Appendix 1 Sequence data

ACCCAGGCAACATCAGGGAGAGCTTTGGGCTGGGAGGAATCTAGGGTGTGGGGGCAGCTGGCCTTCCTCATAGGACAGACCCTCCCACGCGTTCAGGGAGGTGGAGCACAGGTGGCAGTAGTATCTGCATCCCCTGACTCTCTCTCCACAGTTCC
TGGGTAAATGCCTGCTGGTGACCTAGGCCTGCCACACCCTTCCCAGTTTACCCATGTGGTGCCTCCATGGACAAATTATTTGCTTTTGTGAGTGCTGTGTTGACCTAAAAACACCATTAAGCTAGAGCATTGGTGGTCATGCCCCCTGCCTGCTGG
GCCTCCCACCAGGCCCTCCTCCCCTCCCTGCCCCAGCACTTCCTGATCTTTGAATGAAGTCCGAGTAGGCAGCAGCCTGTGTGTGCCTGGGTTCTCTCTGTCCCGGAATGTGCCAACAGTGGAGGTGTTTACCTGTCTCAGACCAAGGACCTCTCT
GCAGCTGCATGGGGCTGGGGAGGGAGAACTGCAGGGAGTATGGGAGGGGAAGCTGAGGTGGGCCTGCTCAAGAGAAGGTGCTGAACCATCCCCTGTCCTGAGAGGTGCCAGGCCTGCAGGCAGTGGCTCAGAAGCTGGGGAGGAGAGAGGCAT
CCAGGGTTCTACTCAGGGAGTCCCAGCATCGCCACCCTCCTTTGAAATCTCCCTGGTTGAACCCAGTTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAATGCAAGTGCAGGTGCCAGAACATTTC
TCTCATTCTCACCCCTTCCTGCCAGAGGGTAGGTGGCTGGAGTGAGGGTGCTGGCCCTACTCACACTTCCTGTGTCATGGTGACCCTCTGAGAGCAGCCCAGTCAGTGGGGAAGGAGGAAGGGGCTGGGATGCTCACAGCCGGCAGCCCACACCT
GGGGAGACTCTTCAGCAGAGCACCTTGCGGCCTTACTCCTGCACGTCTCCTGCAGTTTGTAAGGTGCATTCAGAACTCACTGTGTGCCCAGCCCTGAGCTCCCAGCTAATTGCCCCACCCAGGGCCTCTGGGACCTCCTGGTGCTTCTGCTTCCTG
TGCTGCCAGCAACTTCTGGAAACGTCCCTGTCCCCGGTGCTGAAGTCCTGGAATCCATGCTGGGAAGTTGCACAGCCCATCTGGCTCCAGCCAGCCTAGGAACACGAGCAGCACTTCCAGCCCAGCCCCTGCCCCACAGCAAGCCTCCCCCTCCA
CACTCACAGTACTGAATTGAGCTTTGGGTAGGGTGGAGAGGACCCTGTCACCGCTTTTCTTCTGGACATGGACCTCTCTGAATTGTTGGGGAGTTCCCTCCCCCTCTCCACCACCCACTCTTCCTGTGCCTCACAGCCCAGAGCATTGTTATTTCAA
CAGAAACACTTTAAAAAATAAACTAAAATCCGACAGGCACGGTGGCTCACACCTGTAATCCCAGTACTTTGGGAGGCTGAGGCGAGAGGATCACCTGAGGTCGGGAGTTTGAGACCAGACTGACCAATATGGAGAAACCCCAGTTATACTAAAAAT
ACAAAATTAGCTGGGTGTGGTGGCGCATGCCTGTAATCCTAGCTACTAGGAAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGTGGAGGTTGAGGTGAGCTGAGATCACGCCATTGCACTCCAGCCTGGGCAACAAGAGCAAAACTCCGTCT
CAAAAAATAAATAAATAAATAAATAAATAAACTAAAATCTATCCATGCTTTCACACACACACATACACACACACACACACACACCCTTTTTTGTGTTACTTAAAGTAGGAGAGTGTCTCTCTTTCCTGTCTCCTCACACCCACCCCCAGAAGAGA
CCAAAATGAAGGGTTTGGAACTCAGCCCATGGGCCCCATCCCATGCTGAGGGAACACAGCTACATCTACAACTACTGCCACAGGCTCTCTTTTTGGACAAAAATACCATCATACTGTAGATACCTGTGTACAACTTCCTATTCTCAGTGAAGTGTCT
CCCCTGCATCCCTTTCAGCCAGTTCATTCAGCTCTGCGCCATTCCACAGTCTCACTGATTATTACTATGTTTCCATCATGATCCCCCCAAAAAATCATGACTTTATTTTTTTATTTTTATTATTATTATTATTATTTTTTTTTTTTTTGTGACGGAGTCT
CGCTCTGTCACCCAGGCTGGAGTGCAGTGGCACAATCTCGGCTCACTGCAAGCTCCACCTCGCAGGTTCACGCCATTCTCCTCCCTCTCAGCCTCCCGAGTAGCTGAGTAGCTGGGACTACAGGCGCCCCCCCACTACGCCTGGCTAATTTTTTCTATTT
TTAATAGAGACAGAGTTTCACTGCATTAGCGAGGATGGTCTCGATCTCCTGACCTCGCATCTGCCCGCCTCAGCCTCCCAATGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTTATGTATTTATTTTTTTGAGACAGAGTCTCGCTGTG
TCGTCAGGCTAGAGTGCTGTGGCACGATCTCGGCTCACTGCAACCTCCAACTCCCTGGTTCAAAGGATTCTCCAGCCTCCACCTCCCGAGTAGCTGGGATTACAGGCGTGCACCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGT
TTCTCCATGTTGGTCAGCCTGGTCTGAACTCCCGACCTCAGCTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGGATTACAGGCGTGAGCCACCGAGCCTGGCCAAACCATCACTTTTCATGAGCAGGGATGCACCCACTGGCACTCCTGCACCT
CCCACCCTCCCCCTCGCCAAGTCCACCCCTTCCTTCCTCACCCCACATCCCCTCACCTACATTCTGCAACCACAGGGGCCTTCTCTCCCCTGTCCTTTCCCTACCCAGAGCCAAGTTTGTTTATCTGTTTACAACCAGTATTTACCTAGCAAGTCTTCC
ATCAGATAGCATTTGGAGAGCTGGGGGTGTCACAGTGAACCACGACCTCTAGGCCAGTGGGAGAGTCAGTCACACAAACTGTGAGTCCATGACTTGGGGCTTAGCCAGCACCCACCACCCCACGCGCCACCCCACAACCCCGGGTAGAGGAGTCT
GAATCTGGAGCCGCCCCCAGCCCAGCCCCGTGCTTTTTGCGTCCTGGTGTTTGTTCCTTCCCGGTGCCTGTCACTCAAGCACACTAGTGACTATCGCCAGAGGGAAAGGGAGCTGCAGGAAGCGAGGCTGGAGAGCAGGAGGGGCTCTGCGCAGA
AATTCTTTTGAGTTCCTATGGGCCAGGGCGTCCGGGTGCGCGCATTCCTCTCCGCCCCAGGATTGGGCGAAGCCTCCCGGCTCGCACTCGCTCGCCCGTGTGTTCCCCGATCCCGCTGGAGTCGATGCGCGTCCAGCGCGTGCCAGGCCGGGGCG
GGGGTGCGGGCTGACTTTCTCCCTCGCTAGAGACGCTCCGGCGCCCGAAAGGAAAGGGTGGCGCTGCGCTCCGGGGTGCACGAGCCGACAGCGCCCGACCCCAACGGGCCGGCCCCGCCAGCGCCGCTACCGCCCTGCTCCCGGGCGAGCGGG
ATGGGCGGGAGTGGAGTGGCGGGTGGAGGGTGGAGACGTCCTGGCCCCCGCCCCGCGTGCACCCCCAGGGGAGGCCGAGCCCGCCGCCCGGCCCCGCGCAGGCCCCGCCCGGGACTCCCCTGCGGTCCAGGCCGCGCCCCGGGCTCCGCGCCA
GCCAATGAGCGCCGCCCGGCCGGGCGTGCCCCCGCGCCCCAAGCATAAACCCTGGCGCGCTCGCGGCCCGGCACTCTTCTGGTCCCCACAGACTCAGAGAGAACCCACCATGGTGCTGTCTCCTGCCGACAAGACCAACGTCAAGGCCGCCTGGG
GTAAGGTCGGCGCGCACGCTGGCGAGTATGGTGCGGAGGCCCTGGAGAGGTGAGGCTCCCTCCCCTGCTCCGACCCGGGCTCCTCGCCCGCCCGGACCCACAGGCCACCCTCAACCGTCCTGGCCCCGGACCCAAACCCCACCCCTCACTCTGCT
TCTCCCCGCAGGATGTTCCTGTCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCGACCTGAGCCACGGCTCTGCCCAGGTTAAGGGCCACGGCAAGAAGGTGGCCGACGCGCTGACCAACGCCGTGGCGCACGTGGACGACATGCCCAACGC
GCTGTCCGCCCTGAGCGACCTGCACGCGCACAAGCTTCGGGTGGACCCGGTCAACTTCAAGGTGAGCGGCGGGCCGGGAGCGATCTGGGTCGAGGGGCGAGATGGCGCCTTCCTCTCAGGGCAGAGGATCACGCGGGTTGCGGGAGGTGTAGC
GCAGGCGGCGGCTGCGGGCCTGGGCCGCACTGACCCTCTTCTCTGCACAGCTCCTAAGCCACTGCCTGCTGGTGACCCTGGCCGCCCACCTCCCCGCCGAGTTCACCCCTGCGGTGCACGCCTCCCTGGACAAGTTCCTGGCTTCTGTGAGCACC
GTGCTGACCTCCAAATACCGTTAAGCTGGAGCCTCGGTAGCCGTTCCTCCTGCCCGCTGGGCCTCCCAACGGGCCCTCCTCCCCTCCTTGCACCGGCCCTTCCTGGTCTTTGAATAAAGTCTGAGTGGGCGGCAGCCTGTGTGTGCCTGGGTTCTC
TCTGTCCCGGAATGTGCCAACAATGGAGGTGTTTACCTGTCTCAGACCAAGGACCTCTCTGCAGCTGCATGGGGCTGGGGAGGGAGAACTGCAGGGAGTATGGGAGGGGAAGCTGAGGTGGGCCTGCTCAAGAGAAGGTGCTGAACCATCCCCT
GTCCTGAGAGGTGCCAGGCCTGCAGGCAGTGGCTCAGAAGCTGGGGAGGAGAGAGGCATCCAGGGTTCTACTCAGGGAGTCCCAGCATCGCCACCCTCCTTTGAAATCTCCCTGGTTGAACCCAGTTAACATACGCTCTCCATCAAAACAAAACGA
AACAAAACAAACTAGCAAAATAGGCTGTCCCCAATGCAAGTGCAGGTGCCAGAACATTTCTCTCATTCTCACCCCTTCCTGCCAGAGGGTAGGTGGCTGGAGTGAGGGTGCTGGCCCTACTCACACTTCCTGTGTCATGGTGACCCTCTGAGAGCA
GCCCAGTCAGTGGGGAAGGAGGAAGGGGCTGGGATGCTCACAGCCGGCAGCCCACACCTGGGGAGACTCTTCAGCAGAGCACCTTGCGGCCTTACTCCTGCACGTCTCCTGCAGTTTGTAAGGTGCATTCAGAACTCACTGTGTGCCCAGCCCTG
AGCTCCCAGCTAATTGCCCCACCCAGGGCCTCTGGGACCTCCTGGTGCTTCTGCTTCCTGTGCTGCCAGCAACTTCTGGAAACGTCCCTGTCCCCGGTGCTGAAGTCCTGGAATCCATGCTGGGAAGTTGCACAGCCCATCTGGCTCTCAGCCAGC
CTAGGAACACGAGCAGCACTTCCAGCCCAGCCCCTGCCCCACAGCAAGCCTCCCCCTCCACACTCACAGTACTGAATTGAGCTTTGGGTAGGGTGGAGAGGACCCTGTCACCGCTTTTCTTCTGGACATGGACCTCTCTGAATTGTTGGGGAGTTC
CCTCCCCCTCTCCACCACCCACTCTTCCTGTGCCTCACAGCCCAGAGCATTGTTATTTCAACAGAAACACTTTAAAAAATAAACTAAAATCCGACAGGCACGGTGGCTCACACCTGTAATCCCAGTACTTTGGGAGGCTGAGGCGAGAGGATCACCT
GAGGTCGGGAGTTTGAGACCAGCCTGACCAATATGGAGAAACCCCAGTTATACTAAAAATACAAAATTAGCTGGGTGTGGTGGCGCATGCCTGTAATCCTAGCTACTAGGAAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGTGGAGGTTG
AGGTGAGCTGAGATCACGCCATTGCACTCCAGCCTGGGCAACAAGAGCAAAACTCCGTCTCAAAAAATAAATAAATAAATAAATAAACTAAAATCTATCCATGCTTTCACACACACACACACACACACACACACACACCCTTTTTTGTGTTA
CTTAAAGTAGGAGAGTGTCTCTCTTTCCTGTCTCCTCACACCCACCCCCAGAAGAGACCAAAATGAAGGGTTTGGAACTCAGCCCATGGGCCCCATCCCATGCTGAGGGAACACAGCTACATCTACAACTACTGCCACAGGCTCTCTTTTTGGACAA
AAATACCATCATACTGTAGATACCTGTGTACAACTTCCTATTCTCAGTGAAGTGTCTCCCCTGCATCCCTTTCAGCCAGTTCATTCAGCTCTGCGCCATTCCACAGTCTCACTGATTATTACTATGTTTCCATCATGATCCCCCCAAAAAATCATGACT
TTATTTTTTTATTTTTATTATTATTATTATTATTTTTTTTTTTTTTTTTGTGACGGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCACAATCTCGGCTCACTGCAAGCTCCACCTCGCAGGTTCACGCCATTCTCCTCCCTCAGCCTCCCGAGT
AGCTGAGTAGCTGGGACTACAGGCGCCCCCCACTACGCCTGGCTAATTTTTTCTATTTTTAATAGAGACAGAGTTTCACTGCATTAGCGAGGATGGTCTCGATCTCCTGACCTCGCATCTGCCCGCCTCAGCCTCCCAATGTGCTGGGATTACAGG
CGTGAGCCACCGCGCCCGGCCTTATGTATTTATTTTTTTGAGACAGAGTCTCGCTGTGTCGTCAGGCTAGAGTGCTGTGGCACGATCTCGGCTCACTGCAACCTCCAACTCCCTGGTTCAAAGGATTCTCCAGCCTCCACCTCCCGAGTAGCTGGG
ATTACAGGCGTGCACCACCACACCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCTCCATGTTGGTCAGCCTGGTCTCGAACTCCCGACCTCAGCTGATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGA
GCCTGGCCAAACCATCACTTTTCATGAGCAGGGATGCACCCACTGGCACTCCTGCACCTCCCACCCTCCCCCTCGCCAAGTCCACCCCTTCCTTCCTCACCCCACATCCCCTCACCTACATTCTGCAACCACAGGGGCCTTCTCTCCCCTGTCCTTTC
CCTACCCAGAGCCAAGTTTGTTTATCTGTTTACAACCAGTATTTACCTAGCAAGTCTTCCATCAGATAGCATTTGGAGAGCTGGGGGTGTCACAGTGAACCACGACCTCTAGGCCAGTGGGAGAGTCAGTCACACAAACTGTGAGTCCATGACTTG
GGGCTTAGCCAGCACCCACCACCCCACGCGCCACCCCACAACCCCGGGTAGAGGAGTCTGAATCTGGAGCCGCCCCCAGCCCAGCCCCGTGCTTTTTGCGTCCTGGTGTTTGTTCCTTCCCGGTGCCTGTCACTCAAGCACACTAGTGACTATCGC
CAGAGGGAAAGGGAGCTGCAGGAAGCGAGGCTGGAGAGCAGGAGGGGCTCTGCGCAGAAATTCTTTTGAGTTCCTATGGGCCAGGGCGTCCGGGTGCGCGCATTCCTCTCCGCCCCAGGATTGGGCGAAGCCTCCCGGCTCGCACTCGCTCGCC
CGTGTGTTCCCCGATCCCGCTGGAGTCGATGCGCGTCCAGCGCGTGCCAGGCCGGGGCGGGGGTGCGGGCTGACTTTCTCCCTCGCTAGAGACGCTCCGGCGCCCGAAAGGAAAGGGTGGCGCTGCGCTCCGGGGTGCACGAGCCGACAGCGC
CCGACCCCAACGGGCCGGCCCCGCCAGCGCCGCTACCGCCCTGCTCCCGGGCGAGCGGGATGGGCGGGAGTGGAGTGGCGGGTGGAGGGTGGAGACGTCCTGGCCCCCGCCCCGCGTGCACCCCCAGGGGAGGCCGAGCCCGCCGCCCCGGCCC
CGCGCAGGCCCCGCCCGGGACTCCCCTGCGGTCCAGGCCGCGCCCCGGGCTCCGCGCCAGCCAATGAGCGCCGCCCGGCCGGGCGTGCCCCCGCGCCCCAAGCATAAACCCTGGCGCGCTCGCGGCCCGGCACTCTTCTGGTCCCCACAGACTCA
GAGAGAACCCACCATGGTGCTGTCTCCTGCCGACAAGACCAACGTCAAGGCCGCCTGGGGTAAGGTCGGCGCGCACGCTGGCGAGTATGGTGCGGAGGCCCTGGAGAGGTGAGGCTCCCTCCCCTGCTCCGACCCGGGCTCCTCGCCCGCCCGG
ACCCACAGGCCACCCTCAACCGTCCTGGCCCCGGACCCAAACCCCACCCCTCACTCTGCTTCTCCCCGCAGGATGTTCCTGTCCTTCCCCACCACCAAGACCTACTTCCCGCACTTCGACCTGAGCCACGGCTCTGCCCAGGTTAAGGGCCAGGCA
AGAAGGTGGCCGACGCGCTGACCAACGCCGTGGCGCACGTGGACGACATGCCCAACGCGCTGTCCGCCCTGAGCGACCTGCACGCGCACAAGCTTCGGGTGGACCCGGTCAACTTCAAGGTGAGCGGCGGGCCGGGAGCGATCTGGGTCGAGG
GGCGAGATGGCGCCTTCCTCGCAGGGCAGAGGATCACGCGGGTTGCGGGAGGTGTAGCGCAGGCGGCGGCTGCGGGCCTGGGCCCTCGGCCCCACTGACCCTCTTCTCTGCACAGCTCCTAAGCCACTGCCTGCTGGTGACCCTGGCCGCCCAC
CTCCCCGCCGAGTTCACCCCTGCGGTGCACGCCTCCCTGGACAAGTTCCTGGCTTCTGTGAGCACCGTGCTGACCTCCAAATACCGTTAAGCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCTCCCCTTCCTG
CACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGAGTGGGCGGCAGCCTGTGTGTGCCTGAGTTTTTTCCCTCAGCAAACGTGCCAGGCATGGGCGTGGACAGCAGCTGGGACACACATGGCTAGAACCTCTCTGCAGCTGGATAGGGTAGGAAA
AGGCAGGGGCGGGAGGAGGGGATGGAGGAGGGAAAGTGGAGCCACCGCGAAGTCCAGCTGGAAAAACGCTGGACCCTAGAGTGCTTTGAGGATGCATTTGCTCTTTCCCGAGTTTTATTCCCAGACTTTTCAGATTCAATGCAGGTTTGCTGAAA
TAATGAATTTATCCATCTTTACGTTTCTGGGCACTCTTGTGCCAAGAACTGGCTGGCTTTCTGCCTGGGACGTCACTGGTTTCCCAGAGGTCCTCCCACATATGGGTGGTGGGTAGGTCAGAGAAGTCCCACTCCAGCATGGCTGCATTGATCCCC
CATCGTTCCCACTAGTCTCCGTAAAACCTCCCAGATACAGGCACAGTCTAGATGAAATCAGGGGTGCGGGGTGCAACTGCAGGCCCCAGGCAATTCAATAGGGGCTCTACTTTCACCCCCAGGTCACCCCAGAATGCTCACACACCAGACACTGAC
GCCCTGGGGCTGTCAAGATCAGGCGTTTGTCTCTGGGCCCAGCTCAGGGCCCAGCTCAGCACCCACTCAGCTCCCCTGAGGCTGGGGAGCCTGTCCCATTGCGACTGGAGAGGAGAGCGGGGCCACAGAGGCCTGGCTAGAAGGTCCCTTCTCC
CTGGTGTGTGTTTTCTCTCTGCTGAGCAGGCTTGCAGTGCCTGGGGTATCAGAGGAGGGTTCCCGGAGCTGGTAGCCATAAAGCCCTGGCCCTCAACTGATAGGAATATCTTTTATTCCCTGAGCCCATGAATCACCCTTGGTAAACACCTATGG
CAGGCCCTCTGCCTGCGTTTGTGATGTCCTTCCCGCAGCCTGTGGGTACAGTATCAACTGTCAGGAAGACGGTGTCTTCGTTATTTCATCAGGAAGAATGGAGGTCTGACCTAAAGGTAGAAATATGTCAAATGTACAGCAGAGGGCTGGTTGGA
GTGCAGCGCTTTTTACAATTAATTGATCAGAACCAGTTATAAATTTATCATTTCCTTCTCCACTCCTGCTGCTTCAG

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-4309