**Gene Expression Project Part-1 Bacterial Genes**

*Legend:* **blue box**=start codon; **red box**=stop codon

**1. DnaE (E. coli) 100 bp upstream plus coding region**

**a-Nucleotide FASTA file:**

>DnaE\_coding\_region\_plus\_100bp

CGCGACCGAA CACCATCGGC GCAGCTTTGG GCCTGTCAAA CGCGCACTGG GACTTGCGTC

CTGATTCTTG TGTCGAGATT AAGTAAACCG GAATCTGAAG **atg**TCTGAAC CACGTTTCGT

ACACCTGCGG GTGCACAGCG ACTACTCGAT GATCGATGGC CTGGCCAAAA CCGCACCGTT

GGTAAAAAAG GCGGCGGCGT TGGGTATGCC AGCACTGGCG ATCACCGATT TCACCAACCT

TTGTGGTCTG GTGAAGTTCT ACGGAGCGGG ACATGGCGCA GGGATTAAGC CTATCGTCGG

GGCAGATTTT AACGTCCAGT GCGACCTGCT GGGTGATGAG TTAACCCACC TGACGGTACT

GGCGGCGAAC AATACCGGCT ATCAGAATCT GACGTTGCTG ATCTCAAAAG CGTATCAGCG

CGGGTACGGT GCCGCCGGGC CGATCATCGA TCGCGACTGG CTTATCGAAT TAAACGAAGG

GTTGATCCTT CTTTCCGGCG GACGCATGGG CGACGTCGGA CGCAGTCTTT TGCGTGGTAA

CAGCGCGCTG GTAGATGAGT GTGTCGCGTT TTATGAAGAA CACTTCCCGG ATCGCTATTT

TCTCGAGCTG ATCCGCACCG GCAGGCCGGA TGAAGAAAGC TATCTGCACG CGGCGGTGGA

ACTGGCGGAA GCGCGCGGTT TGCCCGTCGT GGCGACCAAC GACGTGCGCT TTATCGACAG

CAGCGACTTT GACGCACACG AAATCCGCGT CGCGATCCAC GACGGCTTTA CCCTCGACGA

TCCTAAACGC CCGCGTAACT ATTCGCCGCA GCAATATATG CGTAGCGAAG AGGAGATGTG

TGAGCTGTTT GCCGACATCC CCGAAGCCCT TGCCAACACC GTTGAGATCG CCAAACGCTG

TAACGTAACC GTGCGTCTTG GTGAATACTT CCTGCCGCAG TTCCCGACCG GGGACATGAG

CACCGAAGAT TATCTGGTCA AGCGTGCAAA AGAGGGCCTG GAAGAGCGTC TGGCCTTTTT

ATTCCCTGAT GAGGAAGAAC GTCTTAAGCG CCGCCCGGAA TATGACGAAC GTCTGGAGAC

TGAACTTCAG GTTATCAACC AGATGGGCTT CCCGGGCTAC TTCCTCATCG TTATGGAATT

TATCCAGTGG TCGAAAGATA ACGGCGTACC GGTAGGGCCA GGCCGTGGCT CCGGTGCGGG

TTCACTGGTG GCCTACGCGC TGAAAATCAC CGACCTCGAT CCGCTGGAAT TTGACCTGCT

GTTCGAACGT TTCCTTAACC CGGAACGTGT CTCCATGCCT GACTTCGACG TTGACTTCTG

TATGGAGAAA CGCGATCAGG TTATCGAGCA CGTAGCGGAC ATGTACGGTC GTGATGCGGT

ATCGCAGATC ATCACCTTCG GTACAATGGC GGCGAAAGCG GTGATCCGCG ACGTAGGCCG

CGTGCTGGGG CATCCGTACG GCTTTGTCGA TCGTATCTCG AAACTGATCC CGCCCGATCC

GGGGATGACG CTGGCGAAAG CGTTTGAAGC CGAGCCGCAG CTGCCGGAAA TCTACGAAGC

GGATGAAGAA GTTAAGGCGC TGATCGACAT GGCGCGCAAA CTGGAAGGGG TCACCCGTAA

CGCCGGTAAG CACGCCGGTG GGGTGGTTAT CGCGCCGACC AAAATTACCG ATTTTGCGCC

GCTTTACTGC GATGAAGAGG GCAAACATCC GGTCACCCAG TTTGATAAAA GCGACGTTGA

ATACGCCGGA CTGGTGAAGT TCGACTTCCT TGGTTTGCGT ACGCTCACCA TCATCAACTG

GGCGCTGGAG ATGATCAACA AGCGGCGGGC GAAGAATGGC GAGCCGCCGC TGGATATCGC

TGCGATCCCG CTGGATGATA AGAAAAGCTT CGACATGCTG CAACGCTCGG AAACCACGGC

GGTATTCCAG CTTGAATCGC GCGGCATGAA GGACCTGATC AAGCGTCTAC AACCTGACTG

CTTCGAAGAT ATGATCGCCC TAGTGGCACT GTTCCGCCCC GGTCCGTTGC AATCAGGGAT

GGTGGATAAC TTTATCGACC GTAAACATGG TCGTGAAGAG ATCTCCTATC CGGACGTACA

GTGGCAGCAT GAAAGCCTGA AACCGGTACT GGAGCCAACC TACGGCATTA TCCTGTATCA

GGAACAGGTC ATGCAGATTG CGCAGGTGCT TTCTGGTTAT ACCCTCGGTG GCGCGGATAT

GCTGCGTCGT GCGATGGGTA AGAAAAAGCC GGAAGAGATG GCTAAGCAAC GTTCTGTATT

TGCTGAAGGT GCAGAAAAGA ACGGAATCAA CGCTGAACTG GCGATGAAAA TCTTCGACCT

GGTGGAGAAA TTCGCTGGTT ACGGATTTAA CAAATCGCAC TCTGCGGCCT ATGCTTTGGT

GTCATATCAA ACGTTATGGC TGAAAGCGCA CTATCCTGCG GAGTTTATGG CGGCGGTAAT

GACCGCCGAT ATGGACAACA CCGAGAAGGT GGTGGGTCTG GTGGATGAGT GCTGGCGGAT

GGGGCTGAAA ATCCTGCCAC CAGATATAAA CTCCGGTCTT TACCATTTCC ACGTCAACGA

CGACGGCGAA ATCGTGTATG GTATTGGCGC GATCAAAGGG GTCGGTGAAG GTCCGATTGA

GGCCATCATC GAAGCCCGTA ATAAAGGCGG CTACTTCCGC GAACTGTTTG ATCTCTGCGC

CCGTACCGAC ACCAAAAAGT TGAACCGTCG CGTGCTGGAA AAACTGATCA TGTCCGGGGC

GTTTGACCGT CTTGGGCCAC ATCGCGCAGC GCTGATGAAC TCGCTGGGCG ATGCGTTAAA

AGCGGCAGAT CAACACGCGA AAGCGGAAGC TATCGGTCAG GCCGATATGT TCGGCGTGCT

GGCCGAAGAG CCGGAACAAA TTGAACAATC CTACGCCAGC TGCCAACCGT GGCCGGAGCA

GGTGGTATTA GATGGGGAAC GTGAAACGTT AGGCCTGTAC CTGACCGGAC ACCCTATCAA

CCAGTATTTA AAAGAGATTG AGCGTTATGT CGGAGGCGTA AGGCTGAAAG ACATGCACCC

GACAGAACGT GGTAAAGTCA TCACGGCTGC GGGGCTCGTT GTTGCCGCGC GGGTTATGGT

CACCAAGCGC GGCAATCGTA TCGGTATCTG CACGCTGGAT GACCGTTCCG GGCGGCTGGA

AGTGATGTTG TTTACTGACG CCCTGGATAA ATACCAGCAA TTGCTGGAAA AAGACCGCAT

ACTTATCGTC AGCGGACAGG TCAGCTTTGA TGACTTCAGC GGTGGGCTTA AAATGACCGC

TCGCGAAGTG ATGGATATTG ACGAAGCCCG GGAAAAATAT GCTCGCGGGC TTGCTATCTC

GCTGACGGAC AGGCAAATTG ATGACCAGCT TTTAAACCGA CTCCGTCAGT CTCTGGAACC

CCACCGCTCT GGGACAATTC CAGTACATCT CTACTATCAG AGGGCGGATG CACGCGCGCG

GTTGCGTTTT GGCGCGACGT GGCGTGTCTC TCCGAGCGAT CGTTTATTAA ACGATCTCCG

TGGCCTCATT GGTTCGGAGC AGGTGGAACT GGAGTTTGAC **taa**

**b-NCBI ORF Finder screen print: DnaE**

**Graphical user interface, application, table

Description automatically generated**

**c-Protein FASTA file: DnaE**

>DnaE\_DNA\_polymerase\_III\_subunit\_Alpha\_protein

MSEPRFVHLRVHSDYSMIDGLAKTAPLVKKAAALGMPALAITDFTNLCGLVKFYGAGHGAGIKPIVGADF

NVQCDLLGDELTHLTVLAANNTGYQNLTLLISKAYQRGYGAAGPIIDRDWLIELNEGLILLSGGRMGDVG

RSLLRGNSALVDECVAFYEEHFPDRYFLELIRTGRPDEESYLHAAVELAEARGLPVVATNDVRFIDSSDF

DAHEIRVAIHDGFTLDDPKRPRNYSPQQYMRSEEEMCELFADIPEALANTVEIAKRCNVTVRLGEYFLPQ

FPTGDMSTEDYLVKRAKEGLEERLAFLFPDEEERLKRRPEYDERLETELQVINQMGFPGYFLIVMEFIQW

SKDNGVPVGPGRGSGAGSLVAYALKITDLDPLEFDLLFERFLNPERVSMPDFDVDFCMEKRDQVIEHVAD

MYGRDAVSQIITFGTMAAKAVIRDVGRVLGHPYGFVDRISKLIPPDPGMTLAKAFEAEPQLPEIYEADEE

VKALIDMARKLEGVTRNAGKHAGGVVIAPTKITDFAPLYCDEEGKHPVTQFDKSDVEYAGLVKFDFLGLR

TLTIINWALEMINKRRAKNGEPPLDIAAIPLDDKKSFDMLQRSETTAVFQLESRGMKDLIKRLQPDCFED

MIALVALFRPGPLQSGMVDNFIDRKHGREEISYPDVQWQHESLKPVLEPTYGIILYQEQVMQIAQVLSGY

TLGGADMLRRAMGKKKPEEMAKQRSVFAEGAEKNGINAELAMKIFDLVEKFAGYGFNKSHSAAYALVSYQ

TLWLKAHYPAEFMAAVMTADMDNTEKVVGLVDECWRMGLKILPPDINSGLYHFHVNDDGEIVYGIGAIKG

VGEGPIEAIIEARNKGGYFRELFDLCARTDTKKLNRRVLEKLIMSGAFDRLGPHRAALMNSLGDALKAAD

QHAKAEAIGQADMFGVLAEEPEQIEQSYASCQPWPEQVVLDGERETLGLYLTGHPINQYLKEIERYVGGV

RLKDMHPTERGKVITAAGLVVAARVMVTKRGNRIGICTLDDRSGRLEVMLFTDALDKYQQLLEKDRILIV

SGQVSFDDFSGGLKMTAREVMDIDEAREKYARGLAISLTDRQIDDQLLNRLRQSLEPHRSGTIPVHLYYQ

RADARARLRFGATWRVSPSDRLLNDLRGLIGSEQVELEFD

**2. your second gene from the table: araC**

*Legend:* **blue box**=start codon; **red box**=stop codon

**a-Nucleotide FASTA file:**

>araC\_coding\_region\_plus\_100bp

TTGGGTTAGC GAGAAGAGCC AGTAAAAGAC GCAGTGACGG CAATGTCTGA TGCAATATGG

ACAATTGGTT TCTTCTCTGA ATGGTGGGAG TATGAAAAGT **atg**GCTGAAG CGCAAAATGA

TCCCCTGCTG CCGGGATACT CGTTTAACGC CCATCTGGTG GCGGGTTTAA CGCCGATTGA

GGCCAACGGT TATCTCGATT TTTTTATCGA CCGACCGCTG GGAATGAAAG GTTATATTCT

CAATCTCACC ATTCGCGGTC AGGGGGTGGT GAAAAATCAG GGACGAGAAT TTGTCTGCCG

ACCGGGTGAT ATTTTGCTGT TCCCGCCAGG AGAGATTCAT CACTACGGTC GTCATCCGGA

GGCTCGCGAA TGGTATCACC AGTGGGTTTA CTTTCGTCCG CGCGCCTACT GGCATGAATG

GCTTAACTGG CCGTCAATAT TTGCCAATAC GGGTTTCTTT CGCCCGGATG AAGCGCACCA

GCCGCATTTC AGCGACCTGT TTGGGCAAAT CATTAACGCC GGGCAAGGGG AAGGGCGCTA

TTCGGAGCTG CTGGCGATAA ATCTGCTTGA GCAATTGTTA CTGCGGCGCA TGGAAGCGAT

TAACGAGTCG CTCCATCCAC CGATGGATAA TCGGGTACGC GAGGCTTGTC AGTACATCAG

CGATCACCTG GCAGACAGCA ATTTTGATAT CGCCAGCGTC GCACAGCATG TTTGCTTGTC

GCCGTCGCGT CTGTCACATC TTTTCCGCCA GCAGTTAGGG ATTAGCGTCT TAAGCTGGCG

CGAGGACCAA CGCATTAGTC AGGCGAAGCT GCTTTTGAGC ACTACCCGGA TGCCTATCGC

CACCGTCGGT CGCAATGTTG GTTTTGACGA TCAACTCTAT TTCTCGCGAG TATTTAAAAA

ATGCACCGGG GCCAGCCCGA GCGAGTTTCG TGCCGGTTGT GAAGAAAAAG TGAATGATGT

AGCCGTCAAG TTGTCA**taa**

**b-NCBI ORF Finder screen print: araCGraphical user interface, application

Description automatically generated**

**c-Protein FASTA file: araC**

>araC\_DNA-binding\_transcriptional\_dual\_regulator

MAEAQNDPLLPGYSFNAHLVAGLTPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQGREFVCRPGD

ILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSIFANTGFFRPDEAHQPHFSDLFGQIINA

GQGEGRYSELLAINLLEQLLLRRMEAINESLHPPMDNRVREACQYISDHLADSNFDIASVAQHVCLSPSR

LSHLFRQQLGISVLSWREDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAGC

EEKVNDVAVKLS

**3. your third gene from the table: rpoE**

**a-Nucleotide FASTA file:**

*Legend:* **blue box**=start codon; **red box**=stop codon

>rpoE\_coding\_region\_plus\_100bp

AAAAACGAGA CACTCTAACC CTTTGCTTGC TCAAATTGCA GCTAATGGAG TGGCGTTTCG

ATAGCGCGTG GAAATTTGGT TTGGGGAGAC TTTACCTCGG **atg**AGCGAGC AGTTAACGGA

CCAGGTCCTG GTTGAACGGG TCCAGAAGGG AGATCAGAAA GCCTTTAACT TACTGGTAGT

GCGCTATCAG CATAAAGTGG CGAGTCTGGT TTCCCGCTAT GTGCCGTCGG GTGATGTTCC

CGATGTGGTA CAAGAAGCTT TTATTAAAGC CTATCGTGCG CTGGATTCGT TCCGGGGAGA

TAGCGCTTTT TATACATGGC TGTATCGGAT TGCTGTAAAT ACAGCGAAAA ATTACCTGGT

TGCTCAGGGG CGTCGTCCAC CTTCCAGTGA TGTGGATGCC ATTGAAGCTG AAAACTTCGA

AAGTGGCGGC GCGTTGAAAG AAATTTCGAA CCCTGAGAAC TTAATGTTGT CAGAAGAACT

GAGACAGATA GTTTTCCGAA CTATTGAGTC CCTCCCGGAA GATTTACGCA TGGCAATAAC

CTTGCGGGAG CTGGATGGCC TGAGCTATGA AGAGATAGCC GCTATCATGG ATTGTCCGGT

AGGTACGGTG CGTTCACGTA TCTTCCGAGC GAGGGAAGCT ATTGATAACA AAGTTCAACC

GCTTATCAGG CGT**tga**

**b-NCBI ORF Finder screen print: rpoEGraphical user interface, application, Word

Description automatically generated**

**c-Protein FASTA file: rpoE**

>rpoE\_RNA\_polymerase\_sigma\_factor

MSEQLTDQVLVERVQKGDQKAFNLLVVRYQHKVASLVSRYVPSGDVPDVVQEAFIKAYRALDSFRGDSAF

YTWLYRIAVNTAKNYLVAQGRRPPSSDVDAIEAENFESGGALKEISNPENLMLSEELRQIVFRTIESLPE

DLRMAITLRELDGLSYEEIAAIMDCPVGTVRSRIFRAREAIDNKVQPLIRR

**4. your fourth gene from the table: holA**

*Legend:* **blue box**=start codon; **red box**=stop codon

**a-Nucleotide FASTA file:**

>holA\_coding\_region\_plus\_100bp

GCATCCGTGC TGCGGATATT CGTTCCGACG AAGAACAGAC GTCGACCACA ACGGATACTC

CGGCAACGCC TGCACGCGTC TCCACCACGC TGGGTAACTG **atg**ATTCGGT TGTACCCGGA

ACAACTCCGC GCGCAGCTCA ATGAAGGGCT GCGCGCGGCG TATCTTTTAC TTGGTAACGA

TCCTCTGTTA TTGCAGGAAA GCCAGGACGC TGTTCGTCAG GTAGCTGCGG CACAAGGATT

CGAAGAACAC CACACTTTTT CCATTGATCC CAACACTGAC TGGAATGCGA TCTTTTCGTT

ATGCCAGGCT ATGAGTCTGT TTGCCAGTCG ACAAACGCTA TTGCTGTTGT TACCAGAAAA

CGGACCGAAT GCGGCGATCA ATGAGCAACT TCTCACACTC ACCGGACTTC TGCATGACGA

CCTGCTGTTG ATCGTCCGCG GTAATAAATT AAGCAAAGCG CAAGAAAATG CCGCCTGGTT

TACTGCGCTT GCGAATCGCA GCGTGCAGGT GACCTGTCAG ACACCGGAGC AGGCTCAGCT

TCCCCGCTGG GTTGCTGCGC GCGCAAAACA GCTCAACTTA GAACTGGATG ACGCGGCAAA

TCAGGTGCTC TGCTACTGTT ATGAAGGTAA CCTGCTGGCG CTGGCTCAGG CACTGGAGCG

TTTATCGCTG CTCTGGCCAG ACGGCAAATT GACATTACCG CGCGTTGAAC AGGCGGTGAA

TGATGCCGCG CATTTCACCC CTTTTCATTG GGTTGATGCT TTGTTGATGG GAAAAAGTAA

GCGCGCATTG CATATTCTTC AGCAACTGCG TCTGGAAGGC AGCGAACCGG TTATTTTGTT

GCGCACATTA CAACGTGAAC TGTTGTTACT GGTTAACCTG AAACGCCAGT CTGCCCATAC

GCCACTGCGT GCGTTGTTTG ATAAGCATCG GGTATGGCAG AACCGCCGGG GCATGATGGG

CGAGGCGTTA AATCGCTTAA GTCAGACGCA GTTACGTCAG GCCGTGCAAC TCCTGACACG

AACGGAACTC ACCCTCAAAC AAGATTACGG TCAGTCAGTG TGGGCAGAGC TGGAAGGGTT

ATCTCTTCTG TTGTGCCATA AACCCCTGGC GGACGTATTT ATCGACGGT**t ga**

**b-NCBI ORF Finder screen print: holAGraphical user interface, application

Description automatically generated**

**c-Protein FASTA file: holA**

>holA\_DNA\_polymerase\_III\_subunit\_delta

MIRLYPEQLRAQLNEGLRAAYLLLGNDPLLLQESQDAVRQVAAAQGFEEHHTFSIDPNTDWNAIFSLCQA

MSLFASRQTLLLLLPENGPNAAINEQLLTLTGLLHDDLLLIVRGNKLSKAQENAAWFTALANRSVQVTCQ

TPEQAQLPRWVAARAKQLNLELDDAANQVLCYCYEGNLLALAQALERLSLLWPDGKLTLPRVEQAVNDAA

HFTPFHWVDALLMGKSKRALHILQQLRLEGSEPVILLRTLQRELLLLVNLKRQSAHTPLRALFDKHRVWQ

NRRGMMGEALNRLSQTQLRQAVQLLTRTELTLKQDYGQSVWAELEGLSLLLCHKPLADVFIDG

tomato sandwich: <https://nurturedhomes.com/tomato-sandwich>



porkchop sandwich: <https://raymonds.recipes/pork-chop-sandwiches/>

