genome-wide variant effect prediction using deep learning-derived splice scores. *Genome Med.* **13**, 31 (2021).
66. Hundal, J. et al. pVACtools: a computational toolkit to identify and visualize cancer neoantigens. *Cancer Immunol. Res.* **8**, 409–420 (2020).
67. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. & Pappu, R. V. CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21 (2017).
68. Cock, P. J. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
69. Scott, M. S., Troshin, P. V. & Barton, G. J. NoD: a nucleolar localization sequence detector for eukaryotic and viral proteins. *BMC Bioinformatics* **12**, 317 (2011).
70. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).
71. Bunn, K. J. et al. Mutations in *DVL1* cause an osteosclerotic form of Robinow syndrome. *Am. J. Hum. Genet.* **96**, 623–630 (2015).
72. White, J. et al. *DVL1* frameshift mutations clustering in the penultimate exon cause autosomal-dominant Robinow syndrome. *Am. J. Hum. Genet.* **96**, 612–622 (2015).

Acknowledgements We thank the patients and their families for participating in this study; V. Suckow, G. Hildebrand, V. Johnston and Y. Zhang for technical assistance; T. Aktas (MPI-MG) for comments on the manuscript; A. Papantonis (GAU, Göttingen) for advice on HMGB1 immunofluorescence; and U. Marchfelder and E. Weiß from the Flow Cytometry Facility of the MPI-MG for their assistance with cell sorting; D. Meierhofer and B. Lukaszewska-McGreal for performing mass spectrometry; and J. Naderi for mCherry constructs and human iPS cells. M.A.M. is a participant in the Digital Clinician Scientist Program and H.L.S. in the Junior Clinician Scientist Program founded by the late D. Dragun and funded by the Berlin Institute of Health and the Charité–Universitätsmedizin Berlin. This work was partially funded by the Max Planck Society. Work in the Hnisz Lab is supported by the Deutsche Forschungsgemeinschaft (DFG) Priority Program Grants HN 4/1-1 and HN 4/3-1 and a Worldwide Cancer Research grant (20-0232). M.S. is supported by grants from the DFG (SP1532/3-2, SP1532/4-1 and SP1532/5-1) and the Deutsches Zentrum für Luft- und Raumfahrt (DLR 01GM1925). S.M. was funded by grant MU 880/16-1 from the DFG. H. Niskanen is supported by fellowships from the Orion Research Foundation and the Instrumentarium Science Foundation. C.G.-C. acknowledges a graduate fellowship from MINECO (PRE2018-084684) and X.S. acknowledges funding from AGAUR (2017 SGR 324), MINECO (PID2019-110198RB-I00) and the European Research Council (CONCERT, contract number 648201). IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Government of Spain).

Author contributions M.A.M., H. Niskanen and A.P.M. conceived and planned the study with input from S. Basu, S.M., M.S., D. Hnisz and D. Horn. M.A.M. managed the collection, analysis and interpretation of patient clinical and molecular data with M.S. and D. Horn. H. Niskanen designed and performed the cell biology experiments. M.L.K. performed the puromycin experiments. H. Niskanen designed and performed the biochemistry experiments with S. Basu. H. Niskanen, S. Basu and R.B. performed image analyses. A.P.M. performed AlphaFold2 modelling and built the variant catalogue with input from M.A.M., H. Niskanen and D. Hnisz. C.G.-C. performed the CD experiments. M.A.M., H. Niskanen, M.S., D. Hnisz and D. Horn wrote the manuscript with contributions from A.P.M. and S. Basu. M.S., D. Hnisz and D. Horn supervised the study. All other authors contributed clinical or molecular data. All authors approved the final manuscript.

Funding Open access funding provided by Max Planck Society.

Competing interests D. Hnisz and X.S. are founders and scientific advisors of Nuage Therapeutics. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05682-1.

Correspondence and requests for materials should be addressed to Malte Spielmann, Denise Horn or Denes Hnisz.

Peer review information *Nature* thanks Marco Bianchi, Diana Mitrea and Stephen Robertson for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at http://www.nature.com/reprints.

3. Diskussion

Meine hier vorliegenden Arbeiten untersuchen den diagnostischen Nutzen des fazialen NGP am Beispiel von DeepGestalt und des NGS (insbesondere des WGS) am Beispiel von Fällen mit angeborenen Extremitätenfehlbildungen.

3.1 Einsatzfähigkeit von DeepGestalt

Es zeigte sich, dass Systeme zum fazialen NGP grundsätzlich einen zusätzlichen diagnostischen Nutzen haben und tatsächlich helfen können, mehrere der Hürden der klassischen Phänotypisierung zu senken (Danyel et al. 2019; Hsieh et al. 2019; Jean Tori Pantel et al. 2020; Mensah et al. 2023).

So war DeepGestalt auch bei Patienten mit vergleichsweise milder fazialer Gestalt (z.B. beim Loeys-Dietz-Syndrom) in der Lage, die korrekte Verdachtsdiagnose zu benennen.
Es war allerdings auffällig, dass die Sensitivität des Systems bei derartigen Erkrankungen mit milden fazialen Zeichen geringer war. Diese deutliche Varianz der syndromspezifischen Sensitivität von DeepGestalt zeigt, wie wichtig es ist, die Leistungsfähigkeit des Systems an verschiedenen Kohorten zu testen. Die Varianz macht außerdem deutlich, dass Vorsicht bei der klinischen Verwendung von DeepGestalt als ein Testsystem von einheitlicher Qualität für alle möglichen Syndrome geboten ist, als welches es angeboten wird. Zeigt sie doch an, dass dasselbe System eine hervorragende Genauigkeit bei einem Syndrom und eine schlechte Spezifität und Sensitivität bei einem anderen Syndrom aufweisen kann. Dies gilt auch für andere faziale NGP-Systeme (Mensah et al. 2022).
DeepGestalt konnte zum Untersuchungszeitpunkt mindestens 238 verschiedene Diagnosen vorschlagen. Damit hat es schon jetzt das Potenzial, einen Arzt auf ihm unbekannte Diagnosen hinzuweisen. Selbst wenn nicht alle vom INSERM in Orphanet gelisteten Erkrankungen eine genetische Ursache und eine typische Fazies aufweisen, bedeutet dies aber auch, dass DeepGestalt sein Potenzial noch nicht ausgeschöpft hat. Künftige Versionen und vergleichbare Anwendungen müssen dazu für weitere Diagnosen trainiert werden. Es ist davon auszugehen, dass dies gelingen wird und die Zahl der von DeepGestalt und vergleichbaren Anwendungen unterstützten Syndrome zunimmt. Während des Untersuchungszeitraums war dies bereits zu beobachten. Um Systeme der fazialen

Phänotypisierung für die Erkennung eines Syndroms zu trainieren, bedarf es einer bestimmten Anzahl an Bildern von Patienten dieses Syndroms. Es wurde vermutet, dass die Zahl an Bildern, die für die Konstruktion eines bestmöglichen Klassifikators in DeepGestalt notwendig ist, noch nicht für alle bereits unterstützten Syndrome erreicht wurde (Pantel et al. 2018). D.h., mit der zu erwartenden zunehmenden Größe des Trainingsdatensatzes dürfte auch die Genauigkeit des Systems zunehmen.

Die Untersuchungen der Fazies des *DONSON*-assoziierten Kleinwuchses und von BPTAS mittels DeepGestalt zeigen nicht nur, dass es zur diagnostischen Entscheidungsunterstützung dienen, sondern auch, dass es bei bisher undefinierten Syndrome verwendet werden und die Technologie zur Definition solcher Syndrome beitragen kann. So konnte gezeigt werden, dass die Vermutung zutrifft, dass es einen charakteristischen fazialen Aspekt des *DONSON*-assoziierten Kleinwuchses gibt. Interessanterweise gaben die Daten, wenn auch in geringerem Maße, ebenfalls einen Hinweis auf ein typisches Gesicht der Fanconi-Anämie, dessen Existenz immer wieder vermutet wird (Avila et al. 2014). Mit einem nur relativ schwach positiven MCC ist sie auch nach der mit DeepGestalt durchgeführten Bildklassifikation nicht abschließend zu klären. Zusätzlich zu der fehlenden onkologischen und hämatologischen Symptomatik und den spezifischen Skelettanomalien zeigen diese Ergebnisse der fazialen Phänotypisierung allerdings, dass es sich bei dem *DONSON*-assoziierten Kleinwuchs um eine andere (möglicherweise eigenständige) Krankheitsentität und nicht um einen Subtyp der FA handelt.

Jüngste klinische und funktionelle Daten geben einen Hinweis auf diese Entität. DONSON ist ein essentielles Initiator-Protein für die Replisom-Bildung in Wirbeltieren und spielt damit eine wichtige Rolle bei der DNA-Replikation (Hashimoto et al. 2023). Das macht es auch zu einem den Zellzyklus steuernden Gen. Es ist wenig überraschend, dass Defekte solcher Gene zu Wachstumsstörungen führen. Entsprechende, einander ähnelnde Phänotypen einer Gruppe von Genen, die an der Steuerung der DNA-Replikation beteiligt sind, wurden unter dem Begriff Meier-Gorlin-Syndrom (MGS) zusammengefasst. (Nielsen-Dandoroff, Ruegg, and Bicknell 2023)

Nach Veröffentlichung meiner hier aufgeführten Arbeiten wurde der Phänotyp auch einiger Träger biallelischer *DONSON*-Varianten als MGS klassifiziert (Knapp et al. 2020). Zu beachten ist allerdings, dass diese Zuordnung nicht für alle Familien mit *DONSON*-Mutationen getroffen werden konnte. Einige hatten auch eine klinische Diagnose eines Seckel-ähnlichen Syndroms (*Seckel-like syndrome*, SLS) bzw. eines femoral-fazialen Syndroms (FFS). (Karaca et al. 2019). MGS ist allerdings durchaus phänotypisch variabel und nur unvollständig penetrant. Die Trias aus Kleinwuchs, Mikrozephalie und Patelladefekten gilt allerdings als charakteristisch. Es fällt auf, dass diese deutlich mit dem klinischen Bild eines SLS bzw. FFS überlappt. Besonders interessant ist, dass für das MGS eine typische Fazies beschrieben wurde, die absteigende Lidachsen, eine volle Unterlippe und eine prominente Nase zeigt. (Nielsen-Dandoroff, Ruegg, and Bicknell 2023)

Dies steht im Einklang mit den hier vorliegenden Arbeiten. Es wirft allerdings die Frage auf, ob MISSLA (bzw. die Formen des *DONSON*-assoziierten mikrozephalen Kleinwuchses) sich als eine Unterform von MGS von anderen Entitäten aus dem MGS-Spektrum mittels Computer-gestützter fazialer Phänotypisierung unterscheiden lässt. Künftige Untersuchungen z.B. mit DeepGestalt an entsprechenden Bildkohorten sind für die Beantwortung dieser Frage notwendig.

Die relativ kleine Kohortengröße der hier aufgeführten Arbeiten kann zu einer Überanpassung (Overfitting) des Modells auf die Daten geführt haben. Eine solche würde in fälschlich hohen Testmetriken resultieren. Ein Overfitting kann auch bei der verwendeten Kreuzvalidierung nicht gänzlich ausgeschlossen werden. Allerdings sprechen die geringen MCCs, die bei einer Zufallsverteilung der Bilder von MISSLA. FA, SLOS und zwei Kontrollgruppen auf die fünf Klassen auftraten, gegen eine Überanpassung des Modells. Künftige, größere Datensätze werden das Risiko eines Overfittings weiter minimieren.

Da DeepGestalt sich allein auf die phänotypische faziale Information stützt, ist es grundsätzlich in der Lage, auch solche Syndrome vorzuschlagen, die zwar definiert sind, deren (genetische) Ursache aber ungeklärt ist. So kann es z.B. bereits das Dubowitz-Syndrom vorschlagen. Dies gibt fazialen NGP-Systemen wie DeepGestalt einen nicht unerheblichen zusätzlichen Nutzen gegenüber sequenzdatenbasierten diagnostischen Entscheidungsunterstützungssystemen.

Die hier aufgeführten Arbeiten zur Untersuchung der Sensitivität von DeepGestalt waren retrospektiv angelegt. Das bedeutet, dass bei allen zur Evaluation der diagnostischen Genauigkeit eingeschlossenen Fällen vor Studieneinschluss bereits auf klassischem Wege (klinisch und/oder molekulargenetisch) eine Diagnose gestellt wurde und DeepGestalt auch für das Erkennen dieser trainiert worden ist. Dessen eingedenk sind auch die für bestimmte Syndrome teils beachtlich hohen Sensitivitäten kritisch zu bewerten. Schließlich bedeutet jede Sensitivität unter 100% dennoch, dass einige Fälle, die diagnostisch geklärt werden können und von DeepGestalt unterstützt werden, von dem System nicht gelöst wurden. Darüber hinaus kann die so gemessene durchschnittliche Sensitivität von DeepGestalt nicht unmittelbar auf den klinischen Alltag übertragen werden, denn nicht alle Patienten, die sich in der klinischen Genetik vorstellen, haben ein von DeepGestalt unterstütztes Syndrom.

Marwaha et al. haben eine prospektive Studie zum diagnostischen Nutzen von DeepGestalt durchgeführt (A. Marwaha et al. 2021). Für die Gruppe der Syndrome, die von DeepGestalt unterstützt werden, berichten sie mit 82% eine zu den in den vorliegenden Arbeiten vergleichbare Top-10-Sensitivität. Bemerkenswert ist insbesondere die berichtete Top-10-Sensitivität von 57% in der gesamten prospektiven Kohorte. Dies lässt vermuten, dass die Mehrzahl der Patienten, die sich in der klinischen Genetik vorstellen, ein von DeepGestalt zum Untersuchungszeitpunkt bereits unterstütztes Syndrom aufwies.

DeepGestalt wurde für die Zuordnung einer Liste möglicher Verdachtsdiagnosen zu dem Bild eines Patienten entwickelt. Der beabsichtigte Zweck bestand nicht in der Unterscheidung von Fotos, die Patienten mit einem fazial dysmorphen Syndrom zeigen, von solchen Aufnahmen, die fazial unauffällige Personen zeigen. Dennoch wird es dazu bei der klinischen Entscheidungsunterstützung verwendet (z.B. wenn eruiert wird, ob einem Patienten weitere genetische Tests angeboten werden sollen). Und tatsächlich erzielte DeepGestalt im Mittel höhere Gestalt Scores bei Fotos von Patienten mit genetischen Syndromen als bei Bildern von fazial unauffälligen Kontrollprobanden. Allerdings ist die erreichte Trennschärfe nur gering (AUROC: 0.72) und die Höhe eines erzielten DeepGestalt Scores als Maß für das tatsächliche Vorliegen eines Syndroms bei der klinischen Entscheidungsunterstützung nur eingeschränkt geeignet.

Die Fähigkeit von DeepGestalt, Bilder von Menschen mit einer syndromalen Erkrankung von Bildern von Menschen ohne eine syndromale Erkrankung zu unterscheiden, war gemessen an der Sensitivität in anderen retrospektiven Studien mitunter größer als in den hier aufgeführten Arbeiten (Liehr et al. 2018; Vorravanpreecha et al. 2018; Carli et al. 2019; Srisraluang and Rojnueangnit 2021). Dies kann unterschiedliche Ursachen haben (z.B. größere Kohorten, fehlendes Matching der Kontrollen).

Die größte Schwäche von DeepGestalt ist, dass es auch Fotos von gesunden Probanden Syndrome zuordnet und eine Klasse “unauffälliges Gesicht” fehlt, um den Anwender darauf aufmerksam zu machen. Meine vorliegenden Arbeiten zeigen, dass die Höhe des GestaltScores kein akkurates Maß ist, um das Vorhandensein bzw. Nicht-Vorhandensein einer vorgeschlagenen Diagnose abzuschätzen. Für den klinischen Einsatz von DeepGestalt ist insbesondere interessant, dass die Wahrscheinlichkeiten, mit denen DeepGestalt bestimmte Syndrome vorschlägt, stark variabel sind. Zwar wurden die meisten Diagnosen nur selten vorgeschlagen, doch bestimmte Syndrome erscheinen in den Vorschlagslisten (Top30) von weit mehr als der Hälfte der untersuchten Bilder und zwar sowohl bei Fotos von unauffälligen Kontrollprobanden als auch bei Fotos von Patienten mit einem genetischen Syndrom.

So wird z.B. das Angelman-Syndrom sehr häufig vorgeschlagen, hat also eine geringe Spezifität. Dies bedeutet in der klinischen Routine, dass der Verdacht auf ein Vorliegen eines Angelman-Syndroms auch dann gering ist, wenn es in der Ergebnisliste genannt wird.

Der Vorschlag beispielsweise eines Crouzon-Syndroms durch DeepGestalt hingegen ist spezifisch. Es taucht nur äußerst selten in den Ergebnislisten von unauffälligen Kontrollen oder Kontrollen mit einer anderen Diagnose auf.

Für die Syndrome, für welche in den vorliegenden Arbeiten eine geringe Spezifität von DeepGestalt gefunden wurde, haben auch andere Gruppen geringe Spezifitäten ermittelt (Vorravanpreecha et al. 2018; A. Marwaha et al. 2021). Es handelt sich folglich nicht um Artefakte, die sich aus der Zusammensetzung der Bildkohorte ergeben.

Die ermittelten Spezifitäten sollten daher beim Einsatz des Systems unbedingt berücksichtigt werden, wünschenswert wäre, wenn diese gemeinsam mit den Ergebnissen angezeigt würden. Bei Patienten einer definierten Kohorte (wie hier am Beispiel von BPTAS) könnte die übereinstimmende Nennung von Verdachtsdiagnosen, welche das System normalerweise nur selten vorschlägt, dann auch einfacher als Maß für einen gemeinsamen fazialen Phänotyp genutzt werden.

Bei einigen Syndromen ist es möglich, dass DeepGestalt für deren Erkennung eine geringe Spezifität aufweist, weil ein charakteristischer fazialer Aspekt, der bei deren Diagnosestellung hilfreich sein könnte, gar nicht existiert. So ist für das Klinefelter-Syndrom (Karyotyp 47,XXY) keine typische Fazies bekannt (Bird and Hurren 2016) und ein relativ geringer Bartwuchs, der als faziales Zeichen auftreten kann, wird von DeepGestalt grundsätzlich nicht berücksichtigt (Gurovich et al. 2019). Dennoch wurde das System zur Erkennung des Syndroms trainiert und es bestand zumindest in den untersuchten Versionen eine Klasse für das Klinefelter-Syndrom. Ich vermute, dass dies geschah, weil für das Klinefelter-Syndrom, dessen Häufigkeit auf 1:500 bis 1:1000 unter der männlichen Bevölkerung geschätzt wird (Los and Ford 2023), eine relativ große Bildkohorte bei der Erstellung von DeepGestalt zur Verfügung stand. Dies unterstreicht die Notwendigkeit, nicht nur informatische, sondern auch spezifisch medizinische Fachkenntnisse bei der Entwicklung derartiger Software einzubringen. Künftige Untersuchungen des Nutzens von DeepGestalt und vergleichbarer Anwendungen sollten auch eine systematische Prüfung der Eignung der unterstützten Syndrome für eine Computer-gestützte faziale Phänotypisierung umfassen.

Frühere Arbeiten ließen vermuten, dass der ethnische Hintergrund eines Patienten die diagnostische Genauigkeit von ersten Versionen von DeepGestalt beeinflusst (Lumaka et al. 2017). Ein solcher Effekt fand sich bei den vorliegenden Arbeiten nicht, was im Einklang mit anderen jüngeren Arbeiten an neueren Versionen von DeepGestalt steht (Vorravanpreecha et al. 2018; Mishima et al. 2019). Es ist nicht auszuschließen, dass dies auf einen inzwischen erweiterten Trainingsdatensatz zurückzuführen ist, aufgrund dessen das System nun auch in ethnisch diversen Populationen akkurat arbeitet.

Die Ergebnisse zeigen, dass DeepGestalt eine relevante Zahl an Diagnosen kennt, eine beachtliche Sensitivität und für die meisten Syndrome, die es unterstützt, auch eine gute Spezifität aufweist. Die Werte sind hoch genug für den ergänzenden Einsatz als diagnostisches Entscheidungsunterstützungssystem für den Einsatz in der klinischen Genetik oder auch in pädiatrischen Abteilungen. Auf keinen Fall eignet sich das System für den primären Einsatz durch Laien.

3.2 Zusätzlicher diagnostischer Nutzen des NGS

Auch das NGS zeigte einen zusätzlichen diagnostischen Nutzen für die klinische Routine in der Humangenetik. WES- oder WGS-Ansätze können im Gegensatz zur gezielten Sequenzierung einzelner Gene bzw. zur Paneldiagnostik ausgewählter Gene nach klassischer Phänotypisierung auch dann eingesetzt werden, wenn ein Patient zwar auffällig erscheint, aber nur unspezifische oder untypische Symptome zeigt.

Dies zeigte sich z.B. bei dem Patienten, bei welchem eine *HOXD13*-Repeat-Expansion vorlag (Individual 9 in der hier vorliegenden Studie über die Genomsequenzierung) (Elsner et al. 2021). Die *HOXD13*-assoziierte Synpolydaktylie ist klinisch äußerst variabel und unvollständig penetrant. Hauptmerkmale sind eine Syndaktylie der Finger 3 und 4 (die von einer zentralen Polydaktylie begleitet sein kann) sowie eine postaxiale Polydaktylie des Fußes. (Guo et al. 2021; Gottschalk et al. 2023)
Die Handfehlbildungen des oben genannten Patienten wiesen beidseits verkürzte proximale Phalangealknochen des 4. Fingers auf, welche nur bei wenigen Fällen der *HOXD13*-assoziierten Synpolydaktylie (dann mit deutlicher Repeatexpansion) auftreten. Eine gezielte *HOXD13*-Analyse war daher nicht erfolgt.

Auch wenn eine vorliegende Erkrankung dem Untersucher unbekannt ist, kann ein exomweites bzw. genomweites NGS zur molekularen Klärung der Diagnose beitragen, wie sich an den hier aufgeführten Beispielen des *DONSON*-assoziierten MISSLA-Syndroms und des BPTAS nachvollziehen lässt.

WES und WGS können neue Krankheitsgene identifizieren, bzw. helfen diese zu validieren. Mit *SEMA3D*, *ALDH1A2*, und *HMGB1* konnten in einer relativ kleinen Kohorte von 69 Fällen mittels WGS drei Kandidatengene identifiziert werden. Mit *UBA2* sogar ein Kandidat als Krankheitsgen validiert werden.
Frameshift-Mutationen in *HMGB1* ließen sich mit vier weiteren Patienten als Ursache des BPTAS validieren (Mensah et al. 2023). Dies war nur möglich, weil die in den vorliegenden Arbeiten analysierten Patienten für den Studieneinschluss ein klar definiertes phänotypisches Spektrum, nämlich Fehlbildungen der Extremitäten, aufweisen mussten. Es ist daher zu vermuten, dass auch mit relativ kleinen Kohorten von Patienten mit ähnlich klar definierten Fehlbildungen anderer Organsysteme (z.B. Nierenfehlbildungen oder Herzfehler) neue Krankheitsgene identifiziert werden können.

Biallelische pathogene Mutationen von *ALDH1A2* wurden inzwischen mit einem komplexen Syndrom assoziiert, das mit Zwerchfellhernien und kardiovaskulären Fehlbildungen einhergeht; welches u.a. aber auch bestimmte skelettale Fehlbildungen wie fehlende Rippen und bemerkenswerter Weise eine Syndaktylie sowohl der Hände als auch Füße verursachen kann (Beecroft et al. 2021). Ob die isolierte zentrale Syndaktylie der Hände und Füße, welche in der hier vorliegenden Arbeit bei zwei Patienten mit einer monoallelischen frameshift-Variante von *ALDH1A2* (Vater und Tochter der Familie 17 (Elsner et al. 2021)) vorlag, als ein Zeichen der Anlageträgerschaft gedeutet werden kann, müssen künftige Arbeiten zeigen. Die Identifikation weiterer heterozygoter *ALDH1A2*-Mutationsträger mit Syndaktylien der oberen und unteren Extremität, die nicht mit Familie 17 verwandt sind, könnte dazu beitragen

Mit 17,4 % fiel die diagnostische Rate des WGS moderat aus. Dabei ist allerdings unbedingt zu beachten, dass die Studie prospektiv an nur solchen Patienten mit unauffälliger Routinediagnostik durchgeführt wurde. D.h., dass das WGS einen zusätzlichen diagnostischen Nutzen bietet, der sich in der Größenordnung einer großen Studie des britischen National Health Service (NHS) zum Potenzial des WGS bewegte (Turro et al. 2020). In diesem Zusammenhang ist besonders interessant, dass es uns gelang, verschiedene Typen krankheitsrelevanter Mutationen mit einem einzigen Testverfahren nachzuweisen (Substitutionen, Indels, Repeatexpansionen, komplexe strukturelle Veränderungen).

Von besonderer Bedeutung ist dabei die in einer Familie (Familie 10 (Elsner et al. 2021)) mit Ektrodaktylie detektierte Inversion am SHFM3 Locus. Diese umfasste nur 105 kb und war damit zu klein für die Detektion im Rahmen einer klassischen, mikroskopischen Chromosomenanalyse. Da eine Inversion kopienzahlneutral ist, hätte auch eine Mikroarray-basierte aCGH diese nicht entdecken können. Auch eine Detektion der Inversion mittel WES ist fraglich, da der proximale Bruchpunkt im intergenischen Bereich zwischen den Genen *BTRC* und *POLL* und der distale Bruchpunkt in Intron 5 von *FBXW4* liegt. Beide Bruchpunkte sind folglich nicht durch ein Exom abgedeckt, so dass die Inversion nicht mittels auffälliger read-pairs hätte gecallt werden können. Ein Calling im Rahmen eines WES über eine veränderte Coverage der Gene *POLL* und *DPCD*, welche die Inversion am SHFM3 Locus umfasst, ist aufgrund der oben genannten Kopienzahlneutralität einer Inversion ebenfalls nicht möglich. WGS war folglich der einzige unvoreingenommene Test, der die Detektion einer solchen Inversion ermöglichte.

Dies zeigt, dass short-read paired-end WGS eine bedeutende Hürde der gezielten genetischen Testung nach klassischer Phänotypisierung senken kann: Relevante Mutationen werden mit geringerer Wahrscheinlichkeit aufgrund des gewählten Testverfahrens übersehen. Die Möglichkeit, WGS als eine “One-test-for-all-Strategie” anzuwenden, birgt auch das Potenzial zur Reduktion von Dauer und Kosten in der genetischen Labordiagnostik.

Die große Zahl auch mittels WGS ungelöster Fälle von Extremitätenfehlbildungen zeigt allerdings auch eindrücklich, dass das funktionelle Verständnis des Genoms und die gegenwärtigen Teststrategien noch nicht zur umfassenden Klärung der molekulargenetischen Ursachen erblicher Erkrankungen ausreichen. Varianten in den 98,5% des Genoms, die nicht für Proteine kodieren, sind mit Sicherheit auch für die Entstehung erblicher Erkrankungen relevant.

Wir konnten solche Varianten auch identifizieren, nicht aber in dem Maße interpretieren, wie es für eine Priorisierung und Diagnosestellung notwendig wäre. Short-read paired-end WGS erfasst trotz seines Potenzials nicht alle möglicherweise relevanten Veränderungen des Erbguts, lange repetitive Sequenzen und ein möglicher Mosaikstatus werden nur unzureichend erfasst, zur Detektion epigenetischer Veränderungen braucht es Spezialtechniken (Zhao et al. 2021; King et al. 2017; Yan et al. 2016). Auch dies könnte einen Teil der in der vorliegenden Arbeit ungelösten Fälle erklären.

Die Ergebnisse zeigen, dass sich das WGS zur Steigerung der diagnostischen Rate und Effizienz der Labordiagnostik in der klinischen Humangenetik eignet. Weitere

Grundlagenforschung ist nötig, um eine Interpretierbarkeit nicht-kodierender, insbesondere intergenischer Varianten zu ermöglichen.

3.3 Integration von fazialem NGP in NGS-Datenauswertung

Meine hier vorliegenden Arbeiten zeigen, dass zur Steigerung der diagnostischen Rate und sicherlich auch zur Senkung der Dauer bis zur Diagnosestellung die Integration von Systemen zum fazialen NGP in konventionelle Algorithmen zur NGS-Datenanalyse beitragen kann. In der PEDIA-Studie konnten wir zeigen, dass ein faziales NGP allein zwar eine geringere Sensitivität als eine HPO-Term-gestützte Exomanalyse aufweist, dass die Kombination beider Verfahren allerdings mit der höchsten Sensitivität einhergeht.

Die Top-10-Sensitivität lag in der Studie, die retrospektiv an einer Kohorte von Patienten mit konventionell gestellter Diagnose durchgeführt wurde, bei 99%. Es konnte also mit einem Verfahren der reversen Phänotypisierung eine vergleichbare Sensitivität wie bei gezielter genetischer Testung nach konventioneller Phänotypisierung erreicht werden. Allerdings muss beachtet werden, dass dieser Wert nur für Syndrome gilt, die zum Untersuchungszeitpunkt von DeepGestalt unterstützt wurden, denn das Vorhandensein eines solchen Syndroms war ein Einschlusskriterium. Mit einer Zunahme der Zahl der unterstützten Syndrome ist allerdings zu rechnen und sie zeigte sich auch schon im Laufe der hier vorgestellten Arbeiten.

Der PEDIA Algorithmus wurde an künstlich konstruierten Exomdaten (Einfügen von einer bzw. zwei pathogenen Mutationen in je ein Gen bei Exomdaten von gesunden Probanden) trainiert und getestet. Es ist zwar anzunehmen, dass dies eine realistische Simulation echter klinischer Exomdaten darstellt, künftige Arbeiten werden die Nützlichkeit von PEDIA jedoch an unsimulierten Daten bestätigen müssen.

Bemerkenswert war, dass der ethnische Hintergrund der verwendeten Exomdaten die Genauigkeit von PEDIA nicht beeinflusste. Das war nicht unbedingt zu erwarten, denn das humane Referenzgenom basiert stark auf den Informationen von Probanden europäischer Herkunft und die genetische Diversität zu und in anderen menschlichen Populationen ist groß, so dass bei Exomdaten von Probanden afrikanischer oder asiatischer Herkunft solche Varianten, die als selten gelten, häufiger zu finden sind, als z.B. bei Europäern.(Popejoy and Fullerton 2016; Wong et al. 2020)

Welche Rolle die Bildqualität, also Faktoren wie Beleuchtung, Kontrast oder Auflösung, aber auch das Tragen einer Brille, der Haarschnitt, oder die Perspektive der Aufnahme, auf die Genauigkeit von DeepGestalt haben, ist bisher nicht untersucht worden. Dass diese Variablen einen Einfluss haben, ist anzunehmen, da die Genauigkeit anderer Anwendungen, die mittels der Techniken des maschinellen Lernens eine Auswertung von Bilddaten durchführen, in Abhängigkeit von diesen Parametern schwankt. Darüberhinaus wurden in der PEDIA-Studie schlechtere DeepGestalt Scores bei schlechterer Bildqualität bemerkt. Eine systematische Untersuchung möglicher Störparameter, die in der Aufnahmequalität begründet liegen, steht allerdings weiterhin aus. Selbst wenn eine schlechte Bildqualität ein System zur fazialen Phänotypisierung erheblich stört, sollte es allerdings möglich sein, qualitativ hochwertige Portraitaufnehmen eines Patienten in der genetischen Routinesprechstunde anzufertigen, da diese dort traditionell zum Zwecke der medizinischen Dokumentation angefertigt werden und die meisten humangenetischen Ambulanzen und Praxen daher über die dafür nötige Ausstattung verfügen.

Darin zeigt sich auch die große Stärke eines Systems wie DeepGestalt, das mit gewöhnlichen zweidimensionalen digitalen Bilddaten, wie sie von handelsüblichen Digitalkameras erstellt werden können, und nicht mit den Daten von 3D-Kameras arbeitet. DeepGestalt erlaubt es, mehrere Bilder eines Patienten auszuwerten. Hier könnten z.B. bei der Auswertung eines Falls Bilder derselben Person in verschiedenen Altersstufen verwendet werden, ob dies einen Effekt auf die Genauigkeit der Syndromvorschläge hat, muss ebenfalls noch systematisch untersucht werden. Entsprechend der Trio-basierten Analyse von NGS-Daten könnte auch eine vergleichende Analyse von Bildern von Indexpatienten und Eltern (oder anderen Verwandten, z.B. im Sinne einer Segregationsanalyse von betroffenen bzw. nicht-betroffenen Geschwistern) von Systemen wie DeepGestalt durchgeführt werden, um die Syndromvorschläge zu filtern oder integrierte phänotypische und Sequenzdaten wie beim PEDIA-Ansatz zu priorisieren. Besonders für die Gruppe der ultra-seltenen Erkrankungen, von denen nur wenige Dutzend Fälle in der Fachliteratur beschrieben sind, sind die Kohorten, die für das Training von DeepGestalt (und folglich auch PEDIA und anderen derivaten Algorithmen) verwendet werden können, sehr klein. Das Beispiel BPTAS zeigt, wie Systeme zur fazialen Phänotypisierung auch bei der Testung auf derartige Syndrome verwendet werden können (Mensah et al. 2023).

Ansätze wie PEDIA können nur dann für die Nutzung der fazialen Information zur Variantenpriorisierung trainiert werden, wenn die genetische Ursache eines erblichen Syndroms geklärt ist. In diesem Punkt sind sie der gezielten genetischen Testung nach konventioneller Phänotypisierung nicht überlegen. Allerdings ermöglichen es Ansätze wie

PEDIA grundsätzlich eine genetische Spezialdiagnostik auch in klinischen Umgebungen anzubieten, in welchen der Umfang an syndromologischem Wissen gering ist, z.B. besonders seltene Syndrome nicht bekannt sind. Dazu zählen insbesondere pädiatrische Stationen und Primärversorger.

3.4 Mögliche künftige Strategien zur Steigerung der diagnostischen Rate seltener genetischer Erkrankungen

NGS und faziales NGP und nicht zuletzt die Kombination von beidem haben in den vorliegenden Arbeiten ihr Vermögen zur Steigerung der diagnostischen Rate in der klinischen Genetik gezeigt. Insbesondere drei der Hürden des klassischen Ansatzes einer gezielten genetischen Testung nach klinischer Phänotypisierung bleiben aber auch für diese neuen Verfahren, welche eine reverse Phänotypisierung nutzen, bestehen. Sie sind nur begrenzt einsetzbar, *(i)* wenn ein Syndrom nicht definiert, *(ii)* wenn die molekulargenetische Ursache eines Syndroms unklar und *(iii)* wenn das gewählte (NGS-)Testverfahren für die Erfassung der Art einer tatsächlich vorliegenden Mutation nicht geeignet ist.

Darüber hinaus kann ein faziales NGP seiner Natur nach nur der automatisierten Erfassung phänotypischer Parameter des Gesichts dienen, aber nicht alle erblichen Erkrankungen zeigen eine typische dysmorphe faziale Gestalt. Zur weiteren Verbesserung der hier vorgestellten Ansätze müssen künftige Verfahren daher auch Patienten mit unklaren Diagnosen und vergleichbaren Phänotypen einander zuordnen können. Ein nachfolgender Vergleich der Erbinformationen der Patienten könnte dann zur Identifikation neuer Krankheitsgene führen. Der kürzlich vorgestellte, auf DeepGestalt basierende, GestaltMatcher könnte dazu dienen (Hsieh et al. 2022).

Der PEDIA-Algorithmus priorisiert Exomdaten. Um bei der automatisierten Analyse auch Strukturvarianten besser erfassen und priorisieren zu können, braucht es künftige Systeme, die auf Genomdaten oder gar Daten aus verschiedenen Quellen (z.B. zusätzlich klassische Karyotypisierung und aCGH) trainiert wurden.

Fehlbildungen und anatomische Varianten, welche nicht das Gesicht betreffen, die für die Diagnose eines bestimmten erblichen Syndroms aber wegweisend sein können, können prinzipiell auch auf Bildern erfasst und für ein NGP genutzt werden. So könnten künftige

Algorithmen z.B. Röntgenaufnahmen von Extremitätenfehlbildungen analysieren (Gottschalk et al. 2023).

4. Zusammenfassung

Da es sehr viele seltene Erkrankungen gibt (das INSERM listet z.B. mehr als 9000), ist ihre kumulative Gesamtprävalenz hoch. Ein Großteil hat genetische Ursachen und Schätzungen gehen von mehreren Millionen betroffenen EU-Bürgern aus. Verschiedene Initiativen z.B. EJP-RD und Translate-NAMSE wollen die Versorgung von Menschen mit seltenen Erkrankungen verbessern. Dazu bedarf es allerdings der Entwicklung und Testung neuer diagnostischer Verfahren in der klinischen Genetik. NGS ist eine solche Technologie, die verspricht, die genetische Diagnostik zu revolutionieren, indem Patienten dadurch eine umfassende, schnelle und relativ kostengünstige genetische Testung erhalten können.

Auch unspezifische, seltene oder primär unklare Symptome können so pathogenen Mutationen zugeordnet werden (reverse Phänotypisierung). Dazu wurden spezielle Software und Leitlinien (ACMG/AMP-Kriterien) entwickelt, die phänotypische, populationsgenetische, evolutionäre und molekulargenetische Informationen für eine genetische Diagnosestellung zusammenführen. Insbesondere die Erfassung des Phänotyps in computerlesbarer Form ist dabei allerdings schwierig. So kann die HPO, der weltweite Standard zur computerlesbaren Beschreibung phänotypischer Merkmale, diese zwar durch explizite, prädefinierte Begriffe erfassen. Allerdings lassen sich viele genetische Syndrome an einem charakteristischen Gesicht erkennen und zur exakten Beschreibung der fazialen Gestalt ist die HPO gerade nicht geeignet. Maschinelles Sehen könnte für eine solche faziale Phänotypisierung verwendet werden (faziales NGP). So wurde z.B. DeepGestalt, ein tiefes neuronales Netzwerk, entwickelt, das gewöhnlichen Portraitfotos von Patienten mit seltenen, fazial-dysmorphen genetischen Erkrankungen automatisiert eine Liste von Verdachtsdiagnosen zuordnet und auch nutzerspezifisch zur Erkennung ausgewählter Syndrome trainiert werden kann.

Diese Arbeit untersucht, wie NGS und faziales NGP die Diagnostik seltener, erblicher Erkrankungen steigern können.

Bei zwei Geschwistern, deren Skelettfehlbildungen und andere Merkmale an eine FA erinnerten und bei denen die Routinediagnostik unauffällig war, konnte ein WES biallelische Mutationen in *DONSON* identifizieren, mit welchem verschiedene Formen des FA-ähnlichen mikrozephalen Kleinwuchses assoziiert worden sind. Die Geschwister waren unterschiedlich stark betroffen, d.h., dass die reverse Phänotypisierung nach WES nicht nur eine Diagnostik ermöglichte, sondern auch zeigte, dass es sich bei den mit *DONSON*-assoziierten Erkrankungen um ein Spektrum nur eines Phänotyps, MISSLA, handelt. MISSLA hat einen fazialen Aspekt, der sich mit DeepGestalt von einer FA unterscheiden ließ. Es ist daher sinnvoll, die neue Entität MISSLA nicht als einen weiteren Subtyp der FA zu definieren.

WES und andere NGS-Daten sind sehr groß und die Filterung und Priorisierung der detektierten Varianten war, auch bei den MISSLA-Geschwistern, sehr aufwendig. Die von DeepGestalt ausgegebenen, gewichteten Listen an Verdachtsdiagnosen ließen sich allerdings bei 679 Patienten mit einer seltenen, molekulargenetisch bestätigten Diagnose erfolgreich nutzen, um die Priorisierung von Exomdaten zu verbessern. Es fiel allerdings bei einer Kohorte von 646 Fotos auch auf, dass DeepGestalt Fotos von Patienten mit einem kraniofazial dysmorphen Syndrom nicht sicher von unauffälligen Kontrollbildern unterscheiden kann und nur 238 verschiedene Diagnosen vorschlagen konnte.

Als umfassender, einem WES überlegener Test, der sich für eine weit größere Zahl an Diagnosen eignet, kommt ein WGS in Frage. Bei 12 von 69 Patienten mit in der Routinediagnostik ungeklärten Extremitätenfehlbildungen konnte das WGS relevante Varianten identifizieren, darunter Varianten in einem aktuell beschriebenen Krankheitsgen (*UBA2)* und drei neuen Kandidatengenen. Eines davon (*HMGB1*) konnte nachfolgend mit Hilfe von vier weiteren Patienten validiert und mit BPTAS assoziiert werden. Bemerkenswert war, dass sich verschiedenste Mutationstypen mit einem einzigen Testverfahren nachweisen ließen. Eine nichtkodierende pathogene Mutation im intergenischen Bereich konnte nicht identifiziert werden.

Systeme zum fazialen NGP wie DeepGestalt können eine erhebliche Sensitivität erreichen und helfen, die Diagnostik bei Menschen mit seltenen genetischen Erkrankungen zu verbessern. Die Spezifität von DeepGestalt ist allerdings nur moderat, so dass ein Einsatz nur durch Experten sinnvoll ist. NGS bietet insbesondere in der Form des WGS das Potenzial, die verschiedenen in der genetischen Diagnostik etablierten Verfahren zu vereinheitlichen, die diagnostische Rate zu steigern und die Zeit bis zur Diagnosestellung zu verkürzen. Besonders hilfreich ist dafür eine Kombination von fazialem NGP und NGS. Zukünftige Systeme werden für das NGP anderer Körperregionen als des Gesichts benötigt und das NGS muss zur akkuraten Erfassung von epigenetischen und strukturellen Varianten in der klinischen Routine weiterentwickelt werden. Für die vollständige Interpretation von WGS-Daten ist künftig vor allem ein genaueres molekulargenetisches Verständnis des nicht-kodierenden Teils des Genoms notwendig.

5. Literaturangaben

1000 Genomes Project Consortium, Adam Auton, Lisa D. Brooks, Richard M. Durbin, Erik P. Garrison, Hyun Min Kang, Jan O. Korbel, et al. 2015. “A Global Reference for Human Genetic Variation.” *Nature* 526 (7571): 68–74.

Amendola, Laura M., Gail P. Jarvik, Michael C. Leo, Heather M. McLaughlin, Yassmine Akkari, Michelle D. Amaral, Jonathan S. Berg, et al. 2016. “Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium.” *American Journal of Human Genetics* 98 (6): 1067–76.

Antonarakis, Stylianos E., Brian G. Skotko, Michael S. Rafii, Andre Strydom, Sarah E. Pape, Diana W. Bianchi, Stephanie L. Sherman, and Roger H. Reeves. 2020. “Down Syndrome.” *Nature Reviews. Disease Primers* 6 (1): 9.

Arts, Peer, Annet Simons, Mofareh S. AlZahrani, Elanur Yilmaz, Eman AlIdrissi, Koen J. van Aerde, Njood Alenezi, et al. 2019. “Exome Sequencing in Routine Diagnostics: A Generic Test for 254 Patients with Primary Immunodeficiencies.” *Genome Medicine* 11 (1): 38.

Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, et al. 2000. “Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium.” *Nature Genetics* 25 (1): 25–29.

Avila, Lucia Fátima de Castro, Wilson Denis Martins, Lisiane Cândido, Sergio Aparecido Ignácio, Carmen Maria S. Bonfim, and Marina de Oliveira Ribas. 2014. “A Study of Facial Pattern in Patients with Fanconi Anemia.” *The Cleft Palate-Craniofacial Journal: Official Publication of the American Cleft Palate-Craniofacial Association* 51 (1): 83–89.

Beecroft, Sarah J., Marcos Ayala, George McGillivray, Vikas Nanda, Emanuele Agolini, Antonio Novelli, Maria C. Digilio, et al. 2021. “Biallelic Hypomorphic Variants in ALDH1A2 Cause a Novel Lethal Human Multiple Congenital Anomaly Syndrome Encompassing Diaphragmatic, Pulmonary, and Cardiovascular Defects.” *Human Mutation* 42 (5): 506–19.

Berche, Patrick. 2022. “History of Measles.” *Presse Medicale* 51 (3): 104149.

Best, Sunayna, Jing Yu, Jenny Lord, Matthew Roche, Christopher Mark Watson, Roel P. J. Bevers, Alex Stuckey, et al. 2022. “Uncovering the Burden of Hidden Ciliopathies in the 100 000 Genomes Project: A Reverse Phenotyping Approach.” *Journal of Medical Genetics*, June. https://doi.org/10.1136/jmedgenet-2022-108476.

Bird, Rebecca J., and Bradley J. Hurren. 2016. “Anatomical and Clinical Aspects of Klinefelter’s Syndrome.” *Clinical Anatomy* 29 (5): 606–19.

Boehringer, Stefan, Manuel Guenther, Stella Sinigerova, Rolf P. Wurtz, Bernhard Horsthemke, and Dagmar Wieczorek. 2011. “Automated Syndrome Detection in a Set of Clinical Facial Photographs.” *American Journal of Medical Genetics. Part A* 155A (9): 2161–69.

Boehringer, Stefan, Tobias Vollmar, Christiane Tasse, Rolf P. Wurtz, Gabriele Gillessen-Kaesbach, Bernhard Horsthemke, and Dagmar Wieczorek. 2006. “Syndrome Identification Based on 2D Analysis Software.” *European Journal of Human Genetics: EJHG* 14 (10): 1082–89.

Bourchany, Aurélie, Christel Thauvin-Robinet, Daphné Lehalle, Ange-Line Bruel, Alice Masurel-Paulet, Nolwenn Jean, Sophie Nambot, et al. 2017. “Reducing Diagnostic Turnaround Times of Exome Sequencing for Families Requiring Timely Diagnoses.” *European Journal of Medical Genetics* 60 (11): 595–604.

Bowling, Kevin M., Michelle L. Thompson, Candice R. Finnila, Susan M. Hiatt, Donald R. Latner, Michelle D. Amaral, James M. J. Lawlor, et al. 2022. “Genome Sequencing as a First-Line Diagnostic Test for Hospitalized Infants.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 24 (4): 851–61.

Canard, B., and R. S. Sarfati. 1994. “DNA Polymerase Fluorescent Substrates with Reversible 3’-Tags.” *Gene* 148 (1): 1–6.

Carli, Diana, Elisa Giorgio, Francesca Pantaleoni, Alessandro Bruselles, Sabina Barresi, Evelise Riberi, Francesco Licciardi, et al. 2019. “NBAS Pathogenic Variants: Defining the Associated Clinical and Facial Phenotype and Genotype-Phenotype Correlations.” *Human Mutation* 40 (6): 721–28.

Cerrolaza, Juan J., Antonio R. Porras, Awais Mansoor, Qian Zhao, Marshall Summar, and Marius George Linguraru. 2016. “Identification of Dysmorphic Syndromes Using Landmark-Specific Local Texture Descriptors.” *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. https://doi.org/10.1109/isbi.2016.7493453.

Chiu, Annie Ting Gee, Claudia Ching Yan Chung, Wilfred Hing Sang Wong, So Lun Lee, and Brian Hon Yin Chung. 2018. “Healthcare Burden of Rare Diseases in Hong Kong - Adopting ORPHAcodes in ICD-10 Based Healthcare Administrative Datasets.” *Orphanet Journal of Rare Diseases* 13 (1): 147.

Choukair, Daniela, Fabian Hauck, Markus Bettendorf, Heiko Krude, Christoph Klein, Tobias Bäumer, Reinhard Berner, et al. 2021. “An Integrated Clinical Pathway for Diagnosis, Treatment and Care of Rare Diseases: Model, Operating Procedures, and Results of the Project TRANSLATE-NAMSE Funded by the German Federal Joint Committee.” *Orphanet Journal of Rare Diseases* 16 (1): 474.

Cipriani, Valentina, Nikolas Pontikos, Gavin Arno, Panagiotis I. Sergouniotis, Eva Lenassi, Penpitcha Thawong, Daniel Danis, et al. 2020. “An Improved Phenotype-Driven Tool for Rare Mendelian Variant Prioritization: Benchmarking Exomiser on Real Patient Whole-Exome Data.” *Genes* 11 (4). https://doi.org/10.3390/genes11040460.

Coutelier, Marie, Manuel Holtgrewe, Marten Jäger, Ricarda Flöttman, Martin A. Mensah, Malte Spielmann, Peter Krawitz, Denise Horn, Dieter Beule, and Stefan Mundlos. 2022. “Combining Callers Improves the Detection of Copy Number Variants from Whole-Genome Sequencing.” *European Journal of Human Genetics: EJHG* 30 (2): 178–86.

Cristofoli, Francesca, Elisa Sorrentino, Giulia Guerri, Roberta Miotto, Roberta Romanelli, Alessandra Zulian, Stefano Cecchin, et al. 2021. “Variant Selection and Interpretation: An Example of Modified VarSome Classifier of ACMG Guidelines in the Diagnostic Setting.” *Genes* 12 (12). https://doi.org/10.3390/genes12121885.

Danyel, Magdalena, Zhuo Cheng, Christine Jung, Felix Boschann, Jean Tori Pantel, Nurulhuda Hajjir, Ricarda Flöttmann, et al. 2019. “Differentiation of MISSLA and Fanconi Anaemia by Computer-Aided Image Analysis and Presentation of Two Novel MISSLA Siblings.” *European Journal of Human Genetics: EJHG* 27 (12): 1827–35.

Deciphering Developmental Disorders Study. 2015. “Large-Scale Discovery of Novel

Genetic Causes of Developmental Disorders.” *Nature* 519 (7542): 223–28.

DePristo, Mark A., Eric Banks, Ryan Poplin, Kiran V. Garimella, Jared R. Maguire, Christopher Hartl, Anthony A. Philippakis, et al. 2011. “A Framework for Variation Discovery and Genotyping Using next-Generation DNA Sequencing Data.” *Nature Genetics* 43 (5): 491–98.

Donk, Roos van der, Sandra Jansen, Janneke H. M. Schuurs-Hoeijmakers, David A. Koolen, Lia C. M. J. Goltstein, Alexander Hoischen, Han G. Brunner, et al. 2019. “Next-Generation Phenotyping Using Computer Vision Algorithms in Rare Genomic Neurodevelopmental Disorders.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 21 (8): 1719–25.

Down, John Langdon Haydon. 1866. “Observations on an Ethnic Classification of Idiots.” *London Hospital Reports* 3: 259–162.

Druschke, D., F. Krause, G. Müller, J. Scharfe, G. F. Hoffmann, J. Schmitt, and TRANSLATE-NAMSE-Consortium. 2021. “Potentials and Current Shortcomings in the Cooperation between German Centers for Rare Diseases and Primary Care Physicians: Results from the Project TRANSLATE-NAMSE.” *Orphanet Journal of Rare Diseases* 16 (1): 494.

Dudding-Byth, Tracy, Anne Baxter, Elizabeth G. Holliday, Anna Hackett, Sheridan O’Donnell, Susan M. White, John Attia, et al. 2017. “Computer Face-Matching Technology Using Two-Dimensional Photographs Accurately Matches the Facial Gestalt of Unrelated Individuals with the Same Syndromic Form of Intellectual Disability.” *BMC Biotechnology* 17 (1): 90.

Elborn, J. Stuart. 2016. “Cystic Fibrosis.” *The Lancet* 388 (10059): 2519–31.

Elsner, Jonas, Martin A. Mensah, Manuel Holtgrewe, Jakob Hertzberg, Stefania Bigoni, Andreas Busche, Marie Coutelier, et al. 2021. “Genome Sequencing in Families with Congenital Limb Malformations.” *Human Genetics* 140 (8): 1229–39.

Esteva, Andre, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. 2019. “A Guide to Deep Learning in Healthcare.” *Nature Medicine* 25 (1): 24–29.

Ferreira, Carlos R. 2019. “The Burden of Rare Diseases.” *American Journal of Medical Genetics. Part A* 179 (6): 885–92.

Ferreira, Carlos R., Ruqaia Altassan, Dorinda Marques-Da-Silva, Rita Francisco, Jaak Jaeken, and Eva Morava. 2018. “Recognizable Phenotypes in CDG.” *Journal of Inherited Metabolic Disease* 41 (3): 541–53.

Ferry, Quentin, Julia Steinberg, Caleb Webber, David R. FitzPatrick, Chris P. Ponting, Andrew Zisserman, and Christoffer Nellåker. 2014. “Diagnostically Relevant Facial Gestalt Information from Ordinary Photos.” *eLife* 3 (June): e02020.

GBD 2021 Sickle Cell Disease Collaborators. 2023. “Global, Regional, and National Prevalence and Mortality Burden of Sickle Cell Disease, 2000-2021: A Systematic Analysis from the Global Burden of Disease Study 2021.” *The Lancet. Haematology*, June. https://doi.org/10.1016/S2352-3026(23)00118-7.

Gene Ontology Consortium, Suzi A. Aleksander, James Balhoff, Seth Carbon, J. Michael Cherry, Harold J. Drabkin, Dustin Ebert, et al. 2023. “The Gene Ontology Knowledgebase in 2023.” *Genetics* 224 (1). https://doi.org/10.1093/genetics/iyad031.

Gilissen, Christian, Alexander Hoischen, Han G. Brunner, and Joris A. Veltman. 2012. “Disease Gene Identification Strategies for Exome Sequencing.”

European Journal of Human Genetics: EJHG 20 (5): 490–97.

Goede, Christian de, Wyatt W. Yue, Guanhua Yan, Shyamala Ariyaratnam, Kate E. Chandler, Laura Downes, Nasaim Khan, Meyyammai Mohan, Martin Lowe, and Siddharth Banka. 2016. “Role of Reverse Phenotyping in Interpretation of next Generation Sequencing Data and a Review of INPP5E Related Disorders.” *European Journal of Paediatric Neurology: EJPN: Official Journal of the European Paediatric Neurology Society* 20 (2): 286–95.

Gordeeva, Veronika, Elena Sharova, Konstantin Babalyan, Rinat Sultanov, Vadim M. Govorun, and Georgij Arapidi. 2021. “Benchmarking Germline CNV Calling Tools from Exome Sequencing Data.” *Scientific Reports* 11 (1): 14416.

Gottschalk, Annika, Henrike L. Sczakiel, Wiebke Hülsemann, Sarina Schwartzmann, Angela T. Abad-Perez, Johannes Grünhagen, Claus-Eric Ott, et al. 2023. “HOXD13-Associated Synpolydactyly: Extending and Validating the Genotypic and Phenotypic Spectrum with 38 New and 49 Published Families.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 25 (11): 100928.

Gudmundsson, Sanna, Moriel Singer-Berk, Nicholas A. Watts, William Phu, Julia K. Goodrich, Matthew Solomonson, Genome Aggregation Database Consortium, Heidi L. Rehm, Daniel G. MacArthur, and Anne O’Donnell-Luria. 2022. “Variant Interpretation Using Population Databases: Lessons from gnomAD.” *Human Mutation* 43 (8): 1012–30.

Guo, Ruiji, Xia Fang, Hailei Mao, Bin Sun, Jiateng Zhou, Yu An, and Bin Wang. 2021. “A Novel Missense Variant of Caused Atypical Synpolydactyly by Impairing the Downstream Gene Expression and Literature Review for Genotype-Phenotype Correlations.” *Frontiers in Genetics* 12 (October): 731278.

Gurovich, Yaron, Yair Hanani, Omri Bar, Guy Nadav, Nicole Fleischer, Dekel Gelbman, Lina Basel-Salmon, et al. 2019. “Identifying Facial Phenotypes of Genetic Disorders Using Deep Learning.” *Nature Medicine* 25 (1): 60–64.

Haendel, Melissa A., Christopher G. Chute, and Peter N. Robinson. 2018. “Classification, Ontology, and Precision Medicine.” *The New England Journal of Medicine* 379 (15): 1452–62.

Halldorsson, Bjarni V., Hannes P. Eggertsson, Kristjan H. S. Moore, Hannes Hauswedell, Ogmundur Eiriksson, Magnus O. Ulfarsson, Gunnar Palsson, et al. 2022. “The Sequences of 150,119 Genomes in the UK Biobank.” *Nature* 607 (7920): 732–40.

Hallgrímsson, Benedikt, J. David Aponte, David C. Katz, Jordan J. Bannister, Sheri L. Riccardi, Nick Mahasuwan, Brenda L. McInnes, et al. 2020. “Automated Syndrome Diagnosis by Three-Dimensional Facial Imaging.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 22 (10): 1682–93.

Harris, Midori A. 2008. “Developing an Ontology.” *Methods in Molecular Biology* 452: 111–24.

Hashimoto, Yoshitami, Kota Sadano, Nene Miyata, Haruka Ito, and Hirofumi Tanaka. 2023. “Novel Role of DONSON in CMG Helicase Assembly during Vertebrate DNA Replication Initiation.” *The EMBO Journal* 42 (17): e114131.

Hsieh, Tzung-Chien, Aviram Bar-Haim, Shahida Moosa, Nadja Ehmke, Karen W. Gripp, Jean Tori Pantel, Magdalena Danyel, et al. 2022. “GestaltMatcher Facilitates Rare Disease Matching Using Facial Phenotype Descriptors.” *Nature Genetics* 54 (3): 349–57.

Hsieh, Tzung-Chien, Martin A. Mensah, Jean T. Pantel, Dione Aguilar, Omri Bar, Allan Bayat, Luis Becerra-Solano, et al. 2019. “PEDIA: Prioritization of Exome Data by Image Analysis.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 21 (12): 2807–14.

Huang, Xiu-Feng, Fang Huang, Kun-Chao Wu, Juan Wu, Jie Chen, Chi-Pui Pang, Fan Lu, Jia Qu, and Zi-Bing Jin. 2015. “Genotype-Phenotype Correlation and Mutation Spectrum in a Large Cohort of Patients with Inherited Retinal Dystrophy Revealed by next-Generation Sequencing.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 17 (4): 271–78.

Ibañez, Kristina, James Polke, R. Tanner Hagelstrom, Egor Dolzhenko, Dorota Pasko, Ellen Rachel Amy Thomas, Louise C. Daugherty, et al. 2022. “Whole Genome Sequencing for the Diagnosis of Neurological Repeat Expansion Disorders in the UK: A Retrospective Diagnostic Accuracy and Prospective Clinical Validation Study.” *Lancet Neurology* 21 (3): 234–45.

Iglesias, Alejandro, Kwame Anyane-Yeboa, Julia Wynn, Ashley Wilson, Megan Truitt Cho, Edwin Guzman, Rebecca Sisson, Claire Egan, and Wendy K. Chung. 2014. “The Usefulness of Whole-Exome Sequencing in Routine Clinical Practice.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 16 (12): 922–31.

INSERM, and Orphanet consortium. 1999. “Orphadata: Free Access Data from Orphanet.” 1999. http://www.orphadata.org/cgi-bin/index.php., zuletzt abgerufen am 30.09.2022

Jackson, Maria, Leah Marks, Gerhard H. W. May, and Joanna B. Wilson. 2018. “The Genetic Basis of Disease.” *Essays in Biochemistry* 62 (5): 643–723.

Jacobsen, Julius O. B., Catherine Kelly, Valentina Cipriani, Peter N. Robinson, and Damian Smedley. 2022. “Evaluation of Phenotype-Driven Gene Prioritization Methods for Mendelian Diseases.” *Briefings in Bioinformatics* 23 (5). https://doi.org/10.1093/bib/bbac188.

Jiang, Yue, Andrei L. Turinsky, and Michael Brudno. 2015. “The Missing Indels: An Estimate of Indel Variation in a Human Genome and Analysis of Factors That Impede Detection.” *Nucleic Acids Research* 43 (15): 7217–28.

Johnson, Britt, Karen Ouyang, Lauren Frank, Rebecca Truty, Susan Rojahn, Ana Morales, Swaroop Aradhya, and Keith Nykamp. 2022. “Systematic Use of Phenotype Evidence in Clinical Genetic Testing Reduces the Frequency of Variants of Uncertain Significance.” *American Journal of Medical Genetics. Part A* 188 (9): 2642–51.

Julkowska, Daria, and EJP RD Initiative. 2019. “Project Structure - EJP-RD.” Ejprarediseases.org. 2019. https://www.ejprarediseases.org/what-is-ejprd/project-structure/., zuletzt abgerufen am 29.10.2023

Karaca, Ender, Jennifer E. Posey, Bret Bostwick, Pengfei Liu, Alper Gezdirici, Gozde Yesil, Zeynep Coban Akdemir, et al. 2019. “Biallelic and De Novo Variants in DONSON Reveal a Clinical Spectrum of Cell Cycle-Opathies with Microcephaly, Dwarfism and Skeletal Abnormalities.” *American Journal of Medical Genetics. Part A* 179 (10): 2056–66.

Karczewski, Konrad J., Laurent C. Francioli, Grace Tiao, Beryl B. Cummings, Jessica Alföldi, Qingbo Wang, Ryan L. Collins, et al. 2020. “The Mutational Constraint Spectrum Quantified from Variation in 141,456 Humans.” *Nature* 581 (7809): 434–43.

Kasianowicz, J. J., E. Brandin, D. Branton, and D. W. Deamer. 1996. “Characterization of Individual Polynucleotide Molecules Using a Membrane Channel.” *Proceedings of the National Academy of Sciences of the United*

States of America 93 (24): 13770–73.

Kazazian, Haig H., Corinne D. Boehm, and William K. Seltzer. 2000. “ACMG Recommendations for Standards for Interpretation of Sequence Variations.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 2 (5): 302–3.

Kernohan, Kristin D., Taila Hartley, Najmeh Alirezaie, Care4Rare Canada Consortium, Peter N. Robinson, David A. Dyment, and Kym M. Boycott. 2018. “Evaluation of Exome Filtering Techniques for the Analysis of Clinically Relevant Genes.” *Human Mutation* 39 (2): 197–201.

King, Daniel A., Alejandro Sifrim, Tomas W. Fitzgerald, Raheleh Rahbari, Emma Hobson, Tessa Homfray, Sahar Mansour, et al. 2017. “Detection of Structural Mosaicism from Targeted and Whole-Genome Sequencing Data.” *Genome Research* 27 (10): 1704–14.

Knapp, Karen M., Rosie Sullivan, Jennie Murray, Gregory Gimenez, Pamela Arn, Precilla D’Souza, Alper Gezdirici, et al. 2020. “Linked-Read Genome Sequencing Identifies Biallelic Pathogenic Variants in as a Novel Cause of Meier-Gorlin Syndrome.” *Journal of Medical Genetics* 57 (3): 195–202.

Knaus, Alexej, Jean Tori Pantel, Manuela Pendziwiat, Nurulhuda Hajjir, Max Zhao, Tzung-Chien Hsieh, Max Schubach, et al. 2018. “Characterization of Glycosylphosphatidylinositol Biosynthesis Defects by Clinical Features, Flow Cytometry, and Automated Image Analysis.” *Genome Medicine* 10 (1): 3.

Koboldt, Daniel C. 2020. “Best Practices for Variant Calling in Clinical Sequencing.” *Genome Medicine* 12 (1): 91.

Koboldt, Daniel C., Karyn Meltz Steinberg, David E. Larson, Richard K. Wilson, and Elaine R. Mardis. 2013. “The next-Generation Sequencing Revolution and Its Impact on Genomics.” *Cell* 155 (1): 27–38.

Köhler, Sebastian, Michael Gargano, Nicolas Matentzoglu, Leigh C. Carmody, David Lewis-Smith, Nicole A. Vasilevsky, Daniel Danis, et al. 2021. “The Human Phenotype Ontology in 2021.” *Nucleic Acids Research* 49 (D1): D1207–17.

Köhler, Sebastian, N. Christine Øien, Orion J. Buske, Tudor Groza, Julius O. B. Jacobsen, Craig McNamara, Nicole Vasilevsky, et al. 2019. “Encoding Clinical Data with the Human Phenotype Ontology for Computational Differential Diagnostics.” *Current Protocols in Human Genetics* 103 (1): e92.

Kruszka, Paul, Tommy Hu, Sungkook Hong, Rebecca Signer, Benjamin Cogné, Betrand Isidor, Sarah E. Mazzola, et al. 2019. “Phenotype Delineation of ZNF462 Related Syndrome.” *American Journal of Medical Genetics. Part A* 179 (10): 2075–82.

Landini, Samuela, Benedetta Mazzinghi, Francesca Becherucci, Marco Allinovi, Aldesia Provenzano, Viviana Palazzo, Fiammetta Ravaglia, et al. 2020. “Reverse Phenotyping after Whole-Exome Sequencing in Steroid-Resistant Nephrotic Syndrome.” *Clinical Journal of the American Society of Nephrology: CJASN* 15 (1): 89–100.

Lappalainen, Tuuli, Alexandra J. Scott, Margot Brandt, and Ira M. Hall. 2019. “Genomic Analysis in the Age of Human Genome Sequencing.” *Cell* 177 (1): 70–84.

Lek, Monkol, Konrad J. Karczewski, Eric V. Minikel, Kaitlin E. Samocha, Eric Banks, Timothy Fennell, Anne H. O’Donnell-Luria, et al. 2016. “Analysis of Protein-Coding Genetic Variation in 60,706 Humans.” *Nature* 536 (7616): 285–91.

Lesmann, Hellen, Hannah Klinkhammer, and Peter M. Krawitz. 2023. “The Future Role of Facial Image Analysis in ACMG Classification Guidelines.” *Medizinische*

Genetik: Mitteilungsblatt Des Berufsverbandes Medizinische Genetik e.V 35 (2): 115–21.

Liehr, T., N. Acquarola, K. Pyle, S. St-Pierre, M. Rinholm, O. Bar, K. Wilhelm, and I. Schreyer. 2018. “Next Generation Phenotyping in Emanuel and Pallister-Killian Syndrome Using Computer-Aided Facial Dysmorphology Analysis of 2D Photos.” *Clinical Genetics* 93 (2): 378–81.

Los, Evan, and George A. Ford. 2023. “Klinefelter Syndrome.” In *StatPearls*. Treasure Island (FL): StatPearls Publishing.

Lumaka, A., N. Cosemans, A. Lulebo Mampasi, G. Mubungu, N. Mvuama, T. Lubala, S. Mbuyi-Musanzayi, et al. 2017. “Facial Dysmorphism Is Influenced by Ethnic Background of the Patient and of the Evaluator.” *Clinical Genetics* 92 (2): 166–71.

Mak, Bryan C., Rossana Sanchez Russo, Michael J. Gambello, Nicole Fleischer, Emily D. Black, Elizabeth Leslie, Melissa M. Murphy, Emory 3q29 Project, and Jennifer Gladys Mulle. 2021. “Craniofacial Features of 3q29 Deletion Syndrome: Application of next-Generation Phenotyping Technology.” *American Journal of Medical Genetics. Part A* 185 (7): 2094–2101.

Manheimer, Kathryn B., Nihir Patel, Felix Richter, Joshua Gorham, Angela C. Tai, Jason Homsy, Marko T. Boskovski, et al. 2018. “Robust Identification of Deletions in Exome and Genome Sequence Data Based on Clustering of Mendelian Errors.” *Human Mutation* 39 (6): 870–81.

Mañú Pereira, María Del Mar, Raffaella Colombatti, Federico Alvarez, Pablo Bartolucci, Celeste Bento, Angelo Loris Brunetta, Elena Cela, et al. 2023. “Sickle Cell Disease Landscape and Challenges in the EU: The ERN-EuroBloodNet Perspective.” *The Lancet. Haematology*, July. https://doi.org/10.1016/S2352-3026(23)00182-5.

Martinez-Monseny, Antonio, Daniel Cuadras, Mercè Bolasell, Jordi Muchart, César Arjona, Mar Borregan, Adi Algrabli, et al. 2019. “From Gestalt to Gene: Early Predictive Dysmorphic Features of PMM2-CDG.” *Journal of Medical Genetics* 56 (4): 236–45.

Marwaha, Ashish, David Chitayat, M. Stephen Meyn, Roberto Mendoza-Londono, and Lauren Chad. 2021. “The Point-of-Care Use of a Facial Phenotyping Tool in the Genetics Clinic: Enhancing Diagnosis and Education with Machine Learning.” *American Journal of Medical Genetics. Part A* 185 (4): 1151–58.

Marwaha, Shruti, Joshua W. Knowles, and Euan A. Ashley. 2022. “A Guide for the Diagnosis of Rare and Undiagnosed Disease: Beyond the Exome.” *Genome Medicine* 14 (1): 23.

McPherson, J. D., M. Marra, L. Hillier, R. H. Waterston, A. Chinwalla, J. Wallis, M. Sekhon, et al. 2001. “A Physical Map of the Human Genome.” *Nature* 409 (6822): 934–41.

Mensah, Martin A., Henri Niskanen, Alexandre P. Magalhaes, Shaon Basu, Martin Kircher, Henrike L. Sczakiel, Alisa M. V. Reiter, et al. 2023. “Aberrant Phase Separation and Nucleolar Dysfunction in Rare Genetic Diseases.” *Nature* 614 (7948): 564–71.

Mensah, Martin A., Claus-Eric Ott, Denise Horn, and Jean T. Pantel. 2022. “A Machine Learning-Based Screening Tool for Genetic Syndromes in Children.” *The Lancet. Digital Health*.

Mishima, Hiroyuki, Hisato Suzuki, Michiko Doi, Mutsuko Miyazaki, Satoshi Watanabe, Tadashi Matsumoto, Kanako Morifuji, et al. 2019. “Evaluation of Face2Gene Using Facial Images of Patients with Congenital Dysmorphic

Syndromes Recruited in Japan.” *Journal of Human Genetics* 64 (8): 789–94.

Moliner, Antoni Montserrat, and Jaroslaw Waligora. 2017. “The European Union Policy in the Field of Rare Diseases.” *Advances in Experimental Medicine and Biology* 1031: 561–87.

Musante, Luciana, Paola Costa, Caterina Zanus, Flavio Faletra, Flora M. Murru, Anna M. Bianco, Martina La Bianca, et al. 2022. “The Genetic Diagnosis of Ultrarare DEEs: An Ongoing Challenge.” *Genes* 13 (3). https://doi.org/10.3390/genes13030500.

Myers, Lynnea, Britt-Marie Anderlid, Ann Nordgren, Karl Lundin, Ralf Kuja-Halkola, Kristiina Tammimies, and Sven Bölte. 2020. “Clinical versus Automated Assessments of Morphological Variants in Twins with and without Neurodevelopmental Disorders.” *American Journal of Medical Genetics. Part A* 182 (5): 1177–89.

Nielsen-Dandoroff, Emily, Mischa S. G. Ruegg, and Louise S. Bicknell. 2023. “The Expanding Genetic and Clinical Landscape Associated with Meier-Gorlin Syndrome.” *European Journal of Human Genetics: EJHG* 31 (8): 859–68.

Ong, Thida, and Bonnie W. Ramsey. 2023. “Cystic Fibrosis: A Review.” *JAMA: The Journal of the American Medical Association* 329 (21): 1859–71.

Pantel, Jean Tori, Nurulhuda Hajjir, Magdalena Danyel, Jonas Elsner, Angela Teresa Abad-Perez, Peter Hansen, Stefan Mundlos, et al. 2020. “Efficiency of Computer-Aided Facial Phenotyping (DeepGestalt) in Individuals With and Without a Genetic Syndrome: Diagnostic Accuracy Study.” *Journal of Medical Internet Research* 22 (10): e19263.

Pantel, Jean T., Max Zhao, Martin A. Mensah, Nurulhuda Hajjir, Tzung-Chien Hsieh, Yair Hanani, Nicole Fleischer, et al. 2018. “Advances in Computer-Assisted Syndrome Recognition by the Example of Inborn Errors of Metabolism.” *Journal of Inherited Metabolic Disease* 41 (3): 533–39.

Pascolini, Giulia, Nicole Fleischer, Alessandro Ferraris, Silvia Majore, and Paola Grammatico. 2019. “The Facial Dysmorphology Analysis Technology in Intellectual Disability Syndromes Related to Defects in the Histones Modifiers.” *Journal of Human Genetics* 64 (8): 721–28.

Petrovski, Slavé, Vimla Aggarwal, Jessica L. Giordano, Melissa Stosic, Karen Wou, Louise Bier, Erica Spiegel, et al. 2019. “Whole-Exome Sequencing in the Evaluation of Fetal Structural Anomalies: A Prospective Cohort Study.” *The Lancet* 393 (10173): 758–67.

Pillay, Nikita Simone, Owen A. Ross, Alan Christoffels, and Soraya Bardien. 2022. “Current Status of Next-Generation Sequencing Approaches for Candidate Gene Discovery in Familial Parkinson´s Disease.” *Frontiers in Genetics* 13 (March): 781816.

Pollard, Katherine S., Melissa J. Hubisz, Kate R. Rosenbloom, and Adam Siepel. 2010. “Detection of Nonneutral Substitution Rates on Mammalian Phylogenies.” *Genome Research* 20 (1): 110–21.

Popejoy, Alice B., and Stephanie M. Fullerton. 2016. “Genomics Is Failing on Diversity.” *Nature*.

Porras, Antonio R., Kenneth Rosenbaum, Carlos Tor-Diez, Marshall Summar, and Marius George Linguraru. 2021. “Development and Evaluation of a Machine Learning-Based Point-of-Care Screening Tool for Genetic Syndromes in Children: A Multinational Retrospective Study.” *The Lancet. Digital Health* 3 (10): e635–43.

Richards, C. Sue, Sherri Bale, Daniel B. Bellissimo, Soma Das, Wayne W. Grody, Madhuri R. Hegde, Elaine Lyon, Brian E. Ward, and Molecular Subcommittee of the ACMG Laboratory Quality Assurance Committee. 2008. “ACMG Recommendations for Standards for Interpretation and Reporting of Sequence Variations: Revisions 2007.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 10 (4): 294–300.

Richards, Sue, Nazneen Aziz, Sherri Bale, David Bick, Soma Das, Julie Gastier-Foster, Wayne W. Grody, et al. 2015. “Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 17 (5): 405–24.

Richter, Trevor, Sandra Nestler-Parr, Robert Babela, Zeba M. Khan, Theresa Tesoro, Elizabeth Molsen, Dyfrig A. Hughes, and International Society for Pharmacoeconomics and Outcomes Research Rare Disease Special Interest Group. 2015. “Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group.” *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research* 18 (6): 906–14.

Rillig, Franziska, Annette Grüters, Tobias Bäumer, Georg F. Hoffmann, Daniela Choukair, Reinhard Berner, Min Ae Lee-Kirsch, et al. 2022. “The Interdisciplinary Diagnosis of Rare Diseases–Results of the Translate-NAMSE Project.” *Deutsches Arzteblatt International*, no. Forthcoming (July). https://doi.org/10.3238/arztebl.m2022.0219.

Rishishwar, Lavanya, Neha Varghese, Eishita Tyagi, Stephen C. Harvey, I. King Jordan, and Nael A. McCarty. 2012. “Relating the Disease Mutation Spectrum to the Evolution of the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR).” *PloS One* 7 (8): e42336.

Robinson, Peter N., Sebastian Köhler, Sebastian Bauer, Dominik Seelow, Denise Horn, and Stefan Mundlos. 2008. “The Human Phenotype Ontology: A Tool for Annotating and Analyzing Human Hereditary Disease.” *American Journal of Human Genetics* 83 (5): 610–15.

Robinson, Peter N., Sebastian Köhler, Anika Oellrich, Sanger Mouse Genetics Project, Kai Wang, Christopher J. Mungall, Suzanna E. Lewis, et al. 2014. “Improved Exome Prioritization of Disease Genes through Cross-Species Phenotype Comparison.” *Genome Research* 24 (2): 340–48.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. “DNA Sequencing with Chain-Terminating Inhibitors.” *Proceedings of the National Academy of Sciences of the United States of America* 74 (12): 5463–67.

Schulze, Thomas G., and Francis J. McMahon. 2004. “Defining the Phenotype in Human Genetic Studies: Forward Genetics and Reverse Phenotyping.” *Human Heredity* 58 (3-4): 131–38.

Seltzsam, Steve, Chunyan Wang, Bixia Zheng, Nina Mann, Dervla M. Connaughton, Chen-Han Wilfred Wu, Sophia Schneider, et al. 2022. “Reverse Phenotyping Facilitates Disease Allele Calling in Exome Sequencing of Patients with CAKUT.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 24 (2): 307–18.

Shamseldin, Hanan E., Sateesh Maddirevula, Eissa Faqeih, Niema Ibrahim, Mais Hashem, Ranad Shaheen, and Fowzan S. Alkuraya. 2017. “Increasing the Sensitivity of Clinical Exome Sequencing through Improved Filtration Strategy.” *Genetics in Medicine: Official Journal of the American College of Medical*

Genetics 19 (5): 593–98.

Shendure, Jay. 2011. “Next-Generation Human Genetics.” *Genome Biology* 12 (9): 408.

Shickh, Salma, Chloe Mighton, Elizabeth Uleryk, Petros Pechlivanoglou, and Yvonne Bombard. 2021. “The Clinical Utility of Exome and Genome Sequencing across Clinical Indications: A Systematic Review.” *Human Genetics* 140 (10): 1403–16.

Smedley, Damian, and Peter N. Robinson. 2015. “Phenotype-Driven Strategies for Exome Prioritization of Human Mendelian Disease Genes.” *Genome Medicine* 7 (1): 81.

Snel, B., G. Lehmann, P. Bork, and M. A. Huynen. 2000. “STRING: A Web-Server to Retrieve and Display the Repeatedly Occurring Neighbourhood of a Gene.” *Nucleic Acids Research* 28 (18): 3442–44.

Solomon, Benjamin D., Margaret P. Adam, Chin-To Fong, Katta M. Girisha, Judith G. Hall, Anna C. E. Hurst, Peter M. Krawitz, et al. 2023. “Perspectives on the Future of Dysmorphology.” *American Journal of Medical Genetics. Part A* 191 (3): 659–71.

Srisraluang, Wewika, and Kitiwan Rojnueangnit. 2021. “Facial Recognition Accuracy in Photographs of Thai Neonates with Down Syndrome among Physicians and the Face2Gene Application.” *American Journal of Medical Genetics. Part A* 185 (12): 3701–5.

Staufner, Christian, Bianca Peters, Matias Wagner, Seham Alameer, Ivo Barić, Pierre Broué, Derya Bulut, et al. 2020. “Defining Clinical Subgroups and Genotype-Phenotype Correlations in NBAS-Associated Disease across 110 Patients.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 22 (3): 610–21.

Stoddart, David, Andrew J. Heron, Ellina Mikhailova, Giovanni Maglia, and Hagan Bayley. 2009. “Single-Nucleotide Discrimination in Immobilized DNA Oligonucleotides with a Biological Nanopore.” *Proceedings of the National Academy of Sciences of the United States of America* 106 (19): 7702–7.

Stranneheim, Henrik, Kristina Lagerstedt-Robinson, Måns Magnusson, Malin Kvarnung, Daniel Nilsson, Nicole Lesko, Martin Engvall, et al. 2021. “Integration of Whole Genome Sequencing into a Healthcare Setting: High Diagnostic Rates across Multiple Clinical Entities in 3219 Rare Disease Patients.” *Genome Medicine* 13 (1): 40.

Swietlik, Emilia M., Matina Prapa, Jennifer M. Martin, Divya Pandya, Kathryn Auckland, Nicholas W. Morrell, and Stefan Gräf. 2020. “‘There and Back Again’-Forward Genetics and Reverse Phenotyping in Pulmonary Arterial Hypertension.” *Genes* 11 (12). https://doi.org/10.3390/genes11121408.

Szklarczyk, Damian, Rebecca Kirsch, Mikaela Koutrouli, Katerina Nastou, Farrokh Mehryary, Radja Hachilif, Annika L. Gable, et al. 2023. “The STRING Database in 2023: Protein-Protein Association Networks and Functional Enrichment Analyses for Any Sequenced Genome of Interest.” *Nucleic Acids Research* 51 (D1): D638–46.

Tang, Haibao, Ewen F. Kirkness, Christoph Lippert, William H. Biggs, Martin Fabani, Ernesto Guzman, Smriti Ramakrishnan, et al. 2017. “Profiling of Short-Tandem-Repeat Disease Alleles in 12,632 Human Whole Genomes.” *American Journal of Human Genetics* 101 (5): 700–715.

Tekendo-Ngongang, Cedrik, and Paul Kruszka. 2020. “Noonan Syndrome on the African Continent.” *Birth Defects Research* 112 (10): 718–24.

Todd, Emily J., Kyle S. Yau, Royston Ong, Jennie Slee, George McGillivray, Christopher P. Barnett, Goknur Haliloglu, et al. 2015. “Next Generation Sequencing in a Large Cohort of Patients Presenting with Neuromuscular Disease before or at Birth.” *Orphanet Journal of Rare Diseases* 10 (November): 148.

Tu, Liyun, Antonio R. Porras, Alec Boyle, and Marius George Linguraru. 2018. “Analysis of 3D Facial Dysmorphology in Genetic Syndromes from Unconstrained 2D Photographs.” In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 347–55. Lecture Notes in Computer Science. Cham: Springer International Publishing.

Turro, Ernest, William J. Astle, Karyn Megy, Stefan Gräf, Daniel Greene, Olga Shamardina, Hana Lango Allen, et al. 2020. “Whole-Genome Sequencing of Patients with Rare Diseases in a National Health System.” *Nature* 583 (7814): 96–102.

Uliana, Vera, and Antonio Percesepe. 2016. “Reverse Phenotyping Comes of Age.” *Molecular Genetics and Metabolism*.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, et al. 2001. “The Sequence of the Human Genome.” *Science* 291 (5507): 1304–51.

Vollmar, Tobias, Baerbel Maus, Rolf P. Wurtz, Gabriele Gillessen-Kaesbach, Bernhard Horsthemke, Dagmar Wieczorek, and Stefan Boehringer. 2008. “Impact of Geometry and Viewing Angle on Classification Accuracy of 2D Based Analysis of Dysmorphic Faces.” *European Journal of Medical Genetics* 51 (1): 44–53.

Vorravanpreecha, Nattariya, Thanayoot Lertboonnum, Rungrote Rodjanadit, Pak Sriplienchan, and Kitiwan Rojnueangnit. 2018. “Studying Down Syndrome Recognition Probabilities in Thai Children with de-Identified Computer-Aided Facial Analysis.” *American Journal of Medical Genetics. Part A* 176 (9): 1935–40.

Walker, Caroline E., Trinity Mahede, Geoff Davis, Laura J. Miller, Jennifer Girschik, Kate Brameld, Wenxing Sun, et al. 2017. “The Collective Impact of Rare Diseases in Western Australia: An Estimate Using a Population-Based Cohort.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 19 (5): 546–52.

Wang, Jinlian, Jun Liao, Jinglan Zhang, Wei-Yi Cheng, Jörg Hakenberg, Meng Ma, Bryn D. Webb, et al. 2015. “ClinLabGeneticist: A Tool for Clinical Management of Genetic Variants from Whole Exome Sequencing in Clinical Genetic Laboratories.” *Genome Medicine* 7 (July): 77.

Weißbach, Stephan, Stanislav Sys, Charlotte Hewel, Hristo Todorov, Susann Schweiger, Jennifer Winter, Markus Pfenninger, et al. 2021. “Reliability of Genomic Variants across Different next-Generation Sequencing Platforms and Bioinformatic Processing Pipelines.” *BMC Genomics* 22 (1): 62.

Weiss, Karin, Hayley P. Lazar, Alina Kurolap, Ariel F. Martinez, Tamar Paperna, Lior Cohen, Marie F. Smeland, et al. 2020. “The CHD4-Related Syndrome: A Comprehensive Investigation of the Clinical Spectrum, Genotype-Phenotype Correlations, and Molecular Basis.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 22 (2): 389–97.

Wells, Constance F., Guilaine Boursier, Kevin Yauy, Nathalie Ruiz-Pallares, Déborah Mechin, Valentin Ruault, Mylène Tharreau, et al. 2022. “Rapid Exome Sequencing in Critically Ill Infants: Implementation in Routine Care from French Regional Hospital’s Perspective.” *European Journal of Human Genetics: EJHG*

30 (9): 1076–82.

Wong, Karen H. Y., Walfred Ma, Chun-Yu Wei, Erh-Chan Yeh, Wan-Jia Lin, Elin H. F. Wang, Jen-Ping Su, et al. 2020. “Towards a Reference Genome That Captures Global Genetic Diversity.” *Nature Communications* 11 (1): 5482.

Wright, Caroline F., Patrick Campbell, Ruth Y. Eberhardt, Stuart Aitken, Daniel Perrett, Simon Brent, Petr Danecek, et al. 2023. “Genomic Diagnosis of Rare Pediatric Disease in the United Kingdom and Ireland.” *The New England Journal of Medicine* 388 (17): 1559–71.

Wright, Caroline F., Tomas W. Fitzgerald, Wendy D. Jones, Stephen Clayton, Jeremy F. McRae, Margriet van Kogelenberg, Daniel A. King, et al. 2015. “Genetic Diagnosis of Developmental Disorders in the DDD Study: A Scalable Analysis of Genome-Wide Research Data.” *The Lancet* 385 (9975): 1305–14.

Yan, Huihuang, Shulan Tian, Susan L. Slager, Zhifu Sun, and Tamas Ordog. 2016. “Genome-Wide Epigenetic Studies in Human Disease: A Primer on -Omic Technologies.” *American Journal of Epidemiology* 183 (2): 96–109.

Zhang, Qianwen, Yu Ding, Biyun Feng, Yijun Tang, Yao Chen, Yirou Wang, Guoying Chang, et al. 2022. “Molecular and Phenotypic Expansion of Alström Syndrome in Chinese Patients.” *Frontiers in Genetics* 13 (February): 808919.

Zhao, Xuefang, Ryan L. Collins, Wan-Ping Lee, Alexandra M. Weber, Yukyung Jun, Qihui Zhu, Ben Weisburd, et al. 2021. “Expectations and Blind Spots for Structural Variation Detection from Long-Read Assemblies and Short-Read Genome Sequencing Technologies.” *American Journal of Human Genetics* 108 (5): 919–28.

Danksagung

Herzlich danken möchte ich Prof. Dr. Denise Horn für ihre hervorragende Betreuung bei den verschiedenen Projekten, die zu dieser Arbeit beigetragen haben, und für die exzellente Anleitung während meiner klinischen Weiterbildung. Danken möchte ich auch Prof. Dr. Malte Spielmann, der mich mit seiner Leidenschaft für die Humangenetik angesteckt hat und mir stets mit einem Rat zur Seite stand, wenn ich ihn brauchte. Prof. Dr. Peter Krawitz danke ich dafür, dass er mich auf die faziale Phänotypisierung als ein mögliches Forschungsgebiet aufmerksam gemacht und immer das Beste aus mir herausgeholt hat. Prof. Dr. Stefan Mundlos danke ich, dass er mir die Mitarbeit an seiner Extremitätenstudie ermöglichte und mir dabei die Freiheit ließ, die es für diese Habilitation brauchte.

Dr. Solveig Schulz, Dr. Magdalena Danyel, Dr. Tzung-Chien Hsieh, Tori Pantel, Nurulhuda Hajjir und Jonas Elsner, Dr. Henri Niskanen sowie allen Koautoren danke ich für die gute und vertrauensvolle Zusammenarbeit bei der Erstellung der hier aufgeführten Arbeiten.

Valerie Johnston, Gabriele Hildebrandt und Susanne Rothe danke ich für die technische Unterstützung im Labor.

Den Kollegen vom Institut für Medizinische Genetik und Humangenetik danke ich für die freundschaftliche Atmosphäre, ohne die eine solche Arbeit nicht möglich gewesen wäre.

Die Forschungsförderung im Rahmen des von Prof. Dr. Duska Dragun gegründeten Junior Clinician Scientist Programms und des Digital Clinician Scientist Programms des Berlin Institute of Health ist von unschätzbarem Wert gewesen, dafür werde ich immer dankbar sein. Prof. Dr. Dominik Seelow danke ich für die Begleitung und die guten Ratschläge, die mir eine Teilnahme daran ermöglichten.

Nicole Fleischer und Yaron Gurovich von FDNA danke ich für ihre geduldigen Erläuterungen der Funktionsweise von DeepGestalt.

Den Patienten und ihren Angehörigen danke ich für ihre Teilnahme an den hier aufgeführten Studien.

Meiner Frau, unseren Kindern, meinen Eltern, und meinen Brüdern danke ich, dass sie immer an mich geglaubt haben und stets für mich da sind, wenn ich sie brauche.

Erklärung

Erklärung § 4 Abs. 3 (k) der HabOMed der Charité

Hiermit erkläre ich, dass

- weder früher noch gleichzeitig ein Habilitationsverfahren durchgeführt oder angemeldet wurde,
- die vorgelegte Habilitationsschrift ohne fremde Hilfe verfasst, die beschriebenen Ergebnisse selbst gewonnen sowie die verwendeten Hilfsmittel, die Zusammenarbeit mit anderen Wissenschaftlern/Wissenschaftlerinnen und mit technischen Hilfskräften sowie die verwendete Literatur vollständig in der Habilitationsschrift angegeben wurden,
- mir die geltende Habilitationsordnung bekannt ist.

Ich erkläre ferner, dass mir die Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis bekannt ist und ich mich zur Einhaltung dieser Satzung verpflichte.

Berlin, ……………………….. ……………………………...

Datum Dr. med. Martin A. Mensah