

SacI

1 CCCTACCCCT ACTCCAAAA TGTCAAAAGAT ACAGTCTCAG AAGACCAAAG GGCTATTGAG
 GGGATGGGGA TGAGGTTTTT ACAGTTTCTA TGTCAGAGTC TTCTGGTTTC CCGATAACTC
 61 ACTTTTCAAC AAAGGGTAAT TTCGGGAAAC CTCCTCGGAT TCCATTGCCC AGCTATCTGT
 TGAAAAGTTG TTTCCCATTA AAGCCCTTG GAGGAGCCTA AGGTAACGGG TCGATAGACA
 121 CACTTCATCG AAAGGACAGT AGAAAAAGGAA GGTGGCTCCT ACAAATGCCA TCATTGCGAT
 GTGAAGTAGC TTTCCGTGCA TCTTTTCCTT CCACCGAGGA TGTTTACGGT AGTAACGCTA
 181 AAAGGAAAGG CTATCATTCA AGATGCCTCT GCCGACAGTG GTCCCAAAGA TGGACCCCA
 TTTCCTTTTCC GATAGTAAGT TCTACGGAGA CGGCTGTCAC CAGGGTTTTCT ACCTGGGGGT
 241 CCCACGAGGA GCATCGTGGA AAAAGAAGAC GTTCCAACCA CGTCTTCAA GCAAGTGGAT
 GGGTGCTCCT CGTAGCACCT TTTTCTTCTG CAAGGTTGGT GCAGAAGTTT CGTTCACCTA
 301 TGATGTGACA TCTCCACTGA CGTAAAGGAT GACGCACAAT CCCACCCCTA CTCCAAAAT
 ACTACACTGT AGAGGTGACT GCATTCCCTA CTGCGTGTTA GGGTGGGGAT GAGGTTTTTA
 361 GTCAAAGATA CAGTCTCAGA AGACCAAAGG GCTATTGAGA CTTTTCAACA AAGGTAATT
 CAGTTTCTAT GTCAGAGTCT TCTGGTTTCC CGATAACTCT GAAAAGTTGT TTCCCATTA
 421 TCGGGAAACC TCCTCGGATT CCATTGCCCA GCTATCTGTC ACTTCATCGA AAGGACAGTA
 AGCCCTTTGG AGGAGCCTAA GGTAACGGGT CGATAGACAG TGAAGTAGCT TTCCTGTCAT
 481 GAAAAGGAAG GTGGCTCCTA CAAATGCCAT CATTGCGATA AAGGAAAGGC TATCATTCAA
 CTTTTCTTTC CACCGAGGAT GTTACGGTA GTAACGCTAT TTCTTTCCG ATAGTAAGTT
 541 GATGCCTCTG CCGACAGTGG TCCCAAAGAT GGACCCCCAC CCACGAGGAG CATCGTGGA
 CTACGGAGAC GGCTGTCACC AGGGTTTCTA CCTGGGGGTG GGTGCTCCTC GTAGACCCTT
 601 AAAGAAGACG TTCCAACCAC GTCTTCAAAG CAAGTGGATT GATGTGACAT CTCACTGAC
 TTTCTTCTGC AAGGTTGGTG CAGAAGTTTC GTTCACCTAA CTACACTGTA GAGGTGACTG
 661 GTAAGGGATG ACGCACAATC CCACTATCCT TCGCAAGACC CTTCTCTAT ATAAGGAAGT
 CATTCCCTAC TGCCTGTTAG GGTGATAGGA AGCGTTCTGG GAAGGAGATA TATTCCTTCA

SpeI BspEI XbaI

721 TCATTTTATT TGGAGAGGAC AGCCAGATC AACTAGTCTT AAGTCCGGAG CTAGCTCTAG
 AGTAAAGTAA ACCTCTCCTG TCGGGTCTAG TTGATCAGAA TTCAGGCCTC GATCGAGATC

XbaI XhoI XmaI NcoI AatII AgeI BamHI

+2 M V S K G E E L F T

781 AGACGTCTCG AGGACCGGTC CCGGGGGATC CATGGTGAGC AAGGGCGAGG AGCTGTTTAC
 TCTGCAGAGC TCCTGGCCAG GGCCCTTAG GTACCACTCG TTCCCGCTCC TCGACAAGTG

+2 T G V V P I L V E L D G D V N G H K F S V

841 CGGGGTGGTG CCCATCCTGG TCGAGCTGGA CGGCGACGTA AACGGCCACA AGTTCAGCGT
 GCCCACCAC GGGTAGGACC AGCTCGACCT GCCGCTGCAT TTGCCGGTGT TCAAGTCGCA

+2 V S G E G E G D A T Y G K L T L K F I C T

901 GTCCGGCGAG GCGAGGGCG ATGCCACCTA CGGCAAGCTG ACCCTGAAGT TCATCTGCAC
 CAGGCCGCTC CCGCTCCCGC TACGGTGGAT GCCGTTTCGAC TGGGACTTCA AGTAGACGTG

+2 T T G K L P V P W P T L V T T L T W G V Q

961 CACCGGCAAG CTGCCCGTGC CCTGGCCAC CCTCGTGACC ACCCTGACCT GGGCGTGCA
 GTGGCCGTTT GACGGGCACG GGACCGGGTG GGAGCACTGG TGGGACTGGA CCCCACAGT

+2 Q C F A R Y P D H M K Q H D F F K S A M P

1021 GTGCTTCGCC CGCTACCCCG ACCACATGAA GCAGCACGAC TTCTTCAAGT CCGCCATGCC
 CACGAAGCGG GCGATGGGGC TGGTGTACTT CGTCGTGCTG AAGAAGTTCA GGCGGTACGG

2041	GCCCCGTTTC	CAGTCGGGAA	ACCTGTCGTG	CCAGCTGCAT	TAATGAATCG	GCCAACGCGC
	CGGGCGAAAAG	GTCAGCCCTT	TGGACAGCAC	GGTCGACGTA	ATTACTTAGC	CGGTTGCGCG
2101	GGGGAGAGGC	GGTTTGCFTA	TTGGGCGCTC	TTCCGCTTCC	TCGCTCACTG	ACTCGCTGCG
	CCCCTCTCCG	CCAAACGCAT	AACCCGCGAG	AAGGCGAAGG	AGCGAGTGAC	TGAGCGACGC
2161	CTCGGTCTGTT	CGGCTGCGGC	GAGCGGTATC	AGCTCACTCA	AAGGCGGTAA	TACGGTTATC
	GAGCCAGCAA	GCCGACGCCG	CTCGCCATAG	TCGAGTGAGT	TTCCGCCATT	ATGCCAATAG
2221	CACAGAATCA	GGGGATAACG	CAGGAAAGAA	CATGTGAGCA	AAAGGCCAGC	AAAAGGCCAG
	GTGTCTTAGT	CCCCTATTGC	GTCCTTTCTT	GTACACTCGT	TTTCCGGTCG	TTTCCGGTTC
2281	GAACCGTAAA	AAGGCCGCGT	TGCTGGCGTT	TTTCCATAGG	CTCCGCCCCC	CTGACGAGCA
	CTTGGCATT	TTCCGGCGCA	ACGACCGCAA	AAAGGTATCC	GAGGCGGGGG	GAAGGTATCC
2341	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG	GCGAAACCCG	ACAGGACTAT	AAAGATACCA
	AGTGTTTTTTA	GCTGCGAGTT	CAGTCTCCAC	CGCTTTGGGC	TGTCCTGATA	TTTCTATGGT
2401	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGCG	CTCTCCTGTT	CCGACCCTGC	CGCTTACCGG
	CCGCAAAGGG	GGACCTTCGA	GGGAGCACGC	GAGAGGACAA	GGCTGGGACG	GCGAATGGCC
2461	ATACCTGTCC	GCCTTTCTCC	CTTCGGGAAG	CGTGGCGCTT	TCTCATAGCT	CACGGTGTAG
	TATGGACAGG	CGGAAAGAGG	GAAGCCCTTC	GCACCGCGAA	AGAGTATCGA	GTGCGACATC
2521	GTATCTCAGT	TCGGTGTAGG	TCGTTCCGTC	CAAGCTGGGC	TGTGTGCACG	AACCCCCCGT
	CATAGAGTCA	AGCCACATCC	AGCAAGCGAG	GTTTCGACCCG	ACACACGTGC	TTGGGGGGCA
2581	TCAGCCCGAC	CGCTGCGCCT	TATCCGGTAA	GTATCGTCTT	GAGTCCAACC	CGGTAAGACA
	AGTCGGGCTG	GCGACGCGGA	ATAGGCCATT	GATAGCAGAA	CTCAGGTTGG	CCCATCTGT
2641	CGACTTATCG	CCACTGGCAG	CAGCCACTGG	TAACAGGATT	AGCAGAGCGA	GGTATGTAGG
	GCTGAATAGC	GGTGACCGTC	GTCGGTGACC	ATTGTCTTAA	TCGTCTCGCT	CCATACATCC
2701	CGGTGCTACA	GAGTTCTTGA	AGTGGTGGCC	TAACACGGC	TACACTAGAA	GGACAGTATT
	GCCACGATGT	CTCAAGAACT	TCACCACCGG	ATTGATGCCG	ATGTGATCTT	CCTGTACATA
2761	TGGTATCTGC	GCTCTGCTGA	AGCCAGTTAC	CTTCGGAAAA	AGAGTTGGTA	GCTCTTGATC
	ACCATAGACG	CGAGACGACT	TCGGTCAATG	GAAGCCTTTT	TCTCAACCAT	CGAGAACTAG
2821	CGGCAAACAA	ACCACCGCTG	GTAGCGGTGG	TTTTTTTTGTT	TGCAAGCAGC	AGATTACGCG
	GCCGTTTGTT	TGGTGGCGAC	CATCGCCACC	AAAAAAACAA	ACGTTCCGTC	TCTAATGCGC
2881	CAGAAAAAAA	GGATCTCAAG	AAGATCCTTT	GATCTTTTCT	ACGGGGTCTG	ACGCTCAGTG
	GTCTTTTTTT	CCTAGAGTTC	TTCTAGGAAA	CTAGAAAAGA	TGCCCCAGAC	TGCGAGTCAC
2941	GAACGAAAAC	TCACGTTAAG	GGATTTTGGT	CATGAGATTA	TCAAAAAGGA	TCTTACACCTA
	CTTGCTTTTG	AGTGCAATTC	CCTAAAACCA	GTACTCTAAT	AGTTTTTCCT	AGAAGTGGAT
		Dral		Dral		
3001	GATCCTTTTA	AATTA AAAAT	GAAGTTTAA	ATCAATCTAA	AGTATATATG	AGTAACTTG
	CTAGGAAAAT	TTAATTTTTA	CTTCAAAATT	TAGTTAGATT	TCATATATAC	TCATTTGAAC
3061	GTCTGACAGT	TACCAATGCT	TAATCAGTGA	GGCACCTATC	TCAGCGATCT	GTCTATTTTCG
	CAGACTGTCA	ATGGTTACGA	ATTAGTCACT	CCGTGGATAG	AGTCGCTAGA	CAGATAAAGC
3121	TTCATCCATA	GTTGCCTGAC	TCCCGTCTGT	GTAGATAACT	ACGATACGGG	AGGGCTTACC
	AAGTAGGTAT	CAACGGACTG	AGGGGCAGCA	CATCTATTGA	TGCTATGCC	TCCCGAATGG
3181	ATCTGGCCCC	AGTGCTGCAA	TGATACCGG	AGACCCACGC	TCACCGGCTC	CAGATTTATC
	TAGACCGGGG	TCACGACGTT	ACTATGGCGC	TCTGGGTGCG	AGTGGCCGAG	GTCTAAATAG
3241	AGCAATAAAC	CAGCCAGCCG	GAAGGGCCGA	GCGCAGAAGT	GGTCTTGCAA	CTTTATCCGC
	TCGTTATTTG	GTCGGTCCGG	CTTCCCGGCT	CGCGTCTTCA	CCAGGACGTT	GAAATAGGCG
3301	CTCCATCCAG	TCTATTAATT	GTTGCCGGGA	AGCTAGAGTA	AGTAGTTCGC	CAGTTAATAG
	GAGGTAGGTC	AGATAATTAA	CAACGGCCCT	TCGATCTCAT	TCATCAAGCG	GTCAATATATC
3361	TTTGCGCAAC	GTTGTTGCCA	TTGCTACAGG	CATCGTGGTG	TCACGCTCGT	CGTTTGGTAT
	AAACCGGTTG	CAACAACGGT	AACGATGTCC	GTAGCACCAC	AGTGCAGCA	GCAAACCATA
3421	GGCTTCATTC	AGCTCCGGTT	CCCAACGATC	AAGGCGAGTT	ACATGATCCC	CCATGTTGTG
	CCGAAGTAAG	TCGAGGCCAA	GGGTTGCTAG	TTCCGCTCAA	TGTACTAGGG	GGTACAACAC
3481	CAAAAAAGCG	GTTAGCTCCT	TCGGTCCCTC	GATCGTTGTC	AGAAGTAAGT	TGGCCGAGT
	GTTTTTTTCG	CAATCGAGGA	AGCCAGGAGG	CTAGCAACAG	TCTTCATTCA	ACCGCGTCA
3541	GTTATCACTC	ATGGTTATGG	CAGCACTGCA	TAATTCTCTT	ACTGTCATGC	CATCCGTAAG
	CAATAGTGAG	TACCAATACC	GTCGTGACGT	ATTAAGAGAA	TGACAGTACG	GTAGGCATTC

Rsal

3601	ATGCTTTTCT TACGAAAAGA	GTGACTGGTG CACTGACCAC	AGTACTCAAC TCATGAGTTG	CAAGTCATTC GTTCCAGTAAG	TGAGAATAGT ACTCTTATCA	GTATGCGGCG CATACGCCGC
------	--------------------------	--------------------------	--------------------------	---------------------------	--------------------------	--------------------------

Dral

3661	ACCGAGTTGC TGGCTCAACG	TCTTGCCCGG AGAACGGGCC	CGTCAATACG GCAGTTATGC	GGATAATACC CCTATTATGG	GCGCCACATA CGCGGTGTAT	GCAGAACTTT CGTCTTGAAA
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Dral

3721	AAAAGTGCTC TTTTACGAG	ATCATTGGAA TAGTAACCTT	AACGTTCTTC TTGCAAGAAG	GGGGCGAAAA CCCCGCTTTT	CTCTCAAGGA GAGAGTTCCT	TCTTACCGCT AGAATGGCGA
------	-------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

3781	GTTGAGATCC CAACTCTAGG	AGTTCGATGT TCAAGCTACA	AACCCACTCG TTGGGTGAGC	TGCACCCAAC ACGTGGGTTG	TGATCTTCAG ACTAGAAGTC	CATCTTTTAC GTAGAAAATG
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

3841	TTTACCAGC AAAGTGGTCG	GTTTCTGGGT CAAAGACCCA	GAGCAAAAAC CTCGTTTTTG	AGGAAGGCAA TCCTTCCGTT	AATGCCGCAA TTACGGCGTT	AAAAGGGAAT TTTTCCCTTA
------	-------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

3901	AAGGGCGACA TTCCCGCTGT	CGGAAATGTT GCCTTTACAA	GAATACTCAT CTTATGAGTA	ACTCTTCCTT TGAGAAGGAA	TTTCAATATT AAAGTTATAA	ATTGAAGCAT TAACTTCGTA
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

3961	TTATCAGGGT AATAGTCCCA	TATTGTCTCA ATAACAGAGT	TGAGCGGATA ACTCGCCTAT	CATATTTGAA GTATAAACTT	TGTATTTAGA ACATAAATCT	AAAATAACA TTTTATTTGT
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------------

4021	AATAGGGGTT TTATCCCAA	CCGCGCACAT GGCGCGTGTA	TTCCCGAAA AAGGGCTTT	AGTGCCACCT TCACGGTGGA	AAATTGTAAG TTAACATTC	CGTTAATATT GCAATTATAA
------	-------------------------	--------------------------	------------------------	--------------------------	-------------------------	--------------------------

4081	TTGTTAAAAT AACAATTTTA	TCGCGTTAAA AGCGCAATTT	TTTTTGTTAA AAAAACAATT	ATCAGCTCAT TAGTCGAGTA	TTTTTAACCA AAAAATTGGT	ATAGGCCGAA TATCCGGCTT
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

4141	ATCGGCAAAA TAGCCGTTTT	TCCCTTATAA AGGGAATATT	ATCAAAAAGAA TAGTTTTCTT	TAGACCGAGA ATCTGGCTCT	TAGGGTTGAG ATCCCAACTC	TGTTGTTCCA ACAACAAGGT
------	--------------------------	--------------------------	---------------------------	--------------------------	--------------------------	--------------------------

4201	GTTTGGAAAC CAAACCTTGT	AGAGTCCACT TCTCAGGTGA	ATTAAGAAGC TAATTTCTTG	GTGGACTCCA CACCTGAGGT	ACGTCAAAGG TGCAGTTTCC	GCGAAAAACC TGCAGTTTTGG
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	---------------------------

4261	GTCTATCAGG CAGATAGTCC	GCGATGGCCC CGCTACCGGG	ACTACGTGAA TGATGCACTT	CCATCACCTT GGTAGTGGGA	AATCAAGTTT TTAGTTCAAA	TTTGGGGTCG AAACCCAGC
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------------

4321	AGGTGCCGTA TCCACGGCAT	AAGCACTAAA TTCGTGATTT	TCGGAACCTT AGCCTTGGGA	AAAGGGAGCC TTTCCCTCGG	CCCATTAG GGGCTAAATC	AGCTTGACGG TCGAACTGCC
------	--------------------------	--------------------------	--------------------------	--------------------------	------------------------	--------------------------

4381	GGAAAGCCGG CCTTTCCGGC	CGAACGTGGC GCTTGCACCG	GAGAAAGGAA CTCTTTCCTT	GGGAAGAAAG CCCTTCTTTC	CGAAAGGAGC GCTTTCCTCG	GGGCGCTAGG CCCGGATCC
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	-------------------------

4441	GCGCTGGCAA CGCGACCGTT	GTGTAGCGGT CACATCGCCA	CACGCTGCGC GTGCGACGCG	GTAACCACCA CATTGGTGGT	CACCCGCCGC GTGGGCGGGC	GCTTAATGCG CGAATTACGC
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

4501	CCGCTACAGG GGCGATGTCC	GCGCGTCCCA CGCGCAGGGT	TTCCGCATTC AAGCGTAAG	AGGCTGCGCA TCCGACGCGT	ACTGTTGGGA TGACAACCCT	AGGGCGATCG TCCCGTAGC
------	--------------------------	--------------------------	-------------------------	--------------------------	--------------------------	-------------------------

4561	GTGCGGGCCT CACGCCCGGA	CTTCGCTATT GAAGCGATAA	ACGCCAGCTG TGCGGTCGAC	GCGAAAGGGG CGCTTTCCCC	GATGTGCTGC CTACACGACG	AAGGCGATTA TTCCGCTAAT
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

4621	AGTTGGGTAA TCAACCCATT	CGCCAGGGTT GCGGTCCCAA	TTCCCAGTCA AAGGGTCAGT	CGACGTTGTA GCTGCAACAT	AAACGACGGC TTTGCTGCCG	CAGTGAATTG GTCACTTAAC
------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------	--------------------------

Sacl

4681	TAATACGACT ATTATGCTGA	CACTATAGGG GTGATATCCC	CGAATTGGAG GCTTAACCTC	CT GA
------	--------------------------	--------------------------	--------------------------	----------