## Access to bioinformatics data for users of Biological Resource Centres

- Access a summary of all DNA and RNA sequences based on a taxonomic name is
  provided from the <u>European Nucleotide Archive</u>: starting from the <u>Advanced Search page</u>,
  select the "Taxon" domain and enter a taxonomic name at any rank, leaving the checking
  the "include subordinate taxa" box unchecked. Browse through to view the result and select
  the "Portal" tab if it is not already showing. This provides a summary of all held sequence
  data with clickable links to browse further into these data. See the example results for
  family *Prochloraceae* here and the example of the query search here.
- List assembled/annotated sequences linked to a given culture collection: using the
   <u>cross-reference search</u> of the <u>European Nucleotide Archive</u>, select a collection of interest,
   such as the Culture Collection of Algae and Protozoa (CCAP), from the "xref" pull down
   menu and choose "sequence" from the "target" pull down menu.
- List proteins linked to a given organism: using the Universal Protein Resource (UniProt), go to the home page <a href="here">here</a>. Go to the search bar in the UniProt banner at the top of the page and click on the dropdown to the left of the search box and select your dataset. The default is UniProtKB. Click on 'advanced' to the right of the search box and select term type "organism" and enter "Pardachirus", then hit the search button. This provides a summary of all held proteins isolated from the genus Pardachirus with clickable links to further browse these data, including different Pardachirus species such as Pardachirus pavoninus and Pardachirus marmoratus. See the example result for Pardachirus here.
- Access all genomes from marine microbial species is provided from the MarRef databases. The MAR databases are MarRef, MarDB and MarCat. Starting from the MAR databases home page, select one of the three databases to search for specific genomes or attributes, or BLAST nucleotide or protein sequences against the content in the MAR databases. The MAR databases, contextual and/or sequence data resources can be downloaded from the MAR Download page which can be found here.
- List marine microbes based upon taxonomy: from MarDB starting page or MarRef starting page select "Browse" and by using the filter panel on the right side of the page select the taxa of interest on a phylum, family or genus. Browse through the results and select the entry of interest.
- Select one specific microbe: from the from MarDB starting page or MarRef starting page and select "Browse" type the microbe of interest in the "Search" box e.g. Verrucosispora maris and the contextual data for the microbe will be provided. Activate the entry by clicking on the entry name either under "MMP ID", "full scientific name", "strain", "type strain", "geographic location" or "collection date". One will get access to all available information for the microbe such as taxon, isolate, host and pathogenicity, assembly, annotation, phenotype and secondary metabolites. See an example results for Verrucosispora maris here.

- List all marine microbes with compound entries in ChEBI or ChEMBL: from the Browse page in <a href="MarDB">MarDB</a> and <a href="MarRef">MarRef</a> enter ChEBI or ChEMBL in the "Search" box on the left side. The results is a list of genomes with links either to ChEBI or ChEMBL and the example results can be seen <a href="here">here</a>.
- View compound predicted secondary metabolites pathways in a particular genome: from the genome page e.g. <a href="Verrucosispora maris">Verrucosispora maris</a> click on the "Secondary metabolites" curtain. Information about AntiSMASH secondary metabolite clusters and types will be available. Link to the result page is available <a href="here">here</a>.