Electron microscope images are affected by the contrast transfer function (CTF) of the transmission electron microscope, which arises from the aberrations of the lenses and from the defocus used in imaging. The CTF introduces spatial frequency-dependent oscillations into the Fourier space representation of the image. These oscillations result in contrast changes and modulation of the spectrum amplitudes, as well as an additional envelope that attenuates high-resolution information. Estimation of the CTF and correction for its effects is thus essential for any image to faithfully represent a projection of the specimen.

We present in this work the "CTF Benchmark" which provides an opportunity to the researchers in the field to carry out a comprehensive evaluation of their CTF estimation methods based on a common database of image (see more details at URL: http://i2pc.cnb.csic.es/3dembenchmark)

A total of 21 different uploads were submitted, covering most of the software packages in the field including: ace, appion, bsoft, ctffind, dudelft, eman, fei, imagic, particle, sparx, spider, xmipp, etc. The main conclusions of this Benchmark are:

* In general, for quality datasets such as the ones in this Benchmark, about 40\% of the datasets are practically not limited by CTF estimation errors while, for the rest, CTF errors set the limit in between 5 to 7 A resolution at most.
* As a rule, and certainly not unexpected, estimations of the mean defocus is much better than astigmatism estimation. Although the error is lower when estimating astigmatism than when only a 1D CTF model is used, indicating that CTF defocus estimation is beneficial to achieve high resolution.
* It is very clear the trend that when a dataset is specially suited for high resolution most software packages provide similar estimations for the CTF parameters. In other words, when a dataset is "good", is "good" for everything. Consequently, there must be image characteristics that dictate their ability to provide high resolution structural information.
* Our initial hypothesis was that micrographs with carbon support and with a high concentration of particles were going to be "the best", simply because of their increased scattering power did not hold. The "better" datasets do not have carbon, and one of the best does not have a particularly large density of particles.
* Finally, synthetic data behave similarly as experimental data with respect to CTF estimation

Acknowledgement: The authors would like to acknowledge economical support from the Spanish Ministry of Economy and Competitiveness through grants AIC–A–2011–0638 and BIO2010-16566, the Comunidad de Madrid through grant CAM(S2010/BMD-2305) and the NSF through grant 1114901, as well as postdoctoral “Juan de la Cierva” grant with reference JCI-2011-10185. C.O.S. Sorzano is recipient of a Ramón y Cajal fellow.