Part 1

1.

```
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:~/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
TAACTACTGATAAGTTACAAAACTGTTTTCTATCCTAAAGGGCAATACAGCCCTAGACTCTCCCAGGTAT
TTGACTCCTGCAGCAAAAAGGGAAATTGAGGAAATAGAGCAAGCTATTTCTCAGAGGCAACTATATCACA
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
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
KALEGFVMVLTAEGDMAYLSENVSKHLGLSQLELIGHSIFDFIHPCDQEELQDALTPPTERCFSLRMKST
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKASGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
```

4. & 5.

```
>sim1
>sim2
>arnt1
>bmal1
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "A.G" sequence5.fasta
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
KALEGFVMVLTAEGDMAYLSENVSKHLGLSQLELIGHSIFDFIHPCDQEELQDALTPPTERCFSLRMKST
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKASGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
WLQRAGGFVWLQSVATVAGSGKSPGEHHVLWVSHVLSQAEGGQT
GASKARRDQINAEIRNLKELLPLAEADKVRLSYLHIMSLACIYTRKGVFFAGGTPLAGPTGLLSAQELEI[Close]
IVAALPGFLLVFTAEGKLLYLSESVSEHLGHSMVDLVAQGDSIYDIIDPADHLTVRQQLTLTDRLFRCRF
EKSKNAARTRREKENSEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
EIERSFFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
EKSKNAAKTRREKENGEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRAVFPEGLGDAWGQPSRAGP
EIERSFFLRMKCVLAKRNAGLTCSGYKVIHCSGYLKIRIVGLVAVGQSLPPSAITEIKLYSNMFMFRASL
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
GSRRSFICRMRCGSSEPHFVVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIF
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
SGARRSFFCRMKCNRPRKSFCTIHSTGYLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "P[^A]L" sequence5.fasta
QLHWQIPPENSPLMERCFICRLRCLLDNSSGFLAMNFQGKLKYLPPQLALFAIATPLQPPSILEIRTKNF
MRMKCTVTNRGRTVNLKSATWKVLHCTGQVKVYEPLLSCLIIMCEPIQHPSHMDIPLDSKTFLSRHSMDM
LTSRGRTLNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPPLGRGAFLSRHSLDMKFTYCD
FTQLMLEALDGFIIAVTTDGSIIYVSDSITPLLGHLPSDVMDQNLLNFLPEQEHSEVYKILSSEYLKSDS
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "VV" sequence5.fasta
AANFREGLNLQEGEFLLQALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IWLQTHYYITYHQWNSRPEFIVCTHTVVSYAEVRAE
TVIYNTKNSQPQCIVCVNYVVSGIIQHDL
QMDNLYLKALEGFIAVVTQDGDMIFLSENISKFMGLTQVELTGHSIFDFTHPCDHEEIRENLSSTERDFF
```

6. & 7.


```
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep "VV" sequence5.fasta
AANFREGLNLQEGEFLLQALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IWLQTHYYITYHQWNSRPEFIVCTHTVVSYAEVRAE
TVIYNTKNSQPQCIVCVNYVVSGIIQHDL
QMDNLYLKALEGFIAVVTQDGDMIFLSENISKFMGLTQVELTGHSIFDFTHPCDHEEIRENLSSTERDFF
KFTYCDDRITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSGQYRMLAKHGGYVWLETQ
DRIAEVAGYSPDDLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSG
QTHYYITYHQWNSKPEFIVCTHSVVSYADVRVE
DYVHPGDHVEMAEQLGMTLERSFFIRMKSTLTKRGVHIKSSGYKVIHITGRLRLRMGLVVVAHALPPPTI
ISESVLIYLGFERSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEMVVRLQAKTGGWAWIYCLLY
EKSKNAARTRREKENSEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
LDNVGRELGSHLLQTLDGFIFVVAPDGKIMYISETASVHLGLSQVELTGNSIYEYIHPADHDEMTAVLTA
LDGVAKELGSHLLQTLDGFVFVVASDGKIMYISETASVHLGLSQVELTGNSIYEYIHPSDHDEMTAVLTA
SYATVVHNSRSSRPHCIVSVNYVLTEIEYKEL
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
GSRRSFICRMRCGSSEPHFVVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLKGQVLSVMFRFRSKNQEWLWMRTSS
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
SGARRSFFCRMKCNRPRKSFCTIHSTGYLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
RWFSFMNPWTKEVEYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLNLQEGEFLLQALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHVDDLENLAKCHEHLMQYGKGKSCYYRFLTKGQQW
KEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLDIEDDMKAQM
NCFYLKALDGFVMVLTDDGDMIYISDNVNKYMGLTQFELTGHSVFDFTHPCDHEEMREMLTHNTQRSFFL
KEKSRDAARCRRSKETEVFYELAHELPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSENESEAEADQ
KFTYCDDRITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSGQYRMLAKHGGYVWLETQ
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
LTSRGRTLNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPPLGRGAFLSRHSLDMKFTYCD
DRIAEVAGYSPDDLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSG
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
```

```
ISESVLIYLGFERSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEMVVRLQAKTGGWAWIYCLLY
EKSKNAARTRREKENSEFYELAKLLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
LDNVGRELGSHLLQTLDGFIFVVAPDGKIMYISETASVHLGLSQVELTGNSIYEYIHPADHDEMTAVLTA
LDGVAKELGSHLLQTLDGFVFVVASDGKIMYISETASVHLGLSQVELTGNSIYEYIHPSDHDEMTAVLTA
SYATVVHNSRSSRPHCIVSVNYVLTEIEYKEL
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
GSRRSFICRMRCGSSEPHFVVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLKGQVLSVMFRFRSKNQEWLWMRTSS
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
SGARRSFFCRMKCNRPRKSFCTIHSTGYLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
RWFSFMNPWTKEVEYIVSTNTVVL
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLNLQEGEFLLQALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHVDDLENLAKCHEHLMQYGKGKSCYYRFLTKGQQW
KEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLDIEDDMKAQM
NCFYLKALDGFVMVLTDDGDMIYISDNVNKYMGLTQFELTGHSVFDFTHPCDHEEMREMLTHNTQRSFFL
KEKSRDAARCRRSKETEVFYELAHELPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSENESEAEADQ
KFTYCDDRITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSGQYRMLAKHGGYVWLETQ
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
LTSRGRTLNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPPLGRGAFLSRHSLDMKFTYCD
DRIAEVAGYSPDDLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSG
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVEMTGSSVFDYI
LEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHIDDLELLARCHQHLMQFGKGKSCCYRFLTKGQQWIWL
SRDAARSRRGKENFEFYELAKLLPLPAAITSQLDKASIIRLTISYLKMRDFANQGDPPWNLRMEGPPPNT
IVAALPGFLLVFTAEGKLLYLSESVSEHLGHSMVDLVAQGDSIYDIIDPADHLTVRQQLTLTDRLFRCRF
EKSKNAARTRREKENSEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
EKSKNAAKTRREKENGEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRAVFPEGLGDAWGQPSRAGP
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
KFVFVDQRATAILAYLPQELLGTSCYEYFHQDDIGHLAECHRQVLQTREKITTNCYKFKIKDGSFITLRS
```

8. & 9.


```
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep 'V/{2/}' sequence5.fasta
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -E "AA|DD" sequence5.fasta
AANFREGLNLQEGEFLLQALNGFVLVVTTDALVFYASSTIQDYLGFQQSDVIHQSVYELIHTEDRAEFQR
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHVDDLENLAKCHEHLMQYGKGKSCYYRFLTKGQQW
KEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLDIEDDMKAQM
NCFYLKALDGFVMVLTDDGDMIYISDNVNKYMGLTQFELTGHSVFDFTHPCDHEEMREMLTHNTQRSFFL
KEKSRDAARCRRSKETEVFYELAHELPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSENESEAEADQ
KFTYCDDRITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSGQYRMLAKHGGYVWLETQ
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
LTSRGRTLNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPPLGRGAFLSRHSLDMKFTYCD
DRIAEVAGYSPDDLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSG
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVEMTGSSVFDYI
LEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHIDDLELLARCHQHLMQFGKGKSCCYRFLTKGQQWIWL
SRDAARSRRGKENFEFYELAKLLPLPAAITSQLDKASIIRLTISYLKMRDFANQGDPPWNLRMEGPPPNT
IVAALPGFLLVFTAEGKLLYLSESVSEHLGHSMVDLVAQGDSIYDIIDPADHLTVRQQLTLTDRLFRCRF
EKSKNAARTRREKENSEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
EKSKNAAKTRREKENGEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRAVFPEGLGDAWGQPSRAGP
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
KFVFVDQRATAILAYLPQELLGTSCYEYFHQDDIGHLAECHRQVLQTREKITTNCYKFKIKDGSFITLRS
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"
SNPSKRHRDRLNTELDRLASLLPFPQDVINKLDKLSVLRLSVSYLRAKSFFDVALKSSPTERNGGQDNCR
QLHWQIPPENSPLMERCFICRLRCLLDNSSGFLAMNFQGKLKYLPPQLALFAIATPLQPPSILEIRTKNF
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
KNNRWTWVQSNARLLYKNGRPDYIIVTQRPLTDEEGTEHLR
VSRNKSEKKRRDQFNVLIKELGSMLPGNARKMDKSTVLQKSIDFLRKHKEITAQSDASEIRQDWKPTFLS
NEEFTQLMLEALDGFFLAIMTDGSIIYVSESVTSLLEHLPSDLVDQSIFNFIPEGEHSEVYKILSTEYLK
SKNQLEFCCHMLRGTIDPKEPSTYEYVKFIGNFKSLYEDRVCFVATVRLATPQFIKEMCTVEEPNEEFTS
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHVDDLENLAKCHEHLMQYGKGKSCYYRFLTKGQQW
```

```
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ grep -v "^>" sequence5.fasta | grep "P"
SNPSKRHRDRLNTELDRLASLLPFPQDVINKLDKLSVLRLSVSYLRAKSFFDVALKSSPTERNGGQDNCR
QLHWQIPPENSPLMERCFICRLRCLLDNSSGFLAMNFQGKLKYLPPQLALFAIATPLQPPSILEIRTKNF
IFRTKHKLDFTPIGCDAKGRIVLGYTEAELCTRGSGYQFIHAADMLYCAESHIRMIKTGESGMIVFRLLT
KNNRWTWVQSNARLLYKNGRPDYIIVTQRPLTDEEGTEHLR
VSRNKSEKKRRDQFNVLIKELGSMLPGNARKMDKSTVLQKSIDFLRKHKEITAQSDASEIRQDWKPTFLS
NEEFTQLMLEALDGFFLAIMTDGSIIYVSESVTSLLEHLPSDLVDQSIFNFIPEGEHSEVYKILSTEYLK
SKNQLEFCCHMLRGTIDPKEPSTYEYVKFIGNFKSLYEDRVCFVATVRLATPQFIKEMCTVEEPNEEFTS
RHSLEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHVDDLENLAKCHEHLMQYGKGKSCYYRFLTKGQQW
IWLQTHYYITYHQWNSRPEFIVCTHTVVSYAEVRAE
KEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISYLRVRKLLDAGDLDIEDDMKAQM
NCFYLKALDGFVMVLTDDGDMIYISDNVNKYMGLTQFELTGHSVFDFTHPCDHEEMREMLTHNTQRSFFL
RMKCTLTSRGRTMNIKSATWKVLHCTGHIHVYKPPMTCLVLICEPIPHPSNIEIPLDSKTFLSRHSLDMK
FSYCDERITELMGYEPEELLGRSIYEYYHALDSDHLTKTHHDMFTKGQVTTGQYRMLAKRGGYVWVETQA
TVIYNTKNSQPQCIVCVNYVVSGIIQHDL
KEKSRDAARCRRSKETEVFYELAHELPLPHSVSSHLDKASIMRLAISFLRTHKLLSSVCSENESEAEADQ
QMDNLYLKALEGFIAVVTQDGDMIFLSENISKFMGLTQVELTGHSIFDFTHPCDHEEIRENLSSTERDFF
MRMKCTVTNRGRTVNLKSATWKVLHCTGQVKVYEPLLSCLIIMCEPIQHPSHMDIPLDSKTFLSRHSMDM
KFTYCDDRITELIGYHPEELLGRSAYEFYHALDSENMTKSHQNLCTKGQVVSGQYRMLAKHGGYVWLETQ
GTVIYNPRNLQPQCIMCVNYVLSEIEKNDV
DAARSRRSQETEVLYQLAHTLPFARGVSAHLDKASIMRLTISYLRMHRLCAAGEWNQVGAGGEPLDACYL
KALEGFVMVLTAEGDMAYLSENVSKHLGLSQLELIGHSIFDFIHPCDQEELQDALTPPTERCFSLRMKST
LTSRGRTLNLKAATWKVLNCSGHMRAYEPPLQCLVLICEAIPHPGSLEPPLGRGAFLSRHSLDMKFTYCD
DRIAEVAGYSPDDLIGCSAYEYIHALDSDAVSKSIHTLLSKGQAVTGQYRFLARSGGYLWTQTQATVVSG
GRGPQSESIVCVHFLISQVEETGV
KEKSRNAARSRRGKENLEFFELAKLLPLPGAISSQLDKASIVRLSVTYLRLRRFAALGAPPWGLRAAGPP
AGLAPGRRGPAALVSEVFEQHLGGHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVEMTGSSVFDYI
HPGDHSEVLEQLGLVQERSFFVRMKSTLTKRGLHVKASGYKVIHVTGRLRALGLVALGHTLPPAPLAELP
LHGHMIVFRLSLGLTILACESRVSDHMDLGPSELVGRSCYQFVHGQDATRIRQSHVDLLDKGQVMTGYYR
WLQRAGGFVWLQSVATVAGSGKSPGEHHVLWVSHVLSQAEGGQT
NKSEKKRRDQFNVLIKELSSMLPGNTRKMDKTTVLEKVIGFLQKHNEVSAQTEICDIQQDWKPSFLSNEE
FTQLMLEALDGFIIAVTTDGSIIYVSDSITPLLGHLPSDVMDQNLLNFLPEQEHSEVYKILSSEYLKSDS
DLEFYCHLLRGSLNPKEFPTYEYIKFVGNFRSYLGKEVCFIATVRLATPQFLKEMCIVDEPLEEFTSRHS
```



```
NKSEKKRRDQFNVLIKELSSMLPGNTRKMDKTTVLEKVIGFLQKHNEVSAQTEICDIQQDWKPSFLSNEE
FTQLMLEALDGFIIAVTTDGSIIYVSDSITPLLGHLPSDVMDQNLLNFLPEQEHSEVYKILSSEYLKSDS
DLEFYCHLLRGSLNPKEFPTYEYIKFVGNFRSYLGKEVCFIATVRLATPQFLKEMCIVDEPLEEFTSRHS
LEWKFLFLDHRAPPIIGYLPFEVLGTSGYDYYHIDDLELLARCHQHLMQFGKGKSCCYRFLTKGQQWIWL
QTHYYITYHQWNSKPEFIVCTHSVVSYADVRVE
SRDAARSRRGKENFEFYELAKLLPLPAAITSQLDKASIIRLTISYLKMRDFANQGDPPWNLRMEGPPPNT
SVKVIGAQRRRSPSALAIEVFEAHLGSHILQSLDGFVFALNQEGKFLYISETVSIYLGLSQVELTGSSVF
DYVHPGDHVEMAEQLGMTLERSFFIRMKSTLTKRGVHIKSSGYKVIHTGRLRLRMGLVVVAHALPPPTI
NEVRIDCHMFVTRVNMDLNIIYCENRISDYMDLTPVDIVGKRCYHFIHAEDVEGIRHSHLDLLNKGQCVT
KYYRWMQKNGGYIWIQSSATIAINAKNANEKNIIWVNYLLSNPEYKDT
GASKARRDQINAEIRNLKELLPLAEADKVRLSYLHIMSLACIYTRKGVFFAGGTPLAGPTGLLSAQELED
IVAALPGFLLVFTAEGKLLYLSESVSEHLGHSMVDLVAQGDSIYDIIDPADHLTVRQQLTLTDRLFRCRF
NTSKSLRRQSAGNKLVLIRGRFHAHNPVFTAFCAPLEPRPRPGPGPGPGPASLFLAMFQSRHAKDLALLD
ISESVLIYLGFERSELLCKSWYGLLHPEDLAHASAQHYRLLAESGDIQAEMVVRLQAKTGGWAWIYCLLY
SEGPEGPITANNYPISDMEAWSLRQQL
EKSKNAARTRREKENSEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRVVFPEGLGEAWGHSSRTSP
LDNVGRELGSHLLQTLDGFIFVVAPDGKIMYISETASVHLGLSQVELTGNSIYEYIHPADHDEMTAVLTA
EIERSFFLRMKCVLAKRNAGLTCGGYKVIHCSGYLKIRNVGLVAVGHSLPPSAVTEIKLHSNMFMFRASL
DMKLIFLDSRVAELTGYEPQDLIEKTLYHHVHGCDTFHLRCAHHLLLVKGQVTTKYYRFLAKHGGWVWVQ
SYATIVHNSRSSRPHCIVSVNYVLTDTEYKGL
EKSKNAAKTRREKENGEFYELAKLLPLPSAITSQLDKASIIRLTTSYLKMRAVFPEGLGDAWGQPSRAGP
LDGVAKELGSHLLQTLDGFVFVASDGKIMYISETASVHLGLSQVELTGNSIYEYIHPSDHDEMTAVLTA
EIERSFFLRMKCVLAKRNAGLTCSGYKVIHCSGYLKIRIVGLVAVGQSLPPSAITEIKLYSNMFMFRASL
DLKLIFLDSRVTEVTGYEPQDLIEKTLYHHVHGCDVFHLRYAHHLLLVKGQVTTKYYRLLSKRGGWVWVQ
SYATVVHNSRSSRPHCIVSVNYVLTEIEYKEL
NHSEIERRRRNKMTAYITELSDMVPTCSALARKPDKLTILRMAVSHMKSLRGTGNTSTDGSYKPSFLTDQ
ELKHLILEAADGFLFIVSCETGRVVYVSDSVTPVLNQPQSEWFGSTLYDQVHPDDVDKLREQLSTSRMCM
GSRRSFICRMRCGSSEPHFVVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLKGQVLSVMFRFRSKNQEWLWMRTSS
FTFQNPYSDEIEYIICTNTNVK
EAHSQIEKRRRDKMNSFIDELASLVPTCNAMSRKLDKLTVLRMAVQHMKTLRGATNPYTEANYKPTFLSD
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
SGARRSFFCRMKCNRPRKSFCTIHSTGYLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
```

Part 3

9., 10., 11.

```
GSRRSFICRMRCGSSEPHFVVVHCTGYIKAKFCLVAIGRLQVTSSPNCTDMSNVCQPTEFISRHNIEGIF
TFVDHRCVATVGYQPQELLGKNIVEFCHPEDQQLLRDSFQQVVKLKGQVLSVMFRFRSKNQEWLWMRTSS
FTFQNPYSDEIEYIICTNTNVK
EAHSQIEKRRRDKMNSFIDELASLVPTCNAMSRKLDKLTVLRMAVQHMKTLRGATNPYTEANYKPTFLSD
DELKHLILRAADGFLFVVGCDRGKILFVSESVFKILNYSQNDLIGQSLFDYLHPKDIAKVKEQLSSSRLC
SGARRSFFCRMKCNRPRKSFCTIHSTGYLKSNLSCLVAIGRLHSHVVPQPVNGEIRVKSMEYVSRHAIDG
KFVFVDQRATAILAYLPQELLGTSCYEYFHQDDIGHLAECHRQVLQTREKITTNCYKFKIKDGSFITLRS
RWFSFMNPWTKEVEYIVSTNTVVL
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 $pattern protein.fasta
KPVKKKKIKREIKILENLRGGPNIITLADIVKDPVSRTPALVFEHVNNTDFKQLYQTLTDYDIRFYMYEI
WERFVHSENQHLVSPEALDFLDKLLRYDHQSRLTAREAMEHPYFYTVVKDQARMGSSSMPGGSTPVSSAN
dhaval@DESKTOP-50ECNBT:~/Lab_session_2$ 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
```
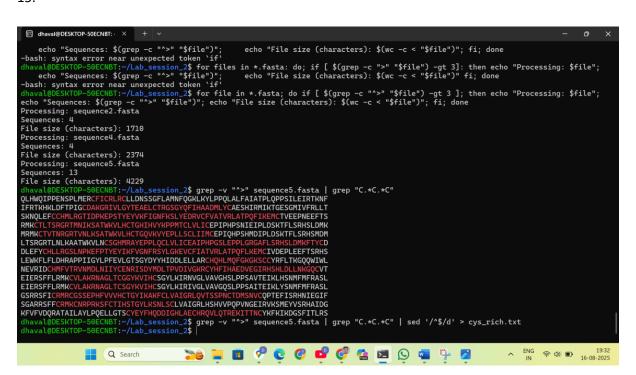
13.


```
 " file size (characters): $(wc - < "$file") done
> for file in *.fasta; do echo "Processing: $file" echo "sequences: $(grep -c ">" "$file")" echo " file size (characters): $(wc - < "$
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): $(wc - < "$file") done
> for file in *.fasta; do echo "Processing: $file" echo "sequences: $(grep -c ">" "$file")" echo " file size (characters): $(wc - < "$
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.

Part 5

15.

Note - sed command I got from Chat-gpt.