

## Die Stunde der unerforschten Gensegmente

**Ein internationales Forschungskonsortium stellt einen Katalog aus 7200 kaum erforschten Segmenten des menschlichen Genoms zusammen und präsentiert einen Fahrplan für ihre Integration in menschliche Genomdatenbanken. Ein Großteil der Gensegmente kodiert möglicherweise für Proteine und könnte erklären, was den Menschen von anderen Tieren unterscheidet. Die neue Studie in „Nature Biotechnology“ ebnet den Weg für neue Forschungsansätze im Bereich der menschlichen Gesundheit und Krankheit.**

Als Forschende im Jahr 2001 mit dem Humangenomprojekt das Erbgut des Menschen vollständig entschlüsselt hatten, gab es eine große Überraschung: Sie hatten nur 20.000 Gene gefunden, die Proteine produzieren. Sollte der Mensch also nur etwa doppelt so viele Gene besitzen wie eine Fliege? Wissenschaftler\*innen hatten mit erheblich mehr gerechnet.

Nun haben Forschende aus 20 Institutionen weltweit mehr als 7200 weitgehend unerforschte Genabschnitte zusammengefasst, die möglicherweise für neue Proteine kodieren. Sie haben dabei eine neue Technologie eingesetzt, die die proteinproduzierende Maschinerie in den Zellen im Detail untersucht, um mögliche Proteine beim Menschen zu finden. Das Humangenomprojekt mit all seinen Bemühungen, die menschlichen Gene zu beschreiben, war erst der Anfang, legt die neue Studie nahe. Das Forschungskonsortium möchte die wissenschaftliche Gemeinschaft dazu ermutigen, ihre Daten in die großen Humangenom-Datenbanken zu integrieren.

Vier gleichberechtigte Forscher haben die Studie, die nun in „Nature Biotechnology“ erschienen ist, geleitet: Dr. Jorge Ruiz-Orera vom Max-Delbrück-Centrum für Molekulare Medizin in der Helmholtz-Gemeinschaft (MDC) in Deutschland, Dr. Sebastiaan van Heesch vom Prinses Máxima Centrum in den Niederlanden, Dr. Jonathan Mudge vom Bioinformatik-Institut am European Molecular Biology Laboratory (EMBL-EBI) in Großbritannien und Dr. John Prensner vom Broad Institute des MIT und Harvard in den USA.

### **Die neuen Gensequenzen blieben unsichtbar**

In den letzten Jahren wurden Tausende, häufig sehr kleine „Open Reading Frames“ im menschlichen Genom entdeckt. Das sind Abschnitte im Erbgut, die Bauanleitungen für Proteine enthalten könnten. Mehrere Autor\*innen der aktuellen Studie haben in der Vergangenheit bereits ORFs gefunden: Van Heesch hat gemeinsam mit den Professoren Norbert Hübner und Uwe Ohler am MDC neue Mini-Proteine im menschlichen Herzen beschrieben und berichtete darüber 2019 in „Cell“; auch Prensner veröffentlichte 2021 zu ORFs in „Nature Biotechnology“. Keine dieser bislang nahezu unerforschten Segmente tauchten jedoch daraufhin in Referenzdatenbanken auf. Viele andere neu entdeckte Sequenzen, die Forschende weltweit z.B. in „Science“ oder „Nature Biochemical Biology“ beschrieben, blieben ebenfalls für den größten Teil der wissenschaftlichen Gemeinschaft unsichtbar – obwohl belegt ist, dass sie RNA-Moleküle produzieren, die daraufhin an die Proteinfabriken der Zelle, die Ribosome, binden.

Traditionell haben Wissenschaftler\*innen proteinkodierende Abschnitte in Genen identifiziert, indem sie DNA-Sequenzen von mehreren Spezies miteinander verglichen. Denn die wichtigsten kodierenden DNA-Sequenzen blieben im Laufe der Evolution von Tieren erhalten. Mit dieser Methode fielen jedoch kodierende Sequenzen durchs Raster, die relativ jung sind, die also erst

während der Entwicklung von Primaten entstanden sind. Sie fehlen in den Datenbanken.

Nun galt es also, die wenig beachteten ORFs in die größten Referenzdatenbanken zu integrieren, denn bislang musste man in der Literatur gezielt nach ihnen suchen, wenn man sie erforschen wollte. In einem ersten Schritt sammelte das internationale Forschungsteam Informationen zu Sequenzen, die mit dem „Ribosom-Profilings“ neu entdeckt wurden – diese Methode ermittelt, mit welchem Teil der Boten-RNA (mRNA) das Ribosom interagiert. Danach fügten sie die Daten zu einem standardisierten Satz zusammen. Keine leichte Aufgabe. Denn Daten aus unterschiedlichen Laboren, die auf verschiedenste Weisen gewonnen wurden, können nicht einfach so miteinander kombiniert werden.

Sobald dies geschafft war, beschäftigte sich das internationale Konsortium mit zentralen Fragen, die unsere Vorstellung vom menschlichen Genom prägen: Was ist ein Gen? Was ist ein Protein? Brauchen wir flexible Vorstellungen davon, ob Ribosomen immer ein Protein produzieren oder vielleicht auch ein ganz anderes zelluläres Signal?

Die Gruppe will nun die Humangenom-Datenbanken überarbeiten, die Forschende weltweit nutzen. Ensembl-Gencode richtet den ORF-Katalog als Bestandteil ihrer Referenz-Annotation-Datenbank ein, viele weitere wie UniProt, HGNC, PeptideAtlas and HUPO wollen folgen.

### **Vermutlich tragen die ORFs zu Volkskrankheiten bei**

„Unsere Forschung bringt das Verständnis des genetischen Aufbaus und der vollständigen Anzahl der Proteine im Menschen einen großen Schritt voran“, sagt van Heesch. „Es ist ungeheuer spannend, die Forschungsgemeinschaft mit unserem Katalog zu unterstützen. Wir können zwar jetzt noch nicht sagen, dass es sich bei allen neuen Sequenzen wirklich menschliche Proteine repräsentieren. Fest steht jedoch, dass ein großer Teil des menschlichen Genoms noch unerforscht ist und die Welt dies zur Kenntnis nehmen sollte.“

„Viel zu lange wurde die wissenschaftliche Gemeinschaft über diese ORFs im Unklaren gelassen“, sagt Mudge. „Wir sind sehr stolz darauf, dass Forschende auf der ganzen Welt sie nun untersuchen können. Ab jetzt sind sie für alle Wissenschaftler\*innen aus der Genomik und Medizin verfügbar – davon versprechen wir uns weitreichende Impulse.“

„Die meisten der 7200 ORFs aus unserem Katalog gibt es nur bei Primaten und stellen möglicherweise evolutionäre Neuerungen dar, die einzigartig sind für unsere Spezies“, sagt Orera, Evolutionsbiologin und Wissenschaftlerin in Hübners Arbeitsgruppe. „Vielleicht verraten sie mehr darüber, was den Menschen wirklich ausmacht.“

Was kommt als nächstes? Prensner meint, dass „diese ORFs mit ziemlicher Sicherheit zu vielen menschlichen Merkmalen und Krankheiten beitragen, auch zu häufigen Krankheiten wie Krebs.“ Die Herausforderung bestehe nun darin, herauszufinden, welche ORFs bei welchen Krankheiten welche Rolle spielen.

\*\*\*Gemeinsame Pressemitteilung des MDC, Prinses Máxima Centrum & EMBL-EBI

### **Über das Max-Delbrück-Centrum für Molekulare Medizin (MDC)**

Das Max-Delbrück-Centrum für Molekulare Medizin in der Helmholtz-Gemeinschaft gehört zu den international führenden biomedizinischen Forschungszentren. Nobelpreisträger Max Delbrück, geboren in Berlin, war ein Begründer der Molekularbiologie. An den MDC-Standorten in Berlin-Buch

und Mitte analysieren Forscher\*innen aus rund 60 Ländern das System Mensch – die Grundlagen des Lebens von seinen kleinsten Bausteinen bis zu organübergreifenden Mechanismen. Wenn man versteht, was das dynamische Gleichgewicht in der Zelle, einem Organ oder im ganzen Körper steuert oder stört, kann man Krankheiten vorbeugen, sie früh diagnostizieren und mit passgenauen Therapien stoppen. Die Erkenntnisse der Grundlagenforschung sollen rasch Patient\*innen zugutekommen. Das MDC fördert daher Ausgründungen und kooperiert in Netzwerken. Besonders eng sind die Partnerschaften mit der Charité – Universitätsmedizin Berlin im gemeinsamen Experimental and Clinical Research Center (ECRC) und dem Berlin Institute of Health (BIH) in der Charité sowie dem Deutschen Zentrum für Herz-Kreislauf-Forschung (DZHK). Am MDC arbeiten 1600 Menschen. Finanziert wird das 1992 gegründete MDC zu 90 Prozent vom Bund und zu 10 Prozent vom Land Berlin.

## **Über das Prinzessin-Máxima-Zentrum für pädiatrische Onkologie**

Wenn ein Kind schwer an Krebs erkrankt ist, zählt nur eines: eine Heilung. Deshalb arbeiten wir im Prinzessin-Máxima-Zentrum für pädiatrische Onkologie mit Leidenschaft zusammen, um die Überlebenschancen und die Lebensqualität von krebskranken Kindern zu verbessern. Jetzt und auf lange Sicht. Denn Kinder haben ihr ganzes Leben noch vor sich. Das Prinzessin-Máxima-Zentrum für pädiatrische Onkologie ist kein gewöhnliches Krankenhaus, sondern ein Forschungs Krankenhaus, das größte Zentrum für Kinderkrebs in Europa. Hier arbeiten mehr als 400 Wissenschaftler und 900 medizinische Fachkräfte eng mit niederländischen und internationalen Krankenhäusern zusammen, um neue Behandlungsmethoden und neue Perspektiven für eine Heilung zu finden. Auf diese Weise bieten wir den Kindern von heute die bestmögliche Behandlung und unternehmen wichtige Schritte zur Verbesserung der Überlebenschancen der Kinder, die noch nicht geheilt sind.

## **Über EMBL-EBI**

Das Europäische Bioinformatik-Institut (EMBL-EBI) ist weltweit führend bei der Speicherung, Analyse und Verbreitung großer biologischer Datensätze. Wir helfen Wissenschaftlern, das Potenzial großer Datenmengen auszuschöpfen, indem wir ihre Fähigkeit verbessern, komplexe Informationen zu nutzen, um Entdeckungen zum Nutzen der Menschheit zu machen. Wir sind führend in der computergestützten biologischen Forschung und arbeiten an Methoden der Sequenzanalyse, der multidimensionalen statistischen Analyse und der datengesteuerten biologischen Entdeckung, von der Pflanzenbiologie bis zur Entwicklung von Säugetieren und Krankheiten. Wir sind Teil des EMBL und befinden uns auf dem Wellcome Genome Campus, einer der weltweit größten Konzentrationen von wissenschaftlichem und technischem Know-how im Bereich der Genomik.

## **Originalpublikation:**

Jorge Ruiz-Orera et al (2022): “A community-driven roadmap to advance research on translated open reading frames”. Nature Biotechnology, DOI: [10.1038/s41587-022-01369-0](https://doi.org/10.1038/s41587-022-01369-0)

## **Weitere Informationen:**

<https://www.mdc-berlin.de/de/news/press/unbekannte-mini-eiweisse-im-herzen>

<https://www.nature.com/articles/s41587-020-00806-2>

[https://www.cell.com/cell/fulltext/S0092-8674\(19\)30508-2](https://www.cell.com/cell/fulltext/S0092-8674(19)30508-2)

<https://pubmed.ncbi.nlm.nih.gov/32139545/>

<https://pubmed.ncbi.nlm.nih.gov/31819274/>

<https://www.mdc-berlin.de/huebner> AG Hübner

<https://research.prinsesmaximacentrum.nl/en/research-groups/van-heesch-group> AG Van Heesch