MAR 25, 2020

OPEN ACCESS

DOI:
dx.doi.org/10.17504/protocols.io.bd3ri8m6

Collection Citation: Graham J Etherington, Darren Heavens, David Baker, Ashleigh Lister, Rose McNelly, Gonzalo Garcia, Bernardo Clavijo, Iain Macaulay, Wilfried Haerty, Federica Di Palma 2020. Protocols for "Sequencing smart: De novo sequencing and assembly approaches for a non-model mammal". **protocols.io** <https://dx.doi.org/10.17504/protocols.io.bd3ri8m6>

License: This is an open access collection distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working
 We use this protocol and it's working

Created: Mar 23, 2020

Last Modified: Mar 26, 2020

COLLECTION integer ID:
 34641

🌐 Protocols for "Sequencing smart: De novo sequencing and assembly approaches for a non-model mammal"

📁 In 1 collection

Graham J Etherington¹, Darren Heavens¹, David Baker¹, Ashleigh Lister¹, Rose McNelly¹, Gonzalo Garcia¹, Bernardo Clavijo¹, Iain Macaulay¹, Wilfried Haerty¹, Federica Di Palma¹

¹The Earlham Institute, Norwich Research Park, Norwich, NR4 7UZ, United Kingdom

GigaScience Press

Graham Etherington
 The Earlham Institute

ABSTRACT

Whilst much sequencing effort has focused on key mammalian model organisms such as mouse and human, little is known about the correlation between genome sequencing techniques for non-model mammals and genome assembly quality. This is especially relevant to non-model mammals, where the samples to be sequenced are often degraded and low quality. A key aspect when planning a genome project is the choice of sequencing data to generate. This decision is driven by several factors, including the biological questions being asked, the quality of DNA available, and the availability of funds. Cutting-edge sequencing technologies now make it possible to achieve highly contiguous, chromosome-level genome assemblies, but relies on good quality high-molecular-weight DNA. Here we use a range of different genomic technologies generated from a roadkill European Polecat (*Mustela putorius*) to assess various assembly techniques on this low-quality sample. We evaluated different approaches for de novo assemblies and discuss their value in relation to biological analyses. The high degree of variability between each de novo assembly method (assessed from the seven key metrics) highlights the importance of carefully devising the sequencing strategy to be able to carry out the desired analysis. Adding more data to genome assemblies does not always results in better assemblies so it is important to understand the nuances of genomic data integration explained here, in order to obtain cost-effective value-for-money when sequencing genomes.

Keywords: polecat,
vertebrate, non-model
organism, Illumina, chromium,
Bionano, assembly,
sequencing

Protocol

NAME

Long Mate Pair Library Construction Protocol

VERSION 1

CREATED BY

Graham EtheringtonThe Earlham Institute

[OPEN](#) →

Protocol

NAME

Amplification Free Paired End Library Construction Protocol

VERSION 1

CREATED BY

Graham EtheringtonThe Earlham Institute

[OPEN](#) →

Protocol

NAME

10x Genomics Library Construction

VERSION 1

CREATED BY

Graham EtheringtonThe Earlham Institute

[OPEN](#) →

Protocol

NAME

Bionano genome mapping from animal tissue

VERSION 1

CREATED BY

Graham EtheringtonThe Earlham Institute

[OPEN](#) →