

GenomA desCodeTzen

Etxebeste Aduriz, Egoitz

Elhuyar Zientziaren Komunikazioa



NHGRI

Zure begien kolorea, ahotsaren tonua, oinen tamaina, atzamar-kopurua, burmuinaren egitura, eta zure eta gainerako bizidunen informazio biologiko guzti-guztia idazteko, nahikoa dira lau letra. Bai, lau letra baino ez ditu genomaren alfabetoak, eta haien ordenaren baitan dago izaki bakoitzaren diseinu osoa. Baina lau letren sekuentziak luzeak dira, oso luzeak, eta haiek osorik irakurtzea erronka handia.

GENOMA KODE BAT DA; G, A, C eta T letren bidez adierazten diren lau nukleotidoren sekuentziaz osatutako kodea. Hori jakin zenetik, genoma nola deskodetu ikertzen aritu da jende asko. Eta, jakina, deskodetzeko, lehenengo irakurri egin behar dira letra horiek, edo, beste modu batera esanda, genoma sekuentziatu egin behar da. Hau da, G, A, C, eta T horiek zein ordenatan dauden jakin behar da.

Hori egiteko, azken 30 urteetan metodo bat erabili da batez ere, nahiz eta urte horietan guztietan hobekuntza nabariak izan dituen. Metodo hori Frederick Sanger-ek eta haren kideek garatu zuten 70eko hamarkadan —bigarren Nobel saria jaso zuen horregatik Sanger-ek—.

Sanger-en metodoaren oinarriko erreakziarako, lau osagai nagusi behar dira: sekuentziatu nahi den DNA-zatia edo 'DNA patroia', lau motatako nukleotido askeak, 'abiarazle' edo *primer* izenez ezagutzen diren 20-30 nukleotidoko DNA-harizpi txikiak eta DNA sintetizatzen duen DNA polimerasa entzima.

Erreakzioa hasteko, DNA-patroia berotu egin behar da, bi harizpiak bereiz daitezten. Hala, abiarazlea harizpietako bati lotuko zaio (sekuentzia osagarria dagoen lekuan). Izan ere, DNA polimerasak luzatu egiten ditu DNA-harizpiak, baina ezin da ezerezetik hasi; horregatik behar da abiarazlea. Behin abiarazlea DNA patroiarri lotu zaionean, DNA polimerasa abiarazlea luzatzen hasten da, nukleotidoz nukleotido: pa-



troian A bat dagoen tokian T bat jarriko du, eta alderantziz, eta G bat dagoen tokian C, eta alderantziz. Hori, gainera, zentzu jakin batean gertatzen da beti; hau da, DNA-harizpi baten bi muturrei 5' eta 3' izena ematen zaie, eta DNA polimerasak 3' muturrari erantzen dizkio nukleotidoak.

Horregatik, DNA patroia baten hainbat kopiatatik abiatuko bagina, patroia horren harizpietako baten beste hainbeste kopia lortuko genituzke. Baina, Sanger-en metodoaren gakoa nukleotido eraldatu batzuk erabiltzean dago: dideoxinukleotidoak. Nukleotido eraldatu horiei 3' muturrean hidroxilo-talde bat falta zaie, eta horren ondorioa da ezin zaiela beste nukleotidorik erantsi.

Lau nukleotido, lau erreakzio

Sanger-en metodoan, lau erreakzio egin behar dira DNA-patroi bakoitzarako. Lau erreakzioetan, goian aipatutako lau osagaiak daude, baina, bakoitzean, lau dideoxinukleotidotariko bakarra jartzen da. Hala, erreakzioetariko batek, esaterako, lau nukleotido normal eta dideoxi-G batzuk izango ditu. Erreakzio horretan, DNA polimerasa abiarazlea luzatzen hasten denean, G bat behar duen bakoitzean, bi gauza gerta daitezke zoriz: G normala edo dideoxi-G bat hartzea. Gehienak G normalak dira, baina, dideoxi-G bat hartzen duenean, polimerasak ezingo du katea luzatzen jarraitu. Hala, patroia hainbat kopiatatik abiatuta,

“Sanger-en metodoaren gakoa nukleotido eraldatu batzuk erabiltzean dago: dideoxinukleotidoak”

luzera desberdineko harizpiak lortuko ditugu, guztiak G batean bukatuak. Eta milioika kopiatatik abiatuta, azkenean, sekuentziako G bakoitzari dagozkion harizpiak izango ditugu. Eta gauza bera gertatzen da gainerako hiru erreakzioetan ere.

Gero, lortutako harizpi berriak tamainaren arabera bereizten dira, elektroforesi baten bidez. Elektroforesian, eremu elektriko batean jartzen dira DNA-harizpiak, eta, DNAk karga ne-

gatiboa duenez, polo negatibotik positibora mugitzen dira. Bidean, mugimendu hori oztopatuko duen zerbait jartzen da —esaterako, gel porotsu bat—, eta, hala, zati handiek txikiak baino denbora gehiago behar izaten dute polo batetik bestera mugitzeko. Modu horretan, nukleotido bakarrek aldearen harizpiak bereiz daitezke.

Elektroforesiaren emaitza ikusi ahal izateko, harizpi berriek nolabait markatuta egon behar dute. Hasieran, erradiaktiboki edo fluoreszentez markatutako nukleotidoak edo abiarazleak erabiltzen ziren horretarako. Hala, autoerradiografia baten bidez, edo izpi ultramoreen bidez, banda gisa ikus daitezke, tamainaren arabera bereizitako harizpiak.

Erreakzio bakoitzarekin lerro bat egingen da elektroforesian, eta azken emaitza da sekuentziaren nukleotido bakoitzeko banda bat ikusiko dugula lau lerroetakoren batean; eta, hala, banda horien ordenari jarraituz, sekuentzia hori irakurri ahal izango dugu.

Automatizazioa

Hori da, funtsean, orain arte sekuentziatutako DNA gehiena sekuentziatzeko erabili den teknika. Baina, Sanger-ek sortu zuenetik, hobekuntza handiak izan ditu. Garrantzitsuenetako bat Leroy Hood-ek 1986an asmatutako DNA automatikoki sekuentziatzeko makina izan zen. Hood-ek lau dideoxinukleotidoetariko bakoitza uhin-luzera (kolore) desberdineko fluoreszentsia



Autoerradiografiaren bidez, banda gisa ikus daitezke luzera desberdinetako DNA-harizpiak.



NHGRI

Sekuentziadore automatikoek aurrerapen handiak izan dituzte.

emateko moduan markatu zuen. Hala, batetik, lau erreakzio egin beharrean, bakarra egitea nahikoa da, eta, bestetik, izpi ultramoreekin argitutakoan igortzen duten uhin-luzera automatikoki detekta daiteke.

Sekuentziadore automatikoetan, lagin bakoitza lerro bakarrean doa, eta kolore desberdinak dituztelako detektatzen dira lau nukleotidoak. Sekuentziadore modernoetan, DNA-harizpiak beirantzuzko kapilar batzuetatik pasaz bereizten dira, muturretik ateratzean laser batek jotzen ditu, eta makinak fluoereszentzia detektatzen du. Lehenengo aterako diren harizpiek abiarazlea gehi nukleotido bat izango dute, eta jarraian aterako da sekuentzia osoa.

Azken hamarkadan, DNA-sekuentziadore automatikoak asko hobetu dira: askoz lagin gehiago prozesa ditzakete, azkarrago, eta erabiltzeko errazagoak dira. Egun, ehunka lagin sekuentziatu daitezke saio batean, eta egun batean 24 bat saio egin daitezke.

Genoma handia

Hala ere, oraindik ez da lan makala genoma oso bat deskodetzea. DNAREN sekuentziazio-teknikek DNA-zati txikiak sekuentziatzeko balio dute, 300-900 nukleotidoko sekuentziak. Izan ere, hori baino harizpi luzeagoetan, ezin dira zehaztasunez bereizi nukleotido bakarreko aldearekin.

“DNA-zatiak bakterioetan gordez, genoma osoa bilduko zuen liburutegia sortu zuten”

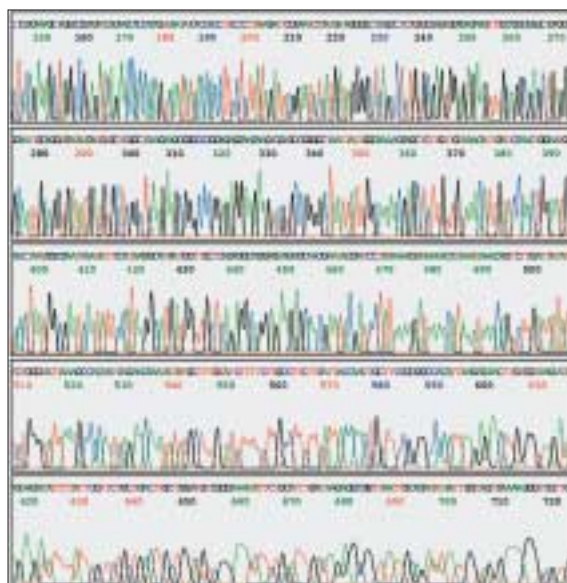
Baina bakterio baten genoma txikiak ere milioika nukleotido ditu; eta giza-kiarenek 3.000.000.000, 23 kromosomatan banatuta. Pentsa, aldizkari honetan letraz letra genoma osoa idazteko, 20.000 bat aldizkari beharko genituzke. Irakurritako sekuentzien batez besteko luzera (500 nukleotido) paragrafo hau litzateke gutxi gor-

behera; eta 20.000 aldizkaritan hone-lako sei milioi paragrafo sartzen dira. Gainera, fidagarritasun nahikoa lortzeko, behin ez, 6-10 aldiz irakurri beharko lirateke aldizkari horiek guztiak.

Horixe egin zuten, hain zuzen ere, Giza Genoma Proiektuan. Zatika-zatika giza genoma osoa irakurri zuten letraz letra. Baina, genoma osoa zatitu eta zati horiek guztiak irakurtzea gauza bat da, eta bestea, gero zati horiek zein ordenatan elkartu behar diren jakitea.

Proiektua hasteko, genomaren mapa bat eraiki zuten lehenengo. Kromosometan zehar galdu gabe ibili ahal izateko, milaka erreferentzia-puntu hartu zituzten. Behin mapa horiek eginda, genoma osoa bilduko zuen DNA-zatien 'liburutegi' bat sortu zuten. Horretarako, DNA-zati horiek bakterioen barnean gorde zituzten, eta erreferentzia-puntuei esker jakin zezaketen zati bakoitza genomaren zein tokitakoa zen. Azken finean, hori da edozein liburutegirik eskaintzen duena: informazio ordenatua. Horri esker, mundu osoko laborategietan modu koordinatuan lan egin ahal izan dute.

DNA-zatiak gure hesteetan bizi ohi diren *E. coli* bakterioetan gorde zituzten, bakterio-kromosoma artifizial (BAC) gisa. BAC bakoitzean, 100.000 edo 200.000 nukleotidoko DNA-zatiak gordezen dira.



Sekuentziadore automatikoek zuzenean irakurtzen dituzte sekuentziak, uhin-luzera desberdinak neurtuz.

NHGRI

E. coli bakterioak nahi adina denbora gorde daitezke izozkailuan. Hala, zientzialariak liburutegiko BAC bat behar duenean, bakterioak 37 °C-an berpiztu besterik ez du egin behar. Gainera, DNA-zatia barruan duen bakterioa ugaltzen denean, DNA-zati hori ere harekin batera ugaltzen da. Hala, bakterioei gau batez hazten uztea nahikoa da barruan duten DNA-zatiaren milioika kopia lortzeko. Horri DNA aplikatzea esaten zaio.

BACak oso handiak dira, oraindik, sekuentziatu ahal izateko. Horregatik, BACak ere zatitu egiten dira zoriz, eta gainjartzen diren zati txikiagoak lortzen dira. Zati horiek bakterioak infektatzen dituzten birusetan edo plasmidoetan sartzen dira, gero, berriz ere, *E. coli* bakterioetan aplikatzeko. Azkenik, bakterioetatik DNA purifikatu eta sekuentziatu egiten da. Gero, zati desberdinetan gainjartzen diren sekuentziak identifikatu eta BACaren sekuentzia osoa osatzen da. Aldi berean, BACak ere gainjarri egiten dira bata bestearekin. Eta, hala, irakurketaz irakurketa, BACez BAC eta kromosomaz kromosoma, genoma osoa irakurri arte.

Lasterketa genomikoa

Genoma oso bat horrela irakurtzeko, denbora, lan eta diru asko behar da. Horregatik, ikertzaileak etengabe ari dira sekuentziazio-teknika azkarragoak eta merkeagoak garatzeko ikertzen.

Sekuentziatuak

Giza genoma ez da sekuentziatu den genoma bakarra, jakina. Lehen DNA-genoma osoa Frederick Sanger-ek sekuentziatu zuen 1977an; Φ-X 174 fagoarena. 1995ean, Venter-ek, Fraser-ek eta Smithek *Haemophilus influenzae* bakterioaren sekuentzia osoa argitaratu zuten, eta, geroztik, 1.800dik gora birusen, ia 600 bakterioaren eta 20 baino izaki eukarioto baino gehiagoren —erdiak baino gehiago onddoak—genomak sekuentziatu dira.

Sekuentziatu diren bakterio gehienak gaitoxotasunak eragiten dituztelako aukeratu dira; eta, gainerako izakiak, ongi ikertutako organismo-ereduak direlako. Legamia, esaterako, (*S. cerevisiae*) zelula eukariotoen eredu garrantzitsua da, eta *Drosophila melanogaster* ozpin-eulia oso garrantzitsua izan da genetikari. Gauza bera gertatzen da *Arabidopsis thaliana* landarearekin, zebra-arrainarekin, saguarekin, arratoiarekin, txinpantzearekin eta beste hainbatekin ere.



A. thaliana-ren 20 barietate.

Metodo berri askok sekuentziazio-ahalmena handitzea bilatzen dute, eta milaka edo milioika sekuentziazioaldi berean egitea. Baina ideia berriak ere ez dira falta; esaterako, DNA-polime-

rasak nukleotidoa harizpiari eranstendion unean bertan detektatzen duen metodo bat, edo harizpiak nanoporetatik pasatzean nukleotido bakoitza identifikatuko luketen korrante elektrikoak detektatzen dituen.

“mila gizakiren genomak sekuentziatzea du helburu nazioarteko proiektu berri batek”

Proiektuak asko dira, eta zorabiatzeko moduko abiaduran ari dira garatzen. Ez da harrizkoa, erakunde publikoak eta pribatuak inbertsio handiak egiten ari dira. 2006an, AEBko National Human Genome Research Institute-k (NHGRI) 13 milioi dolar jarri zituen hainbat proiektu finantzatzeko, DNA sekuentziazioa merkeagoa egingo duten teknologien garapena azkartzeko asmoz. Urte berean, X Prize fundazioak hamar milioi dolarreko saria iragarri zuen, 2006an ehun giza genoma hamar egunetan sekuentziatzeko gai den gailua egiten duen lehen taldearentzat. Eta, ehun ez, mila gizakiren genomak sekuentziatzea du helburu berriki abiarazi duten nazioarteko proiektu batek.

Esan izan da genomaren sekuentziazioa gizakiaren historian egin den aurrerapausorik garrantzitsuenetako bat dela. Mundu osoko ikertzaileak buru-belarri ari dira genomaren sekretu guztiak argitu nahian, kodea deskodetu nahian. Argi dago, **GenomiC**aren **AroAn GA**ude.



DNA-zatiak *E. coli* bakterioetan gordetzen dira.