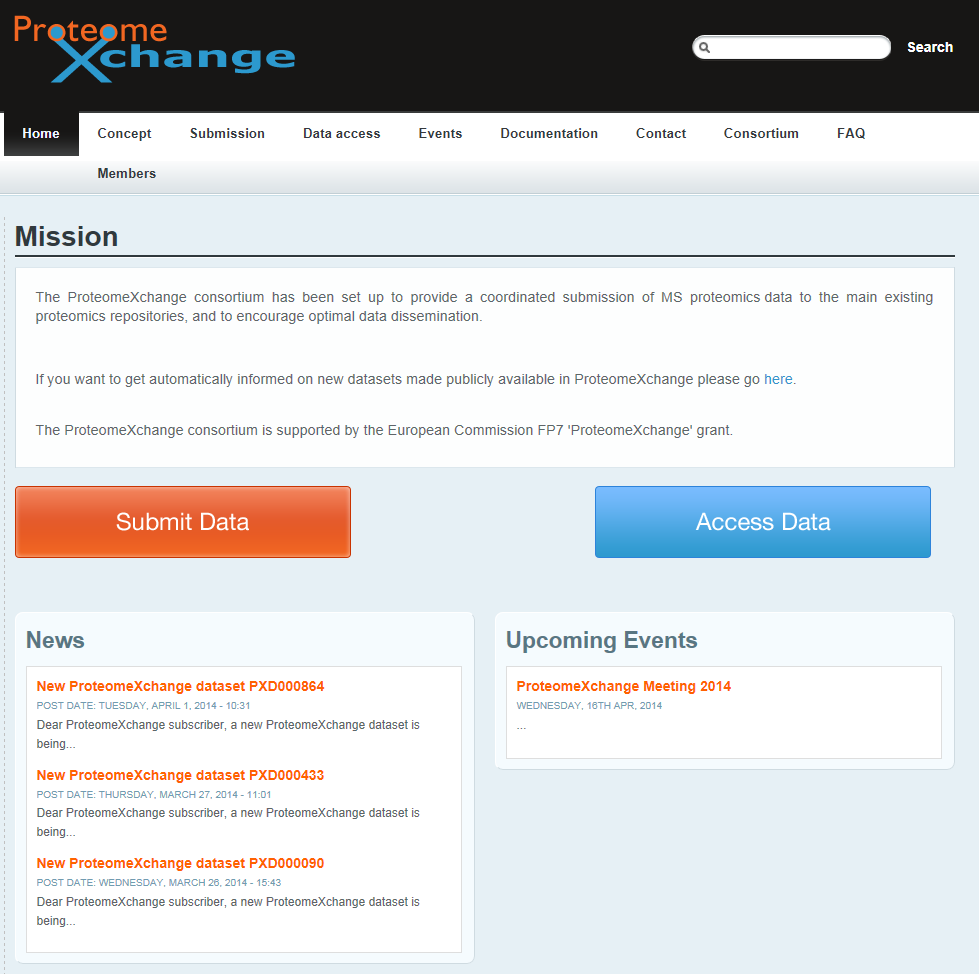
Data Sharing

In the previous chapters we identified proteins using curated database sequences and enriched our results with knowledge from external resources. In order to allow the community to benefit from your results in turn, online repositories are available to enable the exchange of data. Note that making the data public is now required by most journals prior to publication.

**Proteomics Results**

The ProteomeXchange consortium was launched to unify and simplify the sharing of proteomic data1. There, identification results are stored via the proteomics identifications database[2](#_ENREF_1) (PRIDE, <http://www.ebi.ac.uk/pride>). In order to make our data understandable to all, we need to convert it to standard formats and document the processing steps using standardized vocabulary[3](#_ENREF_2). We will here use the PRIDE XML format, however in the near future this will be updated to the new standard format called mzIdentML.[4](#_ENREF_3)

If you go on the ProteomeXchange website (<http://www.proteomexchange.org>) you will see the starting point for submitting and accessing datasets:



In this chapter, we will see how to submit, browse and reprocess online datasets.

References

1. Juan A Vizcaíno et al. ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* **32**, 223–226 (2014).

2. Martens, L. et al. PRIDE: the proteomics identifications database. *Proteomics* **5**, 3537-3545 (2005).

3. Martens, L., Palazzi, L.M. & Hermjakob, H. Data standards and controlled vocabularies for proteomics. *Methods in molecular biology* **484**, 279-286 (2008).

4. Jones, A.R. et al. The mzIdentML data standard for mass spectrometry-based proteomics results. *Mol Cell Proteomics* **11**, M111 014381 (2012).