



Wissenschaftler am EMBL-EBI entwickeln Verfahren zur Speicherung auf DNA All unser Film- und TV-Material in einer Kaffeetasse

Hinxton, 23. Januar 2013 – Wissenschaftler am EMBL-European Bioinformatics Institute (EMBL-EBI) haben eine Methode entwickelt, Daten in der Form von DNA zu speichern – einem extrem langlebigen Material, das mehrere zehntausend Jahre überdauern kann. Diese neue Methode, die heute in der Fachzeitschrift *Nature* vorgestellt wird, ermöglicht die Speicherung von mindestens 100 Millionen Stunden hochauflösendem Videofilm in einer Menge DNA, die in eine Kaffeetasse passt.

Die Menge an digitaler Information weltweit entspricht momentan ungefähr drei Zettabyte (zum Vergleich: 1 Zettabyte entspricht einer 1 mit 21 Nullen) und der immerwährende Zuwachs an neuen digitalen Inhalten stellt die Archivare vor große Herausforderungen. Festplatten sind teuer und benötigen eine ununterbrochene Stromversorgung, während sogar die besten stromunabhängigen Archivierungsmaterialien wie zum Beispiel Magnetbänder innerhalb von zehn Jahren zerfallen. Dies ist vor allem für die Lebenswissenschaften zunehmend ein Problem, denn die Speicherung riesiger Datenvolumen – einschließlich DNA-Sequenzen – macht einen wesentlichen Teil des wissenschaftlichen Datenbestandes aus.

“Wir wissen bereits, dass DNA ein äußerst robustes Speichermedium ist, denn wir können sie aus den Knochen des Wollmammuts extrahieren, die mehrere zehntausend Jahre alt sind und sie enthält immer noch sinnvolle Informationen,” erläutert Nick Goldman vom EMBL-EBI. “Außerdem ist sie unglaublich winzig und besitzt eine hohe Dichte. Und sie braucht zur Speicherung keine Energie, was Transport und Lagerung natürlich stark vereinfacht.”

Das Lesen von DNA ist relativ unkompliziert, bisher war es das Schreiben, das dieser Speicherethode im Weg stand. Denn dabei gibt es zwei Herausforderungen: zunächst einmal erlauben die derzeitigen Methoden lediglich eine Herstellung von kurzen DNA-Strängen und zweitens sind sowohl das Lesen als auch das Schreiben fehleranfällig, besonders dann, wenn sich ein Buchstabe in der DNA wiederholt. Nick und sein Mitautor Ewan Birney, stellvertretender Direktor am EMBL-EBI, machten sich also daran, einen Code zu entwickeln, der beide Probleme angeht.

“Wir waren uns bewusst, dass wir für unseren Code lediglich kurze DNA-Stränge verwenden durften. Außerdem musste er so beschaffen sein, dass eine Wiederholung desselben Buchstabens praktisch ausgeschlossen war. Deshalb kamen wir auf die Idee, den Code in eine Reihe sich überlappender Fragmente aufzubrechen, die sich in beide Richtungen erstrecken. Positionsinformation zeigen an, wo das jeweilige Fragment im Code seinen Platz hat. So entsteht eine Kodierung,



Nick Goldman vom EMBL-EBI mit der synthetisierten DNA.
Credit: EMBL Photolab

die keine Wiederholungen zulässt. Das heißt, derselbe Fehler müsste in vier verschiedenen Fragmenten auftreten, um wirklich Schaden anzurichten – und dies ist extrem selten,” sagt Ewan Birney.

Die neue Methode erfordert eine Synthetisierung von DNA aus der verschlüsselten Information. Dies rief das kalifornische Unternehmen Agilent Technologies Inc. auf den Plan, das seine Unterstützung anbot. Daraufhin schickten Ewan Birney und Nick Goldman verschlüsselte Versionen von folgenden Dateien nach Kalifornien: eine .mp3-Version von Martin Luther Kings Rede “I have a dream”, eine .jpg-Aufnahme des EMBL-EBI, eine .pdf-Datei mit Watson und Cricks bahnbrechender Veröffentlichung “Molecular structure of nucleic acids” sowie eine .txt-Datei mit den Sonetten Shakespeares. Ergänzt wurde dies von einer Datei, in der die Verschlüsselung beschrieben wird.

“Wir haben die Dateien aus dem Web heruntergeladen und damit mehrere hunderttausend DNA-Teilchen synthetisiert – das Ergebnis sieht aus wie eine winzige Menge Staub,” erklärt Emily Leproust von Agilent. Agilent schickte die Probe ans EMBL-EBI, wo es den Wissenschaftlern gelang, die DNA zu sequenzieren und die Dateien fehlerfrei zu dekodieren.

“Wir haben einen fehlertoleranten Code entwickelt, der sich einer molekularen Form bedient, von der wir wissen, dass sie unter günstigen Bedingungen 10 000 Jahre oder länger halten

Kontakt: contactpress@ebi.ac.uk

Mary Todd-Bergman, Senior Communications Officer, Hinxton, UK, Tel: +44 1223 494 665, www.ebi.ac.uk, contactpress@ebi.ac.uk
Isabelle Kling, EMBL Communication Officer, Heidelberg, Germany, Tel: +49 6221 387 8355, www.embl.org, isabelle.kling@embl.de

kann”, so Nick Goldman. “Solange es jemanden gibt, der den Code kennt, wird man ihn lesen können – falls man im Besitz einer Maschine ist, die DNA lesen kann.”

Obwohl es noch eine Reihe praktischer Probleme zu lösen gibt, gilt die DNA aufgrund ihrer Dichte und Langlebigkeit bereits jetzt als attraktives Speichermedium. Der nächste Schritt wird nun sein, die Kodierung zu perfektionieren und die praktischen Probleme zu lösen, um das DNA-Speichermodell wirtschaftlich rentabel zu machen. ●

Quelle

Toward practical high-capacity low-maintenance storage of digital information in synthesised DNA. Advanced online publication in *Nature* on 23 January 2013; DOI: 10.1038/nature11875

Contacts: contactpress@ebi.ac.uk

Mary Todd-Bergman, Senior Communications Officer, Hinxton, UK, Tel: +44 1223 494 665, www.ebi.ac.uk, contactpress@ebi.ac.uk

Isabelle Kling, EMBL Communication Officer, Heidelberg, Germany, Tel: +49 6221 387 8355, www.embl.org, isabelle.kling@embl.de

Über EMBL-EBI

Das Europäische Bioinformatik-Institut (EBI) ist Teil des Europäischen Laboratoriums für Molekularbiologie (EMBL) und befindet sich auf dem Wellcome Trust Genome Campus in Hinxton bei Cambridge (GB). Das EBI entstand aus der Pionierarbeit des EMBL heraus, biologische Datenbanken der Forschungsgemeinschaft öffentlich zugänglich zu machen. Heute ist es Sitz einer Reihe weltweit bedeutender biologischer Datenbanken, unter ihnen ENA (DNA-Sequenzen), UniProt (Protein-Sequenzen), Ensembl (Tiergenome), PDBe (Protein Datenbank in Europe), ArrayExpress (funktionelle Genomik), IntAct (Protein-Protein-Interaktionen) und Reactome (Biochemie-Onlinedatenbank für Molekül-Interaktionen auf zellulärer Ebene). Am EBI arbeitet eine Reihe von Forschungsgruppen kontinuierlich an der Entwicklung neuer bioinformatischer Tools. www.ebi.ac.uk/information.

Über EMBL

Das Europäische Laboratorium für Molekularbiologie (EMBL) ist ein Grundlagenforschungsinstitut, das sich über öffentliche Forschungsgelder aus 20 Mitgliedstaaten und Australien als assoziiertes Mitglied finanziert (Belgien, Dänemark, Deutschland, Finnland, Frankreich, Griechenland, Großbritannien, Irland, Island, Israel, Italien, Kroatien, Luxemburg, Niederlande, Norwegen, Österreich, Portugal, Schweden, Schweiz und Spanien). Etwa 85 unabhängige Forschungsgruppen arbeiten am EMBL zu Themen des gesamten Spektrums der Molekularbiologie. Das Institut ist in fünf Einheiten gegliedert: das Hauptlaboratorium in Heidelberg (900 Mitarbeiter) sowie Außenstellen in Hinxton (Europäisches Bioinformatik-Institut) (400 Mitarbeiter), Grenoble (70 Mitarbeiter), Hamburg (100 Mitarbeiter) und Monterotondo bei Rom (65 Mitarbeiter). Die Kernaufgaben des EMBL sind: molekularbiologische Grundlagenforschung; Ausbildung von Studenten, Wissenschaftlern und Gastwissenschaftlern; Serviceleistungen für Wissenschaftler in den Mitgliedstaaten; Entwicklung neuer Instrumente und Methoden für die Biowissenschaften sowie aktiver Technologietransfer. Im internationalen Doktorandenprogramm des EMBL forschen rund 190 Studenten. Darüber hinaus fördert das Institut den Austausch mit der Öffentlichkeit durch Vortragsreihen, Besucherprogramme und aktive Wissenschaftskommunikation.

Nutzungsbedingungen

EMBL Pressemitteilungen, Photos, Grafiken und Videos unterliegen dem EMBL copyright. Sie können für nicht-kommerzielle Nutzung frei reproduziert und verbreitet werden. Wir bitten um Nennung der Autoren und Institution.