

Գ.Տ. ԿԻՐԱԿՈՍՅԱՆ, Ա.Վ. ՂՈՒԼՅԱՆ, Ե.Ա. ՄԱՆԱՍՅԱՆ, Լ.Կ. ԱՆԴՐԵԱՍՅԱՆ
ԿԱՆԽԱՐԳԵԼԻՉ ԱՆՀԱՏԱԿԱՆԱՑՎԱԾ ՀԻԵՐԱՐԻՏԻԿԱԿԱՆ
ԱՌՈՂՋԱՊԱՀԱԿԱՆ ՀԱՄԱԿԱՐԳԵՐԻ ՄՇԱԿՄԱՆ ՄԵԹՈՂԱԲԱՆՈՒԹՅՈՒՆԸ
ԵՎ ՄՈՏԵՑՈՒՄՆԵՐԸ

Նորարարական կանխարգելիչ անհատականացված բժշկությունը նոր ամբողջական հայեցակարգ է առողջապահության ոլորտում, որը թույլ է տալիս կանխատեսել անհատի հիվանդության հակումները՝ մինչև հիվանդության սկիզբը, ապահովել նպատակային կանխարգելիչ միջոցներ, ձեռնարկել և ստեղծել անհատականացված բուժման ավգորիթմներ, որոնք հարմարեցված են անհատին:

Առանցքային բառեր. կանխարգելիչ անհատականացված բժշկություն, գիտելիք, մեքենայական ուսուցում, մեծածավալ տվյալներ, ամպային ծառայություններ:

Г.Т. КИРАКОСЯН, А.В. ГУЛЯН, Е.А. МАНАСЯН, Л.К. АНДРЕАСЯН

МЕТОДОЛОГИЯ И ПОДХОД К РАЗРАБОТКЕ
ПРОФИЛАКТИЧЕСКИХ ПЕРСОНАЛИЗИРОВАННЫХ
ИЕРАРХИЧЕСКИХ МЕДИЦИНСКИХ СИСТЕМ

Инновационная прогностическая, профилактическая и персонализированная медицина (ПППМ) является новой интегративной концепцией в секторе здравоохранения, что позволяет прогнозировать индивидуальную предрасположенность до появления заболевания, обеспечить целенаправленные предупредительные меры и создать алгоритмы персонализированного лечения с учетом личности.

Ключевые слова: индивидуальная медицина, знание, машинное обучение, большие данные, облачные услуги.

ՀՏԴ 004.3:606:61

Ա.Տ. ԵՍԱՅԱՆ

ՄԱՐԴՈՒ ՄԻՏՈՔՈՆՂՐԻՈՒՄԱՅԻՆ ԴՆԹ-Ի ՏՎՅԱԼՆԵՐԻ ՊԱՀՊԱՆՄԱՆ ԵՎ
ՎԵՐԼՈՒԾՄԱՆ ԽՆԴԻՐՆԵՐԻ ՀԵՏԱԶՈՏՈՒՄ ԵՎ ՆՈՐ ՀԱՄԱԿԱՐԳԻ
ԱՌԱՋԱՐԿՈՒՄ

Դիտարկվում են մարդու միտոքոնդրիումային ԴՆԹ-ի տվյալների պահպանման և վերլուծման խնդիրները, որոնք առաջանում են պոպուլյացիոն գենետիկայում, բժշկական գենետիկայում և էվոլյուցիոն տեսության ոլորտներում կատարվող հետազոտությունների ընթացքում:

Առանցքային բառեր. միտոքոնդրիումային ԴՆԹ, գենոմ, պոպուլյացիոն գենետիկա, բժշկական գենետիկա, էվոլյուցիոն տեսություն:

ԴՆԹ-ի սեկվենավորման մեթոդների բուռն զարգացման դարաշրջանում մայրական տոհմագծով փոխանցվող միտոքոնդրիումային ԴՆԹ-ն (մտԴՆԹ) դարձել է անփոխարինելի գործիք պոպուլյացիոն գենետիկական, բժշկական և էվոլյու-

ցիոն տեսության հետազոտությունների համար [1,2]: ՄտԴՆԹ-ն ներկայացնում է ամբողջ գենոմի մի մասը, 16569 զույգ հիմքերից բաղկացած շրջանաձև երկպարույր ԴՆԹ՝ բաղկացած 37 գեներից, որոնք հիմնականում կատարում են էներգետիկ գործառույթ [3]: Մշտապես կուտակվող մեծածավալ մտԴՆԹ-ի տվյալները արդյունավետ վերլուծության ենթարկելը այսօր ներկայացնում է դժվարին խնդիր: Գլխավոր խոչընդոտը տվյալների պահպանման և վերլուծման միասնական համակարգի բացակայությունն է: Չնայած հայտնի են մի շարք նման տվյալների բազաներ, և գոյություն ունեն բավականին լավ ծրագրային միջոցներ տվյալների վերլուծման համար, սակայն դրանք միմյանց հետ կապակցված չեն, պետք է որոնել և գտնել անհրաժեշտ տվյալները, ներբեռնել դրանք, այնուհետև այն որպես մուտքային ֆայլ փոխանցել վերլուծող ծրագրային միջոցին: Շատ դեպքերում բազայում եղած տվյալների ձևաչափը չի համապատասխանում վերլուծող ծրագրային միջոցների պահանջներին, և ստիպված մեծածավալ աշխատանք ու հաշվարկներ են կատարվում ձեռքով, ինչը, անշուշտ, ժամանակատար է և հանգեցնում է սխալների: Խնդիրը ավելի է բարդանում միաժամանակ մեծ քանակությամբ մարդկանց տվյալների հետ աշխատելիս: Տվյալների վերլուծության ընթացքում ստացվում են ԴՆԹ-ի կարևոր բնութագրիչներ, որոնք տվյալների բազայում պահպանելու դեպքում հետագա հետազոտություններում կարելի է խուսափել միևնույն բնութագրիչը կրկնակի անգամ հաշվարկելուց, սակայն ոչ բոլոր տվյալների բազաներն են, որ ունեն լրացուցիչ բնութագրեր պահպանելու հնարավորություն: Օրինակ, NCBI Genebank-ում (<http://www.ncbi.nlm.nih.gov>), որը համարվում է գենետիկական տվյալների ամենամեծ բազան աշխարհում, պահվում են բոլոր սեկվենավորված մարդկանց տվյալները, սակայն դրանք միայն ԴՆԹ-ի տվյալներ են և որևէ լրացուցիչ բնութագրիչ տվյալ չեն ներառում: Որոնման համակարգի սակավ ֆունկցիոնալ հնարավորությունների պատճառով այստեղից մեծաքանակ տվյալների ներբեռնումը բավականին դժվար է:

ՄտԴՆԹ-ի մեկ այլ հայտնի տվյալների բազա է HmtDB-ն (<http://www.hmtdb.uniba.it/hmtdb>): Սա ֆունկցիոնալ հնարավորություններով բավականին հարուստ տվյալների բազա է, ներառում է ավելի քան 32000 մարդու ամբողջական մտԴՆԹ-ի տվյալներ և հապլոխմբեր: Միտաքոնդրիոմային ԴՆԹ-ի մուտացիաների որոշակի խումբը սահմանվում է որպես մտԴՆԹ-ի հապլոխումբ [4]: Սա կարևոր բնութագրիչ է և որոշվում է առանձին ծրագրային փաթեթների միջոցով (Haplofind, HvrBase++): HmtDB-ում ստեղծված են գործիքներ տվյալների բազմաչափ որոնման, տեսակավորման և ներբեռնման համար: Դիտակվել է նաև mtDB տվյալների բազան (<http://www.mtDB.igp.uu.se>): Տվյալները խմբավորված են ըստ աշխարհագրական դիրքերի և էթնիկ խմբերի, սակայն այստեղ մեծաքանակ մարդկանց տվյալները միաժամանակ ներբեռնելու հնարավորություն չկա: Վերոնշյալ բոլոր տվյալների բազաներից տվյալները կարելի է ներբեռնել FASTA ձևաչափով: Դա գենետիկական տվյալներ պահելու հատուկ մշակված

ձևաչափ է, և շատ ծրագրային փաթեթներ տվյալների վերլուծության համար օգտագործում են հենց FASTA ձևաչափով մուտքային տվյալները [5]: Այնուամենայնիվ, շատ են նաև այն ծրագրային փաթեթները, որոնք որպես մուտք ընդունում են հատուկ ձևաչափով տվյալներ: Դրանցից է, օրինակ, Arlequin ծրագրային փաթեթը, որը լայնորեն կիրառվում է տվյալների վերլուծման ընթացքում, ընդ որում՝ շատ դեպքերում մուտքային տվյալներում պետք է ներառված լինեն այնպիսի բնութագրիչներ, ինչպիսիք են հապլոխումբը, մուտացիայի ենթարկված գեների անվանումները, աշխարհագրական դիրքը, ազգությունը և այլն: Նման ձևաչափով մուտքային տվյալներ ձևավորելը շատ աշխատատար է, քանի որ չկա ծրագրային որևէ միջոց՝ այս ձևաչափով տվյալները ձևավորելու համար:

Հետազոտելով մտԴՆ-ի տվյալների բազաները և տվյալների վերլուծման գործիքները՝ կարելի է նկատել, որ առկա տվյալների բազաների ֆունկցիոնալ հնարավորությունները չեն բավարարում պոպուլյացիոն գենետիկայում, բժշկական գենետիկայում և էվոլյուցիոն տեսության ոլորտներում կատարվող հետազոտությունների արդյունավետ իրականացմանը: Դիտարկված տվյալների բազաներում հնարավոր չէ տվյալները ներբեռնել որպես պատրաստի մուտքային ֆայլ, այլ անհրաժեշտ է ներբեռնելուց հետո ձևափոխել այդ տվյալները՝ դրանք հետագա օգտագործման համար պիտանի դարձնելու համար: Այս գործընթացը շատ ժամանակատար է, և ձևափոխման ընթացքում մեծ է սխալ թույլ տալու հավանականությունը: Առաջարկվում է ստեղծել մի նոր համակարգ, որը կներառի բոլոր առկա տվյալների բազաներում եղած մարդկանց տվյալները, ապիովել տարբեր ձևաչափերով տվյալներ ներբեռնելու հնարավորություն, ստեղծել տվյալների վերլուծության արդյունքում ստացված բնութագրիչները համակարգում պահպանելու հնարավորություն, ինչպես նաև ներդնել ծրագրային միջոցներ՝ տվյալները հենց տեղում վերլուծության ենթարկելու համար:

Այսպիսով, աշխատանքի շրջանակներում դիտարկվել են մտԴՆ-ի հայտնի տվյալների բազաները, քննարկվել են դրանցից օգտվելու անհարմարությունները և թերությունները: Առաջարկվել է ստեղծել մի նոր համակարգ, որը ոչ միայն հնարավորություն կտա տվյալները ներբեռնել հետագա օգտագործման համար պիտանի ձևաչափով, այլև համալրված կլինի տվյալները տեղում վերլուծելու կենսատեխնիկական գործիքներով:

ԳՐԱԿԱՆՈՒԹՅԱՆ ՑԱՆԿ

1. **Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C.** Initial sequencing and analysis of the human genome. - 2001.

2. **Bandelt H.J., Richards M., Macaulay V.** Human mitochondrial DNA and the evolution of Homo sapiens. - 2006. - Vol. 18.
3. **Anderson S., Bankier A.T., Barrell B.G., De Bruijn, M. H.** Sequence and organization of the human mitochondrial genome. – 1981.
4. **Handt O., Meyer S., A von Haeseler.** Compilation of human mtDNA control region sequences – 1998.
5. **Pinho A., Pratas D.** MFCompress: a compression tool for FASTA and multi-FASTA. – 2014.

А.Т. ЕСАЯН

**ИССЛЕДОВАНИЕ ПРОБЛЕМ ХРАНЕНИЯ И АНАЛИЗА ДАННЫХ
ЧЕЛОВЕЧЕСКОГО МИТОХОНДРИАЛЬНОГО ДНК И
ПРЕДЛОЖЕНИЕ НОВОЙ СИСТЕМЫ**

Рассматриваются проблемы, связанные с хранением и анализом данных человеческого митохондриального ДНК, которые встречаются в исследованиях популяционной генетики, медицинской генетики и в теории эволюции.

Ключевые слова: митохондриальный ДНК, популяционная генетика, медицинская генетика, теория эволюции, геном.

A.T. YESAYAN

**INVESTIGATING THE PROBLEMS ON THE STORAGE AND ANALYSIS
OF THE HUMAN MITOCHONDRIAL DNA DATA AND PROPOSING A
NEW SYSTEM**

The problems associated with the storage and analysis of the human mitochondrial DNA data are considered, used in the studies devoted to population genetics, medical genetics and the theory of evolution.

Keywords: mitochondrial DNA, population genetics, genom, medical genetics, evolution theory.