The Microsoft Biology Foundation

Simon Mercer, Michael Zyskowski and Bob Davidson
Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA – msrerbio@microsoft.com

http://mbf.codeplex.com

Available under the Microsoft Public License (http://opensource.org/licenses/ms-pl.html)

The Microsoft Biology Foundation (MBF) is a library of common bioinformatics and genomics functionality built on top of the .NET framework. Functions include parsers and writers for common bioinformatics file formats, connectors to common web services, and algorithms for assembling and aligning DNA sequences. The project is released under the OSI-compliant MS-PL open source license (http://opensource.org/licenses/ms-pl.html) and is available for download from http://mbf.codeplex.com. The MBF project is guided by the user community through a Technical Advisory Board drawn from academia and commerce, with responsibility to maintain code quality and steer future development to respond to the needs of the scientific community. MBF is a community-led and community-curated project, and encourages bug fixes, feature requests and code contributions from all members of the commercial and academic life science community.

The .NET framework provides some advantages in terms of development and use of MBF – code can be added in any .NET compliant language (there are currently around 70), and the functions in MBF can be accessed in many different ways – compiled into an executable, wrapped as workflow activities, accessed on the command-line using a scripting language such as Python, or used in Microsoft Office plug-ins to add biological functionality to applications such as Microsoft Word and Excel.

Microsoft researchers frequently apply computer research to challenges in biology and this leads to the development of unique scientific tools – for example machine learning has been applied to the challenges of HIV vaccine design, leading to the creation of a range of tools relating to epitope binding prediction, phylogenetics, haplotyping and population studies (http://mscompbio.codeplex.com). MBF is being actively used and extended by Microsoft Research projects such as these, as well as by the wider academic and commercial community.