Hkkjrh; df"k vuqla/kku if=dk] 36(1): 25-28 DOI: 10.18805/BKAP265

vxyh ih<h d vuØe.k MVk fo'y"k.k ,o —f"k e bld vuç;kx

'kcuk cxe[1] jkgy cuth

Hkk-—v-i& Hkkjrh; —f"k lkf[;dh vuqla/kku lLFkku] ykbcjh ,oU;] ubZ fnYyh] HkkjrA

ikIr% Qjojh 2021 Lohdr% tuwu 2021

lkjk'k

vxyh ih<h dh vuØe.k (NGS) çk|kfxfd;k vi{kk—r de ykxr ij thukfed MVk d vf/kxg.k e Økfr yk jgh gA ,uth,l çk|kfxfd;k rth l tfVy thukfed v/;;u d –f"Vdk.k dk cny jgh g vkj MVk dh ,d fo'kky ek=k mRiUu dj jgh gA ,uth,l MVk fo'y"k.k }kjk çnku dh xb MVk dh bruh cMh ek=k d ,l cM tfod lxg dk lHkkyu d fy, u, vkj vf/kd 'kfä'kkyh to lpuk foKku d çLrkok dh vko';drk gkrh gA bld vykok] fo'k"k l,¶Vo;j midj.k vkj MVk ,dhdj.k d fy, mUur dEI;V'kuy lk/ku vko';d gA bl y[k e] ge ;gk vxyh ih<h d vuØe.k MVk fo'y"k.k vkj —f"k foKku e bldh Hkfedk d fy, e[; dEI;V'kuy –f"Vdk.k dk o.ku dj jg gA

'kCn dthZ% dEI;V'kuy lk/ku; thukfed v/;;u] vxyh ih<h d vuØe.k MVkA

Bhartiya Krishi Anushandhan Patrika, 36(1): 25-28

Next Generation Sequencing Data Analysis and its Applications in Agriculture

Shbana Begum[1], Rahul Banerjee

ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, Pusa-110 012, New Delhi, India.

Received: February 2021 Accepted: June 2021

ABSTRACT

Next Generation sequencing (NGS) technologies are revolutionizing the acquisition of genomic data at relatively low cost. NGS technologies are rapidly changing approaches to complex genomic studies and generating a vast amount of data. New and more powerful bioinformatics proposals and tools are needed to handle such large biological collections of such huge amounts of data provided by NGS data analysis. In addition, specialized software tools and advanced computational resources are required for data integration. In this article, we are describing here the main computational approach for next generation sequencing data analysis and its role in agriculture for crop improvement.

Key words: Computational approaches, Genomics, Next generation sequencing data.

çLrkouk

vxyh ih<h dh vuØe.k ¼,uth,l½ çk|kfxfd;k] tks Mh,u,/ vkj,u, v.kvk d Hkhrj U;fDy;kVkbM l d vuØe dh tkp dju d fy, mPp&çokg fof/k;k dh i'kd'k djrh g] vc tfod foKku d vuç;kxk e ,d vko';d midj.k cu xb gA db çkxkfex çfrekuk vkj lefir l,¶Vo;j midj.k dk mi;kx djd dk;kfUor fofHkUu xf.krh; vkj lkf[;dh; rjhd tfod vk.kfod] lyyj ;k ijek.k tkudkjh dk fo'y"k.k vkj 0;k[;k dju e l{ke gA ,uth,l çk|kfxfd;k dk viuku l çkIr ,uth,l MVk dh Hkkjh ek=k dk lkf/kr dju vkj fo'y"k.k dju d fy, dEI;V'kuy –f"Vdk.k vkj to lpuk foKku dh vko';drk gA uVod lk/kuk dk fodflr djuk Hkh vko';d g] tks fd fof'k"V tfod {k=k dk lnfHkr

Email: rahuliasri@gmail.com

[1]Hkk—v-i&jk"Vªh; ikni toçk|kfxdh lLFkku] ubZ fnYyh] HkkjrA

[1]ICAR- National Institute for Plant Biotechnology, LBS Centre, Pusa Campus New, Pusa-110 012, New Delhi, India.

MVkcl] ljfpr MVk lxg vkj oKkfud lenk; }kjk lk>k fd, x, 'kCnlxg lyHkA

,uth,l MVk fo'y"k.k dk dk;çokg

,uth,l –f"Vdk.k d vkonu d lkFk tM ekud dk;çokg] vkfeDl MVk d fo'y"k.k l lcf/kr g] fp= 1 e ;ktukc) gA tfod ueuk l vku oky MVk dk eVkxu,feDl] VklfØiV,feDl ;k thukfeDl ij vk/kkfjr ,d v/;;u dk mi;kx djd vuØfer fd;k tkrk g vkj fQj vlcyh }kjk llkf/kr fd;k tkrk g ;k lnHk MVkcl dk mi;kx djd Hkfo";ok.kh fo'y"k.kk d lh/k lid e yk;k tkrk gA vxy pj.k e vyx&vyx tkp 'kkfey g tl fd Qhpj dh igpku vkj VDlkukfed vlkbueVA

lkekU; :i l thukfed] VklfØlVkfed vkj eVkxukfed v/;;uk e mi;kx fd, tku oky vkiu&lkl l,¶Vo;j vkj lnHk MVkcl uhp of.kr fd; x, g] tk le;≤ ij dN çklfxd fopkj djr g vkj çR;d dk;çokg d fy, mUg vyx djr gA

- **thu igpku ,o vuØe fo'y"k.k pj.k**

vuØe fo'y"k.k fofHkUu U;fDyd ,flM fo'k"krkvk dh ,d mfpr le> dk lnHk nrk g vkj vkt to lpuk foKku d lcl yxkrkj vuç;kxk e l ,d gA çkFkfed vuØe fo'y"k.k d fy, mi;kx fd, tku oky midj.k oKkfud lkfgR; e 0;kid :i l of.kr gA

- **io çlLdj.k pj.k**

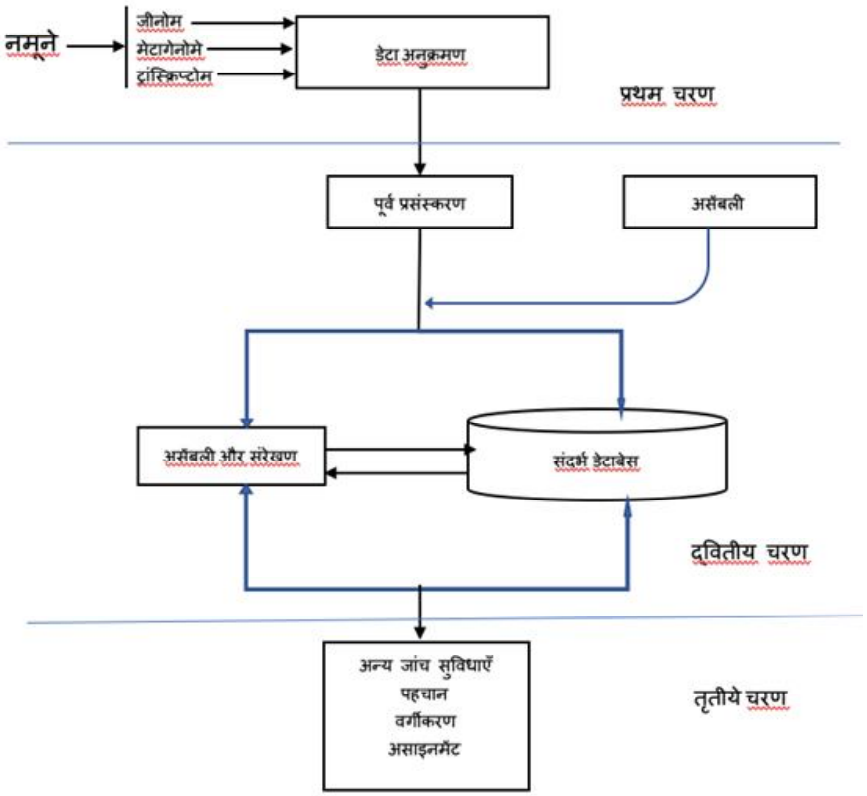
lcl vf/kd bLreky fd, tku oky l,¶Vo;jk e l dN ;gk lphc) gA

1- FASTQC% ,uth,l MVk dh x.koÙkk dh tkp vkj fjikV d fy,A

2- FASTX- VyfdV% vuØe MVk d gjQj vkj mud çk:i :ikrj.k d fy, ,d idtA

- **vlcyh ,o lj[k.k pj.k**

,uth,l fo'y"k.k e lcl egRoi.k pj.kk e l ,d Mh ukok thuke vlcyh gA vlcyh çfØ;k dkQh tfVy g vkj llxr vkj x.koÙkk oky ifj.kke mRiUu dju d fy, vrfufgr rjhdk dh le> vko';d gA vuØek dh ,d cMh ek=k dk lVhd lj[k.k dju d fy, ,d lnHk dk mi;kx djuk gA bu l,¶Vo;j idtk d ihN lkekU; fl)kr lHko lj[k.k LFkkuk dk irk yxku dh dkf'k'k dj



fp= 1% vkfeDl MVk fo'y"k.k dk ekud dk;çokgA

jgk g vkj fQj okLro e lj[k.k dk ijk dj jgk gA uhp] dN lcl vf/kd bLreky dh tku okyh ikjifjd fo/kkuHkk vkj lj[k.k ,Yxkfjne lphc) gA

1½ VELVET% ,d Mh ukok thukfed / VkIfLØIVkfed vlcyj gA

2½ SOAP DE NOVO% ,d M ukok y?k vuØe vlcyj gA

3½ fVfuVh vkj,u,&lhD; MVk dh ,d Mh ukok vlcyh d fy;A

4½ BLAT% ,d BLAST tlk vykbueV Vy] ftl jQjl thuke d vuØe d lj[k.k d fy, fMtkbu fd;k x;k gA

5½ Bowtie% thukfed vuØe e y?k U;fDy;kVkbM vuØek d lj[k.k d fy, ,d dk;ØeA

6½ LVkj ,d vkj,u,&lhd dk thuke ljf[kr dju d fy,A

• **thu Hkfo";ok.kh/ ,ukV'ku pj.k**

mPp x.koÙkk oky ,ukV'ku igy vuØfer thuke d mi;kx dh çHkko'khyrk dk vudyu djr gA okLro e] o thukfed vuØek dk viu tfod dk; l tkMuk lHko cukr gA uhp] dN lcl vf/kd bLreky fd, tku oky vkj ikjifjd ,ukV'ku midj.k fn[kk, x, gA

1½ ,UlEcy thuke ,ukV'ku ,d thu ,ukV'ku ikbiykbu

2½ tuekd ,d thu Hkfo";ok.kh okyk l,¶Vo;j ftle fuxjkuh vkj vn/k&fuxjkuh Vfux 'kkfey gA

3½ ,ulhchvkb thuke ,ukV'ku ,ulhchvkb ¼u'kuy lVj Q,j ck;kVDuky,th bUQ,e'ku½ }kjk tkjh thuke ,ukV'ku ikbiykbuA

fofHkUu lkrk l thukfed vuØe vkj ,ukV'ku MVk dk ,dh—r dju vkj bu MVk d ckmftx] iuikflr vkj fo'y"k.k dh vuefr nu d fy,] thuke ckmtj tl fotvykbt'ku Vy fodflr fd, x, FkA e[; thuke ckmtjk e] ge ,ulcy thuke ckmtj vkj ,ulhchvkb d thuke ckmtj dk mYy[k djr g] tk mijkä l,¶Vo;j l tM gA

• **lnHk MVkcl**

Mh,u,&vuØe.k –f"Vdk.k Økfr u lex vuØe.k ykxr dk cgr de dj fn;k g] ifj.kkeLo:i mRiUu vkdMk dh ek=k e Hkh cMh of) gb gA bld vykok] lxj vuØe.k }kjk fufer yc jhMl dk y?k jhMl ;k mud tkM }kjk cny fn;k x;k FkkA vuØe MVk çc/ku bl çdkj bu ub lfo/kkvk l çHkkfor Fkk] u, to lpuk foKku –f"Vdk.k dh vko';drk gkrh g] ;gk rd fd MVk lxghr dju] n[ku vkj mudk mi;kx dju d lnHk e HkhA u, vkj vf/kd tfVy MVkcl çfØ;k MVk vkj Dojh vuØe MVk d fy, vudfyr bVjQl çnku djuk vko';d gA

uhp] dN lcl vf/kd bLreky fd, tku oky vkj ikjifjd MVkcl dk mYy[k fd;k x;k gA

1- vkjMhih/flYok jkbckIkey vkj,u, thu d fjikftVjh] dk mi;kx dj oxhdj.k Lok;Ùkrk d fy, fd;k tkrk gA

2- dbthth thu d dk;kRed ,ukV'ku d fy, ,dh—r llk/kuA

3- UNIPROT% dk;kRed :il ,ukVV çkVhu vuØek dk ,d MVkclA

4- tucd ,ulhchvkb }kjk fufer] 250]000 l vf/kd çtkfr;k d thuke –';k dk ,d cMk lxg gA MVk dk ,ulhchvkb dh fjVhoy ç.kkyh d ek/;e l ,Dll fd;k tk ldrk g ftl ,UVht dgk tkrk gA lxg e dkfMx vkj vfu;f=r {k=] çekVj] VfeuVj] ,Dl,u] bV,u] fjihV {k= 'kkfey gA

—f"k e ,uth,l MVk fo'y"k.k dh Hkfedk

,uth,l —f"k dk ,d u;h fn'kk çnku dj jgk g vkj ik/kk ,o tkuojk e okNuh; y{k.kk vkj çtuu fu.k;k dh igpku djuk] Vd djuk vkj çlkj djuk lHko cukrk gA nfu;k d 40% [kk| mRiknu l jkx vkj dhV u"V gk x, g] ,uth,l ik/kk dh chekfj;k d lkr dk Vd dju e enn dj jgk g vkj ,lh chekfj;k ;k dhVk dk [kRe dju d mik; l>k jgk gA bll ,l le; e —f"k mRikndrk e l/kkj dh lHkkouk g tc Hkfe mi;kx e deh vk jgh g] nfu;k dh vkcknh yxkrkj c< jgh g] vkj tyok; ifjoru d çHkko l Qlyk d lkeu vku oky tfod vkj vtfod ruko c< jg gA ;g db çtkfr;k l ij thuke dh vuØe.k mud lxBu dk ifjHkkf"kr dju dh vuefr nrk g vkj mudh dk;{kerk dk le>u d fy, 'k#vkrh fcn çnku djrk gA

lcf/kr vk.kfod tkudkjh d mi;ä Kku dh miyfC/k d fy, lckf/kr ç;kl] tl fd VkIfØl'kuy vkj çkfVvke lhDoflx l mRiUu gku oky] thuke dh thu lkexh vkj bldh e[; dk;{kerk dk cgrj <x l fpf=r dju d fy, Hkh vko';d gA bu ç;klk l okLro e lHkh tfod foKku vkj —f"k e Hkh cMh çxfr gbA bld vykok] thu dh tfVyrk vkj mudh uVofdx Hkh ey :il Qlyk ;k i'k/ku d fy, çtuu vH;kl e vuokfnr gku d fy,

ekfyd g] tk mud LokLF;] çfrjk/k vkj mRikndrk e ;kxnku nrk gA

njvly]—f"k e thukfeDl dk ;kxnku fof'k"V çk:i y{k.kk l tM thuk dh igpku vkj gjQj d lkFk&lkFk ofj,V d ekdj&vflLVM p;u }kjk çtuu dju bR;kfn gA e[; mí'; Qlyk ;k i'k/ku thuke d v/;;u d ek/;e l vfHkuo lek/kku [kktuk] [kk| m|kx d fy, lj{k.k vkj fVdkÅ mRikndrk d fy, tkudkjh çkIr djuk gA l;=] feêh vkj i'k/ku ekbØkck;ke Hkh—f"k e egRoi.k Hkfedk fuHkkr g D;kfd ;g ik/k d LokLF; vo feêh d to jklk;fud x.kk dk fu/kkfjr djrk g] vkj mit vkj x.koÙkk y{k.kk nkuk dk çHkkfor djrk gA gkykfd] jkxk.kvk vkj ftu lenk;k e o 'kkfey g] mud fy, cgr de Kku miyC/k gA ,d mnkgj.k d :i e] ;g Lohdkj fd;k tkrk g fd feêh iFoh ij lcl cM dkcu tyk'k;k e l ,d g] vkj çkdfj;kVl feêh ck;kekl dh ,d egRoi.k ek=k dk xBu djr gA

—f"k e eVkxukfeDl dk vuç;kx feêh vkj ik/kk d çdn e l{ethok d lkFk&lkFk fof'k"V Årdk ;k vxk e gku oky bVjD'ku d tfVy iVu dk fpf=r dju d fy, Hkh mi;ä lkfcr gvkA eVkxukfeDl v/;;u ik/kk d ik"k.k e ;k rRok d pØ e feêh d thok.kvk dh Hkfedk dk le>u e Hkh enn dj ldrk gA vkx d vuç;kx u, thu] to&mRiknk] ik/kk d fodkl dk c<kok nu oky l{ethok dh [kkt dk c<kok n ldr g] tk ruko ;k fMfLcvkfll dh çfrfØ;k tl çklfxd igyvk dk le>u d fy, mi;kxh gA nljh vkj] lnHk thuke dh c<rh l[;k u ,dy U;fDy;kVkbM i,yhe,fQTe¼,l,uih½ tl 0;fäxr thuke l mRifjoru dk ,d MVk çkIr dju d fy,] thuke fHkUurk v/;;u dh fn'kk e vul/kku ç;klk dk lckf/kr fd;kA ,l,uih ik/kk e lcl çpj ek=k e miyC/k Mh,u, vuØe fHkUurk dk çdkj g vkj bl cM MVk lxg dk mi;kx—f"k e :fp dh çtkfr;k e dju d fy, d fy, db to lpuk foKku midj.k fodflr fd, x, FkA

fu"d"k

vxyh ih<h d vuØe.k cMh ek=k e MVk mRiUu dj jgk g vkj ,uth,l MVk d fo'y"k.k d fy, dEI;V'kuy ikbiykbu d fodkl dh vko';drk gkrh g] [kkldj tc eYVh&vkfeDl—f"Vdk.k ykx gkr gA vxyh ih<h d vuØe.k MVk d fy, db miU;kl dEI;V'kuy j.kuhfr;k vkj vuç;kxk dk ;gk lckf/kr fd;k x;k gA ;g dEI;V'kuy thukfeDl vkj MVkcl d vuç;kxk e lcf/kr fof/k vkj—f"Vdk.k e l/kkj Hkh djr gA eulfgr] ;g fo'k"k eík 'kk/kdrkvk vkj Nk=k dk vxyh ih<h d vuØe.k MVk fo'y"k.k d ckj e viu Kku dk viMV dju e enn djxkA

InHk

Andrews, K.R. and Luikart, G. (2014). Recent novel approaches for population genomics data analysis. Molecular Ecology. 23(7): 1661-1667.

Del Vecchio, F., Mastroiaco, V., Di Marco, A., Compagnoni, C., Capece, D., Zazzeroni, F. and Tessitore, A. (2017). Next-generation sequencing: recent applications to the analysis of colorectal cancer. Journal of Translational Medicine. 15(1): 246.

Esposito, A., Colantuono, C., Ruggieri, V. and Chiusano, M.L. (2016). Bioinformatics for agriculture in the Next-Generation sequencing era. Chemical and Biological Technologies in Agriculture. 3(1): 9.

Henry, R.J. (2012). Next-generation sequencing for understanding and accelerating crop domestication. Briefings in Functional Genomics. 11(1): 51-56.

Sirangelo, T.M. and Calabrò, G. (2020). Next Generation Sequencing Approach and Impact on Bioinformatics: Applications in Agri-Food Field. Journal of Bioinformatics and Systems Biology. 3: 32-44.

Tang, C.Y., Hung, C.L., Zheng, H., Lin, C.Y. and Jiang, H. (2015). Novel computational technologies for next-generation sequencing data analysis and their applications.

Wadapurkar, R.M. and Vyas, R. (2018). Computational analysis of next generation sequencing data and its applications in clinical oncology. Informatics in Medicine Unlocked. 11: 75-82.