

## Part 1

1.

```
dhaval@DESKTOP-50ECNBT: ~ x + v
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ vi textfile.txt
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ wc textfile.txt
 1  4 16 textfile.txt
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ |
```

## Part 2

2. & 3.

```
dhaval@DESKTOP-50ECNBT: ~ x + v
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ cp ../Lab_session_1/
dhaval/ protein.fasta sequence.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ cp ../Lab_session_1/sequence.fasta ./
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ ls
protein.fasta sequence.fasta sequence1.fasta sequence2.fasta sequence3.fasta sequence4.fasta sequence5.fasta textfile.txt
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ tail -n 4 sequence.fasta
TAACTACTGATAAGTTACAAACTGTTTCTATCCTAAAGGGCAATACAGCCCTAGACTCTCCCAGGTAT
TTGACTCCTGCAGCAAAAAGGAAATTGAGGAAATAGAGCAAGCTATTTCTCAGAGGCAACTATATCAC
TAGACACCCCG
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep ">" sequence
sequence.fasta sequence1.fasta sequence2.fasta sequence3.fasta sequence4.fasta sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep ">" sequence5.fasta
>ahr
>clock
>hif1a
>hif2a
>hif3a
>npas1
>npas2
>npas3
>npas4
>sim1
>sim2
>arnt1
>bmal1
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "A.G" sequence5.fasta
IFRTKHKLDFTPIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
DAARRSRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNOVGAGGEPLDACYL
KALEGFVMVLTAEGDMAYLSENVSKHLGLSQLEIGHSTFDFIHPCDQELQDALTPPTERCFSLRMKST
KEKSRNAARRSRGKENLEFFELAKLLPLPGAITSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEKGFLYISETVSIYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKAAGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
```

4. & 5.

```
dhaval@DESKTOP-50ECNBT: ~$ cat >sim1
>sim2
>arnt1
>bmal1
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "A.G" sequence5.fasta
IFRTKHLDFPTIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVAGGEPDADCYL
KALEGFVMVLTAEGDMAYLSENVSKHLGLSQLEIGHISIFDFIHPCDQEELQDALTPPTERCFLRMKST
KEKSRNAARSRRKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKAAGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
WLQRAAGGFVWLQSVATVAGSGKSPGEHVLWVSHVLSQAEGGQT
GASVARRDQINAEIRNLKELLPLAEADKVRLSYLHIMSLACIYTRKGVFFAGGTPLAGPTGLLSAQELE[Close
IVAALPGFLLVFTAEGLLSESVSEHLGHSMDVLVAQGDSTYDIIIDPADHLTVRQQLTLTDRLFRCRF
EKSNAARTREKENSEFELAKLLPLPSAITSQDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
EIERSSFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSMFMFRASL
EKSNAARTREKENSEFELAKLLPLPSAITSQDKASIIRLTTSYLKMRVVFPEGLGEAWGQPSRAGP
EIERSSFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGQSLPPSAITEIKLHSMFMFRASL
ELKHLILEAADGFLFVSCETGRVVYVSDSVTPVLNQPSQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
GSRSSFICMRGCSSEPHFVVHCTGYIAKFCFLVAIGRLQVTSSPNCMDMSNVCPQTEFISRHNIEGIF
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIKVKQELSSSRLC
SGARRSFFCRMKCNRPKSFCTIHTGYLKSNSCLVAIGRLHSHVVPQPVNGEIRVKSMYVSRHAIDG
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "P[*A]L" sequence5.fasta
QLHWIIPPENSLMERCFCIRLCLLDNSSGFLAMNFQCKLVLPQALFAIATPLQPPSILEIRTWFM
MRMKTCTVNRGRVTNLSKATWVHLCTGQVQVYEPPLLSCLIIMCEPIQHPSHMDIPLDSKTFLSRHSMDH
LTSRGRITNLKAAATWVNLNCSGHMAYEPPLQCLVLICEAIPHGSLPEPLGRGAFLSRHSIDMKFTYCD
FTQLMLEALDGFIIAVTTDGSIIYVSDSITPLLGHLPDSDVMDQNLNLFPEQEHSEVYKILSEYLVKSDS
ELKHLILEAADGFLFVSCETGRVVYVSDSVTPVLNQPSQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "VV" sequence5.fasta
AANFREGLNLQEGEFLQLALNGFVLVVTDDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IWLQTHYYITYHQWNSRPEFIVCTHTVVSVAEVRAE
TVIYNTKNSQPQCIVCNVYVSGIIQHDL
QMDNLYLKALEGFIAVVTQDGDIMFLENISKFMGLTQVELTGHSIFDFTHPCDHEIRENLSSTERDFF
KFTYCDORITELIGYHPEELLGRSAEYFYHALDSENMTKSHQNLCTKGQVSGQYRMLAKHGGYVMLETQ
DRIAEGVAGYSPDDLIGCSAYEYIHALSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQATQAVVSG
QTHYYITYHQWNSKPEFIVCTHSVVSADVRVE
DYVHPGDHVEMAEQMLTSLERSFFIRMKSTLTKRGVHKISSGYKVIHITGRLRLRMLGVVAHALPPPTI
ISESVLYLGFERSSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEMVRLQAKTGGWAWIYCLLY
EKSNAARTREKENSEFELAKLLPLPSAITSQDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
LDNVGRELGSHELLQTLDFGFVVPADGKIMYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
LDGVAHELGSHELLQTLDFGFVVASDGKIMYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
SYATVVHNSRRSRPHCIVSVNYVLTEIEYKEL
ELKHLILEAADGFLFVSCETGRVVYVSDSVTPVLNQPSQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
GSRSSFICMRGCSSEPHFVVHCTGYIAKFCFLVAIGRLQVTSSPNCMDMSNVCPQTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLRDSFQOVVVLKGQVLSVMFRFRSKNQEWLWMRTSS
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIKVKQELSSSRLC
SGARRSFFCRMKCNRPKSFCTIHTGYLKSNSCLVAIGRLHSHVVPQPVNGEIRVKSMYVSRHAIDG
RWFSSMNPWTKEVEYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLNLQEGEFLQLALNGFVLVVTDDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IFRTKHLDFPTIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
RHSLEWKFLLDHRAPPIIGYLPFEVLGTSGDYHYHDDLLENLAKCHELMQYKGKSGCYRFLTKGQQW
KEKSRDAARSRRSKESEVFEYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLIEDDMKAQM
NCFYLKALDGFVMVLTDGDMYIISDNVNKMVGLTQFELTGHSVDFDTHPCDHEEMREMLTHNTQRSFQL
KEKSRDAARCRRSKETEVFEYELAHQLPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSSENESEAEADQ
KFTYCDORITELIGYHPEELLGRSAEYFYHALDSENMTKSHQNLCTKGQVSGQYRMLAKHGGYVMLETQ
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVAGGEPDADCYL
LTSRGRITNLKAAATWVNLNCSGHMAYEPPLQCLVLICEAIPHGSLPEPLGRGAFLSRHSIDMKFTYCD
DRIAEGVAGYSPDDLIGCSAYEYIHALSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQATQAVVSG
KEKSRNAARSRRKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
```

6. & 7.

```
dhaval@DESKTOP-50ECNBT: ~$ cat >sim1
>sim2
>arnt1
>bmal1
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "VV" sequence5.fasta
AANFREGLNLQEGEFLQLALNGFVLVVTDDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IWLQTHYYITYHQWNSRPEFIVCTHTVVSVAEVRAE
TVIYNTKNSQPQCIVCNVYVSGIIQHDL
QMDNLYLKALEGFIAVVTQDGDIMFLENISKFMGLTQVELTGHSIFDFTHPCDHEIRENLSSTERDFF
KFTYCDORITELIGYHPEELLGRSAEYFYHALDSENMTKSHQNLCTKGQVSGQYRMLAKHGGYVMLETQ
DRIAEGVAGYSPDDLIGCSAYEYIHALSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQATQAVVSG
QTHYYITYHQWNSKPEFIVCTHSVVSADVRVE
DYVHPGDHVEMAEQMLTSLERSFFIRMKSTLTKRGVHKISSGYKVIHITGRLRLRMLGVVAHALPPPTI
ISESVLYLGFERSSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEMVRLQAKTGGWAWIYCLLY
EKSNAARTREKENSEFELAKLLPLPSAITSQDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
LDNVGRELGSHELLQTLDFGFVVPADGKIMYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
LDGVAHELGSHELLQTLDFGFVVASDGKIMYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
SYATVVHNSRRSRPHCIVSVNYVLTEIEYKEL
ELKHLILEAADGFLFVSCETGRVVYVSDSVTPVLNQPSQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
GSRSSFICMRGCSSEPHFVVHCTGYIAKFCFLVAIGRLQVTSSPNCMDMSNVCPQTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLRDSFQOVVVLKGQVLSVMFRFRSKNQEWLWMRTSS
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIKVKQELSSSRLC
SGARRSFFCRMKCNRPKSFCTIHTGYLKSNSCLVAIGRLHSHVVPQPVNGEIRVKSMYVSRHAIDG
RWFSSMNPWTKEVEYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLNLQEGEFLQLALNGFVLVVTDDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IFRTKHLDFPTIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
RHSLEWKFLLDHRAPPIIGYLPFEVLGTSGDYHYHDDLLENLAKCHELMQYKGKSGCYRFLTKGQQW
KEKSRDAARSRRSKESEVFEYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLIEDDMKAQM
NCFYLKALDGFVMVLTDGDMYIISDNVNKMVGLTQFELTGHSVDFDTHPCDHEEMREMLTHNTQRSFQL
KEKSRDAARCRRSKETEVFEYELAHQLPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSSENESEAEADQ
KFTYCDORITELIGYHPEELLGRSAEYFYHALDSENMTKSHQNLCTKGQVSGQYRMLAKHGGYVMLETQ
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVAGGEPDADCYL
LTSRGRITNLKAAATWVNLNCSGHMAYEPPLQCLVLICEAIPHGSLPEPLGRGAFLSRHSIDMKFTYCD
DRIAEGVAGYSPDDLIGCSAYEYIHALSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQATQAVVSG
KEKSRNAARSRRKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
```

```
dhaval@DESKTOP-50ECNBT: ~$ cat sequence5.fasta
ISESVLYTLGFERSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEVVVRLOAKTGGWAWIYCLLY
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
LDNVGRELGSHELLQTLDFGFIWVAPDGIKIYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
LDGVAKELGSHLLQTLDFGVFWVADSGKIMYISETASVHLGLSQVELTGNISYIYHPADHDEMTAVLTA
SYATVVHNSRSSRPHCIVSVNYVLTEIYKEL
ELKHLILEAADGFLFIVSCETGRVVVSDSVTPVLNQPSSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
GSRRSFICRMRCGSSEPHFVVVHCTGYIAKFCFLVAIGRLQVTSPPNCTDMSNVCQPTFISRHNIIEGIF
TFVDHRCVATVGYQPQLLGNIVFECHPEDQQLRDSFQVVKLGQVLSVMFRFRSKNQEWLWMRTSS
DELKHLILRAADGFLFVVGCDRGKILFVSESVKILNYSQNDLIGQSLFDYLPKDIKVKKEQLSSSRLC
SGARRSFFCRMCMKRNPRKSCTIHTSYGLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMYVSRHAIDG
RWFSSMNPWTKEVEYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLLNQEGFLLQALNGFVLVVTDDALVFYASSTIQDYLGQQSDVIHQSVVELIHTEDRAEFQR
IFRTKHLDFTPIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIIVFRLLT
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGDYHYHVDLENLAKCHEHLMQYGGKSCYYRFLTKGQQW
KEKSRDAARSRRSKSESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLIEDDMAQM
NCFYLKALDGFVMVLTDDGDMYISDNVNKYMGLTQFELTGHSVDFDTHPCDHEEMREMLTHNTQRSFLL
KEKSRDAARCRRSKETEVEFYELAHQLPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSSESEAEADQ
KFTYCDORITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSQGYRMLAKHGGYVWLETD
DAARSRRSQTEVLVYQLAHTLPFARGVSAHLKASIMRLTISYLRMHRLCAAGEWNVQVAGGEPDACYL
LTSRGRTLNLKAAATWIKVNLCSGHMRAVEPPLQCLVLICEAIPHPGSLPEPLGRGAFLSRHSIDMKFTYCD
DRIAEGVAGYSPDOLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSQ
KEKSRNAARSRRGKENLEFFELAKLLPLPGAITSQDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAAALVSEVFEQHLGGHILQSLDGFVFNQEGKFLYISETVSYLGLSQVEMTGSSVFDYI
LEWKFLFLDHRAPPIIGYLPFEVLGTSGDYHYHVDLENLARCHQHLMQFGKGKSCYYRFLTKGQQWILW
SRDAARSRRGKENFEFYELAKLLPLPAATISQDKASIIRLTISYLMKRVFPEGLGEAWGHSSRTSP
IVAAPGFLLVFTAEGKLLYLSVSEVSEHLGSHMVDLVAQGDYIYDIPADHLTVRQQLTLTDLRFRCRF
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
ELKHLILEAADGFLFIVSCETGRVVVSDSVTPVLNQPSSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
DELKHLILRAADGFLFVVGCDRGKILFVSESVKILNYSQNDLIGQSLFDYLPKDIKVKKEQLSSSRLC
KFVFDQRATAILAVLPQELLGTSCYEFYHQDDIGHLAECRQVLQTREKITTNCYFKIKDGSFITLRS
```

8. & 9.

```
dhaval@DESKTOP-50ECNBT: ~$ cat sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLLNQEGFLLQALNGFVLVVTDDALVFYASSTIQDYLGQQSDVIHQSVVELIHTEDRAEFQR
IFRTKHLDFTPIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIIVFRLLT
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGDYHYHVDLENLAKCHEHLMQYGGKSCYYRFLTKGQQW
KEKSRDAARSRRSKSESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLIEDDMAQM
NCFYLKALDGFVMVLTDDGDMYISDNVNKYMGLTQFELTGHSVDFDTHPCDHEEMREMLTHNTQRSFLL
KEKSRDAARCRRSKETEVEFYELAHQLPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSSESEAEADQ
KFTYCDORITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSQGYRMLAKHGGYVWLETD
DAARSRRSQTEVLVYQLAHTLPFARGVSAHLKASIMRLTISYLRMHRLCAAGEWNVQVAGGEPDACYL
LTSRGRTLNLKAAATWIKVNLCSGHMRAVEPPLQCLVLICEAIPHPGSLPEPLGRGAFLSRHSIDMKFTYCD
DRIAEGVAGYSPDOLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSQ
KEKSRNAARSRRGKENLEFFELAKLLPLPGAITSQDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAAALVSEVFEQHLGGHILQSLDGFVFNQEGKFLYISETVSYLGLSQVEMTGSSVFDYI
LEWKFLFLDHRAPPIIGYLPFEVLGTSGDYHYHVDLENLARCHQHLMQFGKGKSCYYRFLTKGQQWILW
SRDAARSRRGKENFEFYELAKLLPLPAATISQDKASIIRLTISYLMKRVFPEGLGEAWGHSSRTSP
IVAAPGFLLVFTAEGKLLYLSVSEVSEHLGSHMVDLVAQGDYIYDIPADHLTVRQQLTLTDLRFRCRF
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
EKSMAARTTRREKENSEFYEALAKLLPLPSAITSQDKASIIRLTTSYLMKRVFPEGLGEAWGHSSRTSP
ELKHLILEAADGFLFIVSCETGRVVVSDSVTPVLNQPSSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
DELKHLILRAADGFLFVVGCDRGKILFVSESVKILNYSQNDLIGQSLFDYLPKDIKVKKEQLSSSRLC
KFVFDQRATAILAVLPQELLGTSCYEFYHQDDIGHLAECRQVLQTREKITTNCYFKIKDGSFITLRS
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -V "^>" sequence5.fasta | grep "P"
License GPLv3+: GNU GPL version 3 or later <https://gnu.org/licenses/gpl.html>.
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -v ">" sequence5.fasta | grep "P"
SNPSKRHRDLNTELDRLASLPPQDVINKLKLVLRLSVSLRAKSFDOALKSSPTERNGGQDNCR
QLHWQIPPEPSPLMERCFCIRLCLLDNNSGFLAMNFQGLKYLPPQLALFAIATPLQPPSILEIRTKNF
IFRTKHLDFTPIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIIVFRLLT
KNNRWTVQSNARLLYKNGRPDIYITQRPDLTDEEGTEHLR
VSRNKEKKRRDQFNVLIKELGSMPLGNARKMKDKSTVLQKSIDFLRKHEITAQSDASEIRQDWKPTFLS
NEEFQTLMLAALDGFLLAINTDGSIIYVSESVTSLLEHLPSDLVDQSIFFNFIPEGEHSEVYKILSTEYVK
SKNQLFCCHMLRGTDPKPSYEVYKFGNFKSLYEDRVCFVATVRLATPQFIKEMCTVEEPNEEFTS
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGDYHYHVDLENLAKCHEHLMQYGGKSCYYRFLTKGQQW
```

```
dhaval@DESKTOP-50ECNBT: ~$ grep -v ">" sequence5.fasta | grep "P"
SNPSKRRDRRLMTELDRLASLPFPQDVINKLDKLSVLRRLSVSYLRAKSFDFVALKSSPTERNGGQDNCR
QLHWQIIPENSPLMERCFTICRLCLLDNSSGFLAMNFQGLKVLPPQLALFAIATPLQPPSILEIRTKNF
IFRTKHKLDFTPIGCDAGRIVLGYTEALCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
KNNRWTVQSNARLLYKNGRPDIYIVTQRPLTDEEGTEHLR
VSRNKSSEKKRRDQFNVLIKELGSMPLGNARKMKDSTVLQKSIDFLRKHKIEITAQSDASEIRQDWKPTFLS
NEEFTQLMLEALDGFIAIMTDSIIYVSESVTSLLEHLPSDLVDQSIENFIPEGEHSEVYKILSTEYLK
SKNQLEFCCHMLRGTDIPKEPSTYEYVKFIGNFKSLYEDRVCFVATVRLATPQFIKEMCTVEEPEEFTS
RHSLEWKFLDHRAPPIIGYLPFEVLGTSYDYHYHDDLENLAKCHEHLMQYGGKSCYRFLTKGQQW
IWLQTHYYITYHQWNSRPEFIVCTHTVVSYAEVRAE
KEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLIEDDMKAQM
NCFYLKALDGFVMVLTDGDMYIYSDNVNKMGLTQFELTGHVSFDFTHPCDHEEMREMLTHNTORSFLL
RMKCTLTSRGTMMIKSATMKVHLCTGHIHVYKPPMTCLVLICEPIPHPSNIETPLDSKTLRSHSLDMK
FSYCDERITELMGVEPEELLGRSIEYFYHALDSOHLTKTHHDMFTKGQVTTGQYRMLAKRGGYVWVETQA
TVIYNTKNSQPQCIVCVNVVSGIYQHDLL
KEKSRDAARCRRSKETEVFELAHQLPLPHNVSSHLDKASIMRLAISFLRTHKLLSSVCSENESEAEADQ
QMDNLYLKALDGFIAVTVQDGMILFENISKFMGLTQVELTGHVSFDFTHPCDHEEMREMLTSSTERDFF
MRMKCTVTNRGRTVNLKSAWKVHLCTGQVKVYEPILLSCLIMCEPIQHPSHMDIPLDSKTLRSHSLDMK
KFTYCDORITELIGYHPEELLGRSIEYFYHALDSOHLTKTHHDMFTKGQVTTGQYRMLAKRGGYVWVETQA
GTVIYNPNLQPPQIMCVNVVSEIEKNDV
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGENQVAGGEPDLACYL
KALDGFVMVLTAEGDMAYLSENVSKHLGLSQLELIGHSIFDFIHPDQEEQLQDALTPPTERCFLRMKST
LTSRGRTLNLKAAATMKVNLNCSGHMRAYEPPLQLCLVLICEAPIHPSGLEPPLGRGAFLRSHSLDMKFTYCD
DRIAEVAGYSPDILLIGSAEYFYHALDSOHLTKTHHDMFTKGQVTTGQYRMLAKRGGYVWVETQA
GRGPQSEISVCHFLISQVEETGV
KEKSRDAARSRRRKENLEFFELAKLLPLPGAITSQDKASIVRLSVTYLRRLRFAALGAPPWGLRAAGPP
AGLAPGRRGPAAALVSEVFEQHLGGHILQSLDGFVAFALNQEKGFLYISETVSYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKASGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
LHGHIIVFRLSLGLTILACESRVSDHMDLGPSELVGRSCYQFVHGQDATIRQSHVDLLDKGQVMTGYVR
WLORAGGFVWLOQSVATVAGSGKSPGEHHVLWVSHVLSQAEGGQT
NKSEKRRDQFNVLIKELSSMLPGNTRKMDKTTVLEKVIQKHNESVSAQTEICDIQQDWKPSFLSNEE
FTQLMLEALDGFIAVTVQDGMILFENISKFMGLTQVELTGHVSFDFTHPCDHEEMREMLTSSTERDFF
DLEFYCHLLRGSLLNPKFPTYEYIKFVGNFRSYLKGVECFIATVRLATPQFLKEMCIVDEPLEEFTSRHS
LWKFLDHRAPPIIGYLPFEVLGTSYDYHYHDDLENLARCHQHLMQYGGKSCCYRFLTKGQQW
IWLQTHYYITYHQWNSRPEFIVCTHSVSYADVRV
SRDAARSRRRKENLEFFELAKLLPLPAATISQDKASIMRLTISYLRMHRLCAAGENQVAGGEPDLACYL
SVKVIQAQRRRSPSALATEVFEAHLGSHILQSLDGFVAFALNQEKGFLYISETVSYLGLSQVEMTGSSV
DYVHPGDHVEMAEQMLTSLERSFFIRMKSTLTKRGLHVKASGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
NEVRIDCHMFTVRVMDLNIITYCENRISDYMDLTPVDIVGKRCYHFIHAEDVEGIRSHSLDLLNKGQCVT
KYRWMQKNGGYIWIQSSATIATNAKNANEKNIWVNYLLSNPEYKDT
GASKARRDQINAEIRNLKELLPLAEADKVRSLYHIMSLACIYTRKGVFFAGGTPLAGPTGLLSAQELED
IVAALPGFLLVFTAEGKLLYLSESVSEHLGSHMVDLVAQGDSDIYIDPADHLTVRQQLTLTDRLFCRFR
NTSKSLRRQSAGNKLVLIRGRFAHNPVFTAFCALEPRPRPGPGPGPASLFLAMFQSRHAKDALLD
ISESVLYLGFERSSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEVMVRLQAKTGGAWIYCLLY
SEGPEGPITANNYPISDMEAWSLRQQL
EKSNAARTREKENSEFYELAKLLPLPSAITSQDKASIMRLTISYLRMHRLCAAGENQVAGGEPDLACYL
LDNVGRELGSHLLQTLDFGFIYVAPDGKIIMYISETASVHLGLSQVELTGNISYIYIHPADHDEMTAVLTA
EIERSSFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
DMWLIFLDSRVAELTGYEPQDLIEKTYHHVHGCDTFHLRCAHLLLVKGQVTTKYRFLAKHGGWVWVQ
SYATVHNSRSSRPHCIVSVNYVLTDEYKGL
EKSNAARTREKENSEFYELAKLLPLPSAITSQDKASIMRLTISYLRMHRLCAAGENQVAGGEPDLACYL
LDNVGRELGSHLLQTLDFGFIYVAPDGKIIMYISETASVHLGLSQVELTGNISYIYIHPADHDEMTAVLTA
EIERSSFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
DLKLIFLDSRVTETGYEPQDLIEKTYHHVHGCDTFHLRCAHLLLVKGQVTTKYRFLAKHGGWVWVQ
SYATVHNSRSSRPHCIVSVNYVLTDEYKGL
NHSEIERRRRNMKTAYITELSDMVPKCSALARKPDKLTILRMVSHMKSLRGTGNTSDGSYKPSFLTDQ
ELKHLILEAADGFLFVSCETGRVYVSDSVTVLNPQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
GSRRSFICRMRCGSEPHFVVHCTGYIAKFCFLVAIGRLQVTSSPNCTDMSNVCPTEFISRHNIIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQVQVVKLGQVLSVMFRFRSKNQEWLWMRTSS
FTFQNPYSDEIEYIICNTNVK
EAHSQIEKRRDQINAEIRNLKELLPLAEADKVRSLYHIMSLACIYTRKGVFFAGGTPLAGPTGLLSAQELED
DLEKHLILEAADGFLFVSCETGRVYVSDSVTVLNPQSEWFGSTLYDQVHPDDVDKLRQLSTSRMCM
SGARRSFFCRMKNRPRKSFCTIHTGYLKSNSCLVAIGRLHSHVVPQVNGEIRVKSMEYVSRHAIDG
```

### Part 3

9., 10., 11.



```
dhaval@DESKTOP-50ECNBT: ~
GSRRSFTICMRGSGSEPHFVVVHCTGYIKAFCLVAIGRLQVTSSPNCNTDMSNVCPQTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGHNIVEFCHPEDQQLLRDSFQQVVKLGQVLSVMFRRSKNQEWLWMRTSS
FTFQNPVSYDEIEYIICNTNVK
EAHSQIEKRRRDKMNSFIDELASLVPCTNAMSRLDKLTVLRMAVQHMKTLRGATNPYTEANYKPTFLSD
DELKHLILRAADGFLFVVGCDRGKILFVSESVFILNYSQNDLIGQSLFDYLHPKDIKVKQELSSSRLC
SGARRSFFCRMKCNRPKRSFCTIHTSYLKSNLSCSLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
KFVFVDQRATAILAYLPQELLGTSCYEYFHQDDIGHLAECRQVLQTREKITTNCYKFKIKDGSFITLRS
RWFSSMNPWTKREYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ seq="sequence5.fasta"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ echo "Number of sequences in $seq:"
Number of sequences in sequence5.fasta:
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -c $seq
^[A^[[D^[[c^C
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -c ">" $seq
13
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ pattern="G/{2/}"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep $pattern protein.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ pattern="G/{2,/}"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep $pattern protein.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ pattern="G/{2,}"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep $pattern protein.fasta
grep: Unmatched \{
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ pattern="G/{2,}"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep $pattern protein.fasta
KPVKKKKIKREIKILENRGGPNIIITLADIVKDPVSRTPALVFEHVNNTDFKQLYQTLTDYDIRFMYEI
WERFVHSENQHLVSPEALDFLDKLLRYDHQSRLTAREAMEHPYFTVVQDQARMGSSSMPGGSTPVSSAN
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ myvar="Biocomputing"
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ export myvar
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ bash -c 'echo myvar'
myvar
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ bash -c 'echo $myvar'
Biocomputing
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ |
```

## Part 4

12.

```
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ wc -l sequence3.fasta
19 sequence3.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ echo *.fasta
protein.fasta sequence.fasta sequence1.fasta sequence2.fasta sequence3.fasta sequence4.fasta sequence5.fasta
```

13.

```
dhaval@DESKTOP-50ECNBT: ~
" file size (characters): $(wc -c < "$file") done
> for file in *.fasta; do echo "Processing: $file" echo "sequences: $(grep -c ">" "$file)" echo " file size (characters): $(wc -c < "$file") done"C
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for file in *.fasta; do echo "Processing: $file" echo "sequences: $(grep -c ">" "$file)" echo
" file size (characters): $(wc -c < "$file") done
> for file in *.fasta; do echo "Processing: $file" echo "sequences: $(grep -c ">" "$file)" echo " file size (characters): $(wc -c < "$file") done"C
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for file in *.fasta; do
echo "Processing: $file"
echo "Sequences: $(grep -c ">" "$file)"
echo "File size (characters): $(wc -c < "$file)"
done
Processing: protein.fasta
Sequences: 1
File size (characters): 467
Processing: sequence.fasta
Sequences: 1
File size (characters): 79551
Processing: sequence1.fasta
Sequences: 1
File size (characters): 974
Processing: sequence2.fasta
Sequences: 4
File size (characters): 1710
Processing: sequence3.fasta
Sequences: 2
File size (characters): 1000
Processing: sequence4.fasta
Sequences: 4
File size (characters): 2374
Processing: sequence5.fasta
Sequences: 13
File size (characters): 4229
```

Note – I took help from Chat-gpt for arrange this code.

14.

```
Processing: sequence3.fasta
Sequences: 2
File size (characters): 1000
Processing: sequence4.fasta
Sequences: 4
File size (characters): 2374
Processing: sequence5.fasta
Sequences: 13
File size (characters): 4229
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for files in *.fasta: do
> if [ $(grep -c ">" "$file") -gt 3 ]; then
-bash: syntax error near unexpected token `if'
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for files in *.fasta: do; if [ $(grep -c ">" "$file") -gt 3 ]; then echo "Processing: $file";
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
-bash: syntax error near unexpected token `if'
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for files in *.fasta: do; if [ $(grep -c ">" "$file") -gt 3 ]; then echo "Processing: $file";
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
-bash: syntax error near unexpected token `if'
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for files in *.fasta: do; if [ $(grep -c ">" "$file") -gt 3 ]; then echo "Processing: $file";
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
Processing: sequence2.fasta
Sequences: 4
File size (characters): 1710
Processing: sequence4.fasta
Sequences: 4
File size (characters): 2374
Processing: sequence5.fasta
Sequences: 13
File size (characters): 4229
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ |
```

## Part 5

15.

```
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
-bash: syntax error near unexpected token `if'
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for files in *.fasta: do; if [ $(grep -c ">" "$file") -gt 3 ]; then echo "Processing: $file";
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
-bash: syntax error near unexpected token `if'
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ for file in *.fasta; do if [ $(grep -c ">" "$file") -gt 3 ]; then echo "Processing: $file";
echo "Sequences: $(grep -c "^>" "$file)"; echo "File size (characters): $(wc -c < "$file)"; fi; done
Processing: sequence2.fasta
Sequences: 4
File size (characters): 1710
Processing: sequence4.fasta
Sequences: 4
File size (characters): 2374
Processing: sequence5.fasta
Sequences: 13
File size (characters): 4229
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -v "^>" sequence5.fasta | grep "C.*C.*C"
QLHWQIPPNSPLMERCFCICRLCLLDNSSGFLAMNFQGLKYLPPQLALFAIATPLQPPSILEIRTKNF
IFRTHKHLDFTPIGCDAGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
SKNQLFEFCCHMLRGITDPKEPSTYEVYKFIGNFSLYEDRVCFVATVRLATPQFIKEMCTVEEPNEEFTS
RMKCTLTSRGRMTNIIKSAWKVLHCTGHIHVYKPPMTCLVLICEPIPHPSNIEIPLDSKTLFSRHSMDM
WRMKCTVTNRGRVTNIIKSAWKVLHCTGQVAVYEPPLSCLTIIMCEPIQHPSHMDIPLDSKTLFSRHSMDM
LTSRGRITNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPLGRGAFLSRHSMDMKFTYCD
DLEFYCHLLRGLSNPKFEPTYEYIKFVGNFRSYLGKEVCFIATVRLATPQLKEMCIVDEPLEEFTSRHS
LEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYHYHIDLELLARCHQLMQFGKGKSCCYRFLTKGQQWIWL
NEVRIDCHMFVTRVNMMDLNIYCENRISDYMDLTPVDIVGKRCYHFIHAEDVEGIRHSHLDLLNKGQCVT
EIERSSFFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
EIERSSFFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
GSRRSFICRMRCGSSEPHFVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCOPTFISRHNIIEGIF
SGARRSFFCRMRCNRPKRSFCTIHTSYGLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
KFWFVDQRQATAILAYLPQELLGTSCYEFHQDDIGHLAECRQVLQTRKITTNCYKFKIKDGSFITLRS
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -v "^>" sequence5.fasta | grep "C.*C.*C" | sed '/^$/d' > cys_rich.txt
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ |
```

Note - sed command I got from Chat-gpt.