

File S.7: Fasta sequence of strain-specific proteins of NCDO 2118 identified by LC-MS^E analysis.

>NCDO2118_RS04930 NCDO2118_RS04930 hypothetical protein 1000763:1001092 reverse
MW:12688

VDDIQLTKDSKQLLAIYKEYLSKINNGVSKNSAKRIGHISDICELVPDWLPDDVLETMN

ELSRSEYVFNRYGNNTIIDSYLLDKTIIYFENKNINTVKEIADWVFKLI

>NCDO2118_RS07360 NCDO2118_RS07360 hypothetical protein 1537567:1538367 reverse
MW:29298

IGAVAVLIIGGVSFALVKHNETGGVKTNASGLKSSSKKEMTSKTSEINESQKSSSSQTSS

SSQEETTQSQLSVDTKNLTENQFEQWVAITQSNQMQDTHPLRITVYKSDNNLHQAIDYT

ATRIDSVNSYKVNSSGELEAIDFEHYPAMKVVSQSFPKQKQLTTNQINMWVAVQDAL

FEQKGLKPGIDVVYDLETATDGNIANVKVFKIDYTFSANGGADNANKTLNEFRINSDGN

LEIKDVNSETGYRIISKVYMDTSMVQ

>NCDO2118_RS08490 NCDO2118_RS08490 hypothetical protein 1775121:1775813 reverse
MW:26900

VQHVFYFDDSGVLHPNAPERVFVYAGYVFIGKESKDDAIRKYRAVCQKIRDEIKSDGEI

KAFGLSNRNKNRLYQILAKEESVGLTVDIERVYSSILDNKKSIHRYKDYVLKMLVKEKLK

ELIFEQKINPNGSVALHIFVDEQPTSTDGFYKLRDTIYEVKVGITNFQYNKRHAPLFNS

SVKCDVTFVDSSKYPLVQASDILANRIWNSFRLKKPNLREFVNHKCLHMP

>NCDO2118_RS08730 NCDO2118_RS08730 hypothetical protein 1810039:1810539 forward
MW:19075

MNKLEITSFEYAVAVLNELASNNNESFVPFEIVWNTTLGPAKARTIYDGTTPIISETII

GSDYIDRVFYNPRIKESDIFSIRHEIFGYFEDKFSGLQHKSLIERHDLLMIKLELSKL

DITSELTPNYSTIFDFELIDGSMKLPFIHSDSIEIKEPISLISKN

>NCDO2118_RS08735 NCDO2118_RS08735 hypothetical protein 1810517:1810765 reverse
MW:9754

MNPEDVKVTKESEEETARDRYLKNFKRDNLPLDFFGKPKLNEVGLKILDLLKKEDLTHEQ

AYASLQYVYNLIKYESNFLKLN

>NCDO2118_RS08645 NCDO2118_RS08645 hypothetical protein 1801600:1802301 reverse
MW:27406

IDIENKNTDFLNLDLSDTSEYFGESNLFKELDKLPKPDILASPPCESWSNASAMLNGNV

CWYTESTDTMFGQEFVSNEFTIRTRQQLETKN DTPFKKHWWKTVYSRLNGELCAFNTIRI

IERYQPEVWVIENPQSSRIWKYYKQIQDFQGIKNIAHYSAYDSERYSKKPTCFYSNLMFN

LKTTDEQSKLTFQGLGDKGISRSYNVRSEIPLQLIKDILDQCFLKLEKNSKEI