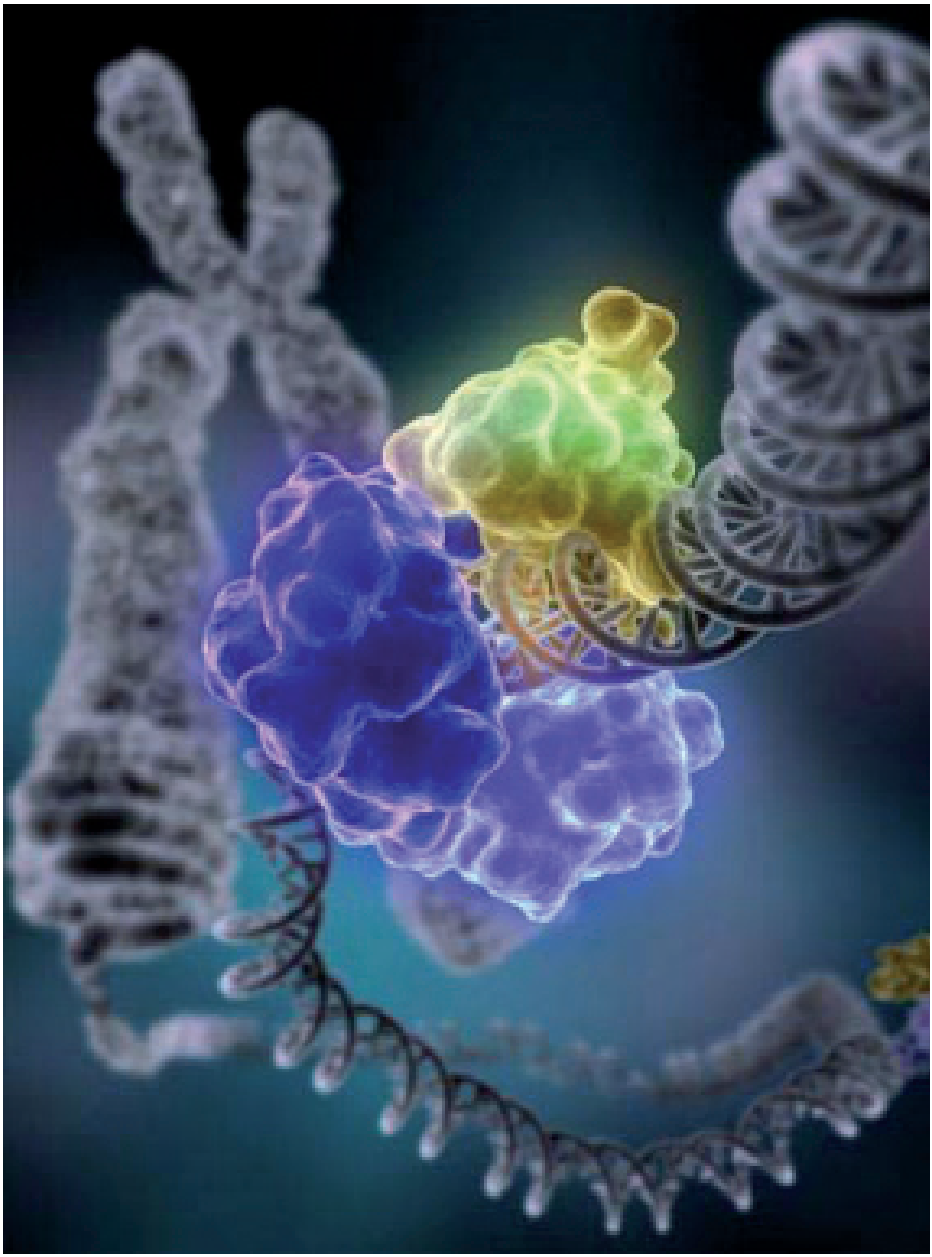
Alumna: Anna Galindo Verdugo
Tutor: Òscar Puigarnau González

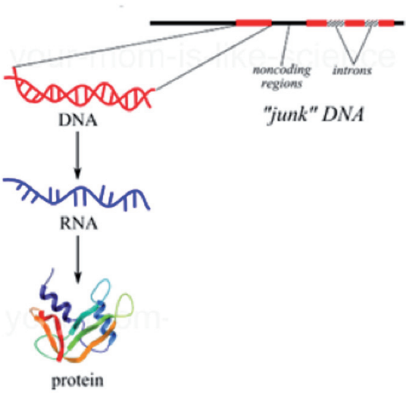
El Projecte Encode, més que gens

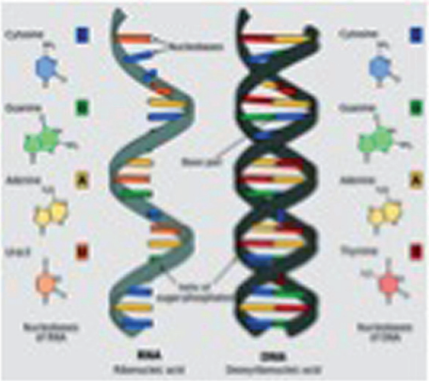
Presentació
En aquest treball, «El Projecte Encode, més que gens», es pretén donar a conèixer les investigacions que s'han dut a terme al llarg dels anys en l'estudi del genoma humà i fer una anàlisi específica de la seva importància en aplicacions pràctiques.

He realitzat el meu treball de recerca sobre el genoma humà i el projecte Encode perquè tots aquests estudis han capgirat el panorama del coneixement de la complexitat del genoma i han obert llum a noves investigacions pel que fa al diagnòstic i a la teràpia de moltes malalties genètiques.

Metodologia
Aquest treball ha estat fonamentat i documentat amb els estudis oficials que es van fer en el desenvolupament tant del Projecte Genoma Humà com del Projecte Encode. A la vegada, ha estat necessària la consulta de llibres i altres fonts pel que fa la comprensió teòrica de tot el cos del treball.







Cos del treball

Introducció teòrica

Els àcids nucleics són macromolècules complexes que contenen la informació genètica. Es distingeixen dos tipus de pentoses que poden formar els àcids nucleics: la desoxiribosa que donarà lloc a DNA i la ribosa que formarà una cadena de RNA. Alhora, les bases nitrogenades possibles en els dos àcids nucleics són diferents; adenina, citosina i guanina estan presents tant en el DNA com en el RNA, mentre que la timina només apareix en el DNA i l'uracil és exclusiu del RNA.

L'àcid desoxiribonucleic forma el material genètic de tots els organismes cel·lulars i quasi bé de tots els virus. El DNA conté la informació necessària per sintetitzar proteïnes a partir d'un procés de duplicació, transcripció i traducció d'aquest.

L'àcid ribonucleic o RNA està constituït per nucleòtids de ribosa i és gairebé sempre monocatenari, excepte en alguns virus. A la cèl·lula trobem diferents tipus de RNA depenent de la seva funció.

Una de les fonts de variabilitat genètica que han fet possible l'evolució de les espècies han estat les mutacions. Es tracta de canvis heretables en la seqüència de nucleòtids del DNA d'un organisme. Les mutacions ocorren per causa atzarosa tot i que són poc freqüents. L'existència de mutacions provoca la diversitat genètica i les

lleis de l'herència estableixen com aquestes mutacions i la informació genètica es transmeten a la descendència. L'herència es caracteritza per una enorme regularitat que segueix uns patrons predicibles. Les regles bàsiques de l'herència les devem als treballs realitzats per Gregor Mendel a partir de plantes pesoleres abans de saber que els caràcters hereditaris es transmetien d'una generació a la següent.

El Projecte Genoma Humà

El Projecte Genoma Humà (PGH) és un programa d'investigació a escala mundial per determinar la localització precisa dels gens que formen el genoma humà, així com determinar quina funció realitzen cadascun d'aquests gens.

El PGH representa el primer gran esforç de coordinació internacional en la història de la recerca biològica. Es proposa determinar la seqüència completa dels conjunts de bases que formen el genoma localitzant amb exactitud tots els gens i la resta de material hereditari de la nostra espècie. Amb aquesta recerca s'espera un impuls en el coneixement dels processos biològics, des de l'escala molecular fins a l'evolutiva, i de la fisiologia i patologia dels éssers humans, que es tradueix en múltiples aplicacions, tant tècniques com comercials, en matèries com el diagnòstic i la teràpia de malalties.

D'aquesta manera aquest projecte es marca dos objectius concrets. D'una banda es va plantejar crear un mapa genètic que permetés als investigadors situar cada gen. D'altra banda es va marcar com a objectiu dur a terme la seqüenciació del DNA, ja que la cadena exacta de nucleòtids del DNA facilita la investigació a l'hora de saber com el gen pot donar lloc a una malaltia.

Com a resultat del Projecte Genoma Humà es va obtenir la seqüència completa dels 3.200 milions de nucleòtids o bases nitrogenades que el componen, es va aconseguir dissenyar el mapa que ubica els prop de 30.000 que hi tenen lloc i l'anàlisi de prop de 1.400 gens causants de malalties.

A la vegada, es va demostrar que els éssers humans compartim un 99,9% d'aquesta seqüència i que el 0,1% restant varia entre cada individu, essent les variacions més comunes aquelles en les quals només es canvia una base. Així doncs, aquesta variació genòmica provoca que cada membre de l'espècie tingui característiques úniques.

El genoma humà

El genoma és el conjunt de gens continguts en els cromosomes, és a dir, la totalitat de la informació genètica que posseeixen els organismes.

El genoma humà descriu la totalitat d'informació genètica continguda en les cèl·lules humanes. Realment abraça dos genomes, el genoma nuclear, que correspon al 99,9995% del total de la informació genètica, i la part restant, corresponent al 0,0005%, que representa el genoma mitocondrial.

El DNA mitocondrial està format per una doble cadena circular de 16,6 kilobases de llargada. Està format per 37 gens i la seva herència només depèn del genoma mitocondrial provinent de l'òvul femení.
El genoma humà nuclear està format per aproximadament 3.200 megabases, és a dir, tres mil dos-cents milions de parells de bases. El DNA nuclear forma una cadena lineal i es troba en forma de cromosomes. Una cèl·lula diploide d'un ésser humà conté 23 parells de cromosomes, 22 dels quals són autosomes i un parell són cromosomes sexuals. De tota la informació continguda en els cromosomes una porció molt petita correspon als gens. Es calcula que per cada 109 bases del DNA nuclear, 3 d'elles formen gens i per tant codifiquen proteïnes. La majoria dels gens determinen quins aminoàcids formaran la proteïna i en quin ordre se situaran. Aquest DNA es coneix amb el nom de DNA codificant. Tot i això, la major part del DNA humà està format per seqüències que no codifiquen proteïna però que realitzen altres funcions.
A partir del Projecte Genoma Humà es va poder realitzar una primera classificació del genoma, que distingeix d'una banda els gens i les seqüències relacionades amb els gens i d'altra banda tot el DNA que es troba entre els gens que s'anomena DNA extragènic que no codifica cap proteïna. La major part del genoma humà, concretament un 70%, està format per aquest tipus de DNA, mentre que un 30% del genoma inclou seqüències lligades amb els gens. D'aquest 30%, només un 5% està format per DNA codificant, mentre que la resta és DNA no codificant associat a gens. En conseqüència, només un 1,5% del total del genoma humà representa DNA codificant i per tant sintetitza polipèptids.

El Projecte Encode
Amb el Projecte Genoma Humà es va poder saber que de tot el genoma, únicament una petita part correspon als gens. A la vegada, d'aquests gens, només una part minoritària conté la informació necessària que donarà lloc a proteïnes. De tots els gens que hi ha, no només hi ha els que tenen una proteïna com a producte, sinó que també n'hi ha molts que tenen com a producte final molècules de RNA. Tot i això, una petita part del genoma humà s'expressa en productes gènics. Així doncs, què fa la resta del genoma? Amb anterioritat, com que no se sabia la funció d'aquest DNA se l'anomenava DNA brossa. En canvi, ara se sap que aquest DNA compleix diverses funcions reguladores. És aquí on entra en joc el Projecte Encode que treballa per generar una Enciclopèdia d'Elements de DNA, és a dir, una llista de tots els elements funcionals del genoma humà, tant dels gens com d'altres seqüències. Així doncs, el Projecte Encode té com a objectiu final aconseguir explicar com la informació dels gens humans es regula i s'expressa en elements funcionals.

Encode es tracta d'un projecte d'anàlisi minuciosa del genoma humà. A partir d'aquest s'ha pogut obtenir una imatge detallada del genoma que mostra tots els transcrits primaris i madurs, així com la localització de les principals modificacions d'histones, els llocs d'unió dels factors de transcripció, llocs d'inici de la transcripció... A la vegada s'han obtingut dades d'expressió gènica, de replicació i del nombre de còpies d'aquestes regions.
Els resultats obtinguts indiquen que el 80% del genoma humà conté elements relacionats amb funcions bioquímiques, deixant per falses les pròpies definicions de DNA brossa o sense funció. Per comprendre el treball Encode s'ha de conèixer amb precisió què és una seqüència funcional. Encode la defineix com una regió de DNA que genera un producte definit o que realitza una funció bioquímica concreta, per tant engloba gens, RNA no codificants, regions promotores, elements de transcripció.
La regulació gènica determina quan, com i on s'expressa un gen concret. S'ha observat gràcies a aquest projecte que és un procés molt més complex del que s'havia plantejat amb anterioritat. Tant és així, que un cop acabada la segona fase del projecte es va anunciar que les dades que s'havien obtingut farien canviar diversos conceptes recollits fins ara en els llibres de text. S'ha pogut saber que el 62% de la seqüència del genoma humà dóna lloc a molècules de RNA de diferents tipus. D'una banda, aquelles que mitjançant un procés de traducció donen lloc a proteïnes, que representa un 2,9% del total de la informació continguda en el genoma. El segon lloc, casos en què a partir del DNA s'obté RNA com a producte final, ja que no donarà lloc a una proteïna sinó que funcionarà com un element regulador de l'expressió d'altres gens. En aquest tipus de molècules es distingeixen els llargs no codificants i els curts no codificants, en els quals trobem els RNA de transferència, RNA nuclear i RNA nucleolar. Aquests últims, els RNA no codificants, són molt importants per a la regulació genètica.
Com s'ha explicat anteriorment, s'han descobert formes de regulació que impliquen el processament del RNAm. En alguns casos un mateix gen produeix una proteïna en un teixit i en un altre de diferent se sintetitza un altre pèptid. El mecanisme pel qual es poden obtenir d'un mateix gen dues proteïnes s'anomena splicing o empalmament alternatiu. Aquest procés consisteix a unir diferents combinacions d'axons dels pre-RNA obtenint dos RNAm madurs amb diferent informació i, per tant, dos productes que són diferents pel que fa a la seqüència d'aminoàcids.
Sembla, doncs, que el gran mèrit d'aquest projecte ha estat mostrar-nos el que fins ara es desconeixia o no se sabia. Està tenint repercussions importants que van des de la redefinició del concepte de gen fins a l'obtenció de noves vies per a la descoberta de les causes de les malalties genètiques. Ha transformat també la visió que teníem sobre el nostre genoma, deixant enrere la idea de DNA brossa.

Conclusions

Amb la realització d'aquest treball he pogut arribar a la conclusió que el coneixement sobre el DNA i tot el relacionat amb el genoma ha estat molt important per a la recerca biològica i per tal de dotar de veracitat conceptes que amb anterioritat es tenien definits amb l'enunciat de DNA brossa.

Aquests projectes han capgirat el punt de vista que es tenia del genoma i totes les seves seqüències ja que s'ha dotat de funcionalitat regions que aparentment no realitzaven cap funció perquè no codificaven cap producte proteic.

Gràcies a la realització del meu treball de recerca he pogut concloure que és de vital importància conèixer la totalitat del genoma i de les seves seqüències funcionals per tal de desenvolupar tècniques que puguin crear sistemes de recerca de causes de malalties així com de processos bioquímics que es realitzen en el genoma.

Bibliografia

Llibres: — CARDONA, Ll. Genética. Editorial Océano, 2001. — PERETÓ, J.; SENDRA, R.; PAMBLANCO, M.; BAÑÓ, C. Fonaments de bioquímica. Universitat de València, 2005.
Webs: <http://www.genome.gov/> — <http://www.ugr.es/> — <https://www.encodeproject.org/> — <http://www.nature.com/encode/>.