

Notes on the vaccinia virus open reading frame library

- 1) This collection of vaccinia virus (VACV) ORF clones was made using re-amplified PCR products generated originally for integration into yeast for a 2-hybrid project: McCraith S, Holtzman T, Moss B, and Fields S. (2000). Genome-wide analysis of vaccinia virus protein-protein interactions. Proc. Natl. Acad. Sci. USA 97: 4879-84.
- 2) The ORF definition for the original 2-hybrid project was based on an unpublished sequence of vaccinia virus strain WR in collaboration with Dr Bernard Moss at NIAID, NIH, Bethesda, MD.
- 3) The spreadsheet listing all the ORFs includes 1) an ORF name that relates only to this library and no published genome annotation, 2) the name of the ORF (if present) in VACV strain Copenhagen (GenBank accession M35027) and 3) the name of the ORF (if annotated) in VACV strain WR sequence (GenBank accession AY243312), which was published after these clones were made.
- 4) There are some differences between the ORFs defined by this collection of clones and the ORFs annotated on the published VACV WR sequence. Some arise from sequence errors in the unpublished sequence, others from differences in annotation. Please note that the amino acid sequence given for each ORF in the spreadsheet is derived from the unpublished sequence and has some errors.
- 5) All ORFs retain the native stop codon and nearly all have a common sequence added to each end that allowed all ORFs in the original set to be re-amplified by PCR using a single primer pair. These common sequences are: 5' end: CGAATTCCAGCTGACCACC and 3' end: CCGATCCCCGGGAATTGC. **NB:** some ORFs were cloned directly from VACV genomic DNA for my project and lack these common sequences, these are VVorfs 8, 27, 30, 31, 107, 151, 243.
- 6) For all ORFs flanked by the common sequences, Bam HI excises a product that is roughly the size of the insert (unless there are internal BamHI sites); there is a Bam HI site in the 3' common sequence and one just 5' of the TOPO insertion site in the vector.
- 7) Up to roughly 500 bp from the 5' end of most clones was sequenced and many have base changes within the sequenced region (primer and PCR fidelity errors). The extent of my sequencing and any clones with known errors are the last 2 columns of the spreadsheet.
- 8) No attempt has been made to test how many of these ORFs express a protein after transfection into mammalian cells.
- 9) Given all the above, the clone collection remains a useful (but by no means ideal) set for expression screening and should not be relied on as a complete source of error-free ORFs.

For any more information feel free to contact me: david.tscharke@anu.edu.au

David Tscharke

(5th Oct 2006)