MetPetDB: A database for metamorphic geochemistry
(v. 04 Dec 2007)

Frank S. Spear, Joseph M. Pyle
Department of Earth and Environmental Sciences

Sibel Adali, Boleslaw K. Szymanski, Anthony Waters, Zak Linder, Cagri Ozcalar, Shawn O. Pearce,
Department of Computer Sciences

Rensselaer Polytechnic Institute
Troy, NY, 12180, USA

Abstract

We present a data model for the initial implementation of MetPetDB - a geochemical database specific to metamorphic rock samples. The database is designed around the concept of preservation of spatial relationships, at all scales, of chemical analyses and their textural setting. Virtual objects in the database (samples) represent physical rock samples; each sample contains one or more subsamples with associated geochemical and image data. Samples, subsamples geochemical data and images are described with attributes (some required, some optional); these attributes also serve as search delimiters. All data in the database are classified as *public* (i.e. archived or published data) or *private*. Public data may be freely searched and downloaded. All private data is owned; permission to view, edit, download and otherwise manipulate private may be granted only by the data owner; all such editing operations are recorded by the database to create a data version log. The sharing of data permissions among a group of collaborators researching a common sample is done by the sample owner through the *project manager*. User interaction with MetPetDB is hosted by a web-based platform based upon the Java servlet application programming interface, with the PostgreSQL relational database. The database web portal includes modules that allow the user to interact with the database: *guest* users may search, view public data and *registered* users may save and download public data, upload private data,

create projects, and assign permission levels to project collaborators. An Image Viewer module is available within the web portal for spatial integration of image and geochemical data. A toolkit consisting of plotting and geochemical calculation software will be available for data analysis. Future issues to address include population of the database, integration with other geochemical databases, development of the analysis toolkit, creation of data models for derivative data, and building a community-wide user base. It is believed that this and other geochemical databases will enable more productive collaborations, generate more efficient research efforts, and foster new developments in basic research in the field of solid earth geochemistry.

Introduction

Current research in the geological sciences generates a huge amount of visual and numeric data, and the recognition of the need to catalog and share this data (e.g., see white papers at www.communitytechnology.org/nsf_ci_report, www.adec.edu/nsf/nsfcyberinfrastructure.html, www.nsf.gov/news/special_reports/cyber/index.jsp) as a value-added research tool has spurred the development of a number of organizations (EARTHCHEM http://www.earthchem.org/earthchemWeb/index.jsp; GEON - Geosciences Earth network - http://www.geongrid.org/; SESAR http://www.geosamples.org/) and databases (GEOROC http://georoc.mpch-mainz.gwdg.de/georoc/Start.asp; NAVDAT http://navdat.kgs.ku.edu/; and PETDB http://www.petdb.org/petdbWeb/index.jsp) specific to handling, sharing, and manipulating data generated in the research and analysis of earth materials.

A compelling justification for this development is the belief that fundamentally new science will emerge from the scientific disciplines that possess highly evolved cyberinfrastructure. A leading example in the geosciences is the use of ocean basalt chemistry data mined from PETDB (Lehnert et al., 2000); GeoRef searches utilizing "PetDB" as a delimiting term return hits on several papers published in the last three years (Spiegelman and Kelemen, 2003; Salters and Stracke, 2004; Hirschmann et al., 2003). These papers make use of large data sets that would have been virtually impossible for an individual, or small group of collaborators, to collate until the advent of the aforementioned databases.

Although some metamorphic rocks are included in existing petrologic databases, and at least one database has been developed specifically to address some of the needs of metamorphic petrology (Schmatz et al., 1995), there is no current global database that incorporates the special requirements of metamorphic geochemistry. Interpretations of metamorphic parageneses require not only high precision chemical analyses of minerals present, but evaluation of the textural context of those analyses. Because of this, a large component of the existing metamorphic data is images (photomicrographs, back-scattered, secondary electron (BSE and SE) and cathodoluminescence (CL) images, X-ray maps, etc.) and the location of analyses with respect to these images is a critical component of a metamorphic database. Unfortunately, many of the images and analyses collected during the course of a study never get published owing to limitations of print media. Indeed, it is estimated by the authors that less than one percent of all data collected on metamorphic rocks is published, which suggests that a potential treasure trove of information is waiting to be mined, given the proper infrastructure.

The preservation of spatial relationships at a variety of scales thus differentiates a geochemical database for metamorphic rocks from other geochemical databases, and provides a framework for its design. In our realization, the basic components of a database specific to metamorphic petrology should (1) include the incorporation of bulk rock and mineral analyses; (2) provide for the incorporation of images of any type; (3) preserve the spatial relationship among the various images collected on a thin section; (4) preserve the relationship between analyses and textural setting of the analyses on the relevant images; (5) provide an intuitive user interface that allows for searching, uploading, and downloading sample information as well as facilitating collaborations among researchers; (6) provide a set of tools for recalculation, plotting, and analysis of data; and (7) interface with other geological databases. We envision a database that is populated not only with published, archive data, but incorporates unpublished data that complements what is published, and that can also serve as a research level collaborative tool for researchers globally.

Ideally, a metamorphic database should incorporate both raw data (analyses, images, etc., collected on a sample) as well as derivative or interpretative data for an area (P-T conditions, P-T-t path, cooling history, etc.). The present communication focuses only on the development of the first part of the database for raw data. How to treat derivative/interpretative data is far more complex and will be discussed in a future communication.

The Data Model

The MetPetDB data model (Fig. 1) is built around the concept of a *sample*. The sample is a physical entity (a piece of rock) that exists in the database as a virtual entity with associated information, similar to the approach taken by other geochemical databases (e.g., Lehnert et al., 2000). A number of considerations have been taken into account in the design of the data model. Typically, a thin section is cut from a sample and analytical work (e.g. electron microprobe analysis) is done on a polished thin section that may be different from the original thin section. Whole rock chemical analyses are done on yet different parts of the rock, and other parts of the sample might be used for mineral separates (e.g. zircons). It is considered imperative that the chain of evidence of these distinct pieces of a sample be kept intact, so we have also introduced the concept of a *subsample*, which is, as the name implies, a piece of the original sample. *Chemical analyses* are always done on subsamples (unless the entire sample has been crushed for analysis). Each thin section is a subsample and a piece of the rock used for mineral separates is yet another subsample. *Images* may be associated with the sample or an individual subsample (e.g. thin section). It is also possible that an image is associated with more than one sample (e.g. a photograph of an outcrop), which we have called a *supersample image*. Each of these data types has its own set of attributes, which will be discussed in this section.

Sample

Sample attributes are shown in Table 1 with required attributes labeled "R". The minimum required information for adding a sample to the database are (1) sample number; (2) location (latitude and longitude); (3) rock type (see Table 2); and (4) sample owner, although, of course, additional information is highly desirable. When available, the ISGN (International Geo Sample Number) can provide the unique sample identifier and in the future we expect that all samples will be registered and have ISGNs. Latitude and longitude provide location, with the option to provide a location error if a sample is not well located. Although modern topographic maps and GPS provides location within a few meters, many older data are not so precisely located, and it is deemed important that the location of a sample is not misrepresented. The rock types shown in Table 2 is restrictive relative to the total range of metamorphic rock names that have been used in the literature. However, it was decided that a compact, simple list would be preferable to a comprehensive list because it would help avoid the use of ambiguous names. It is fully recognized that many researchers will wish to use their own favorite (in some cases locally derived) rock names, and this is encouraged. MetPetDB will accommodate as many of these as desired in comment fields. Next, if the data is from a published article (i.e., *archived data*) a unique reference to the publication is required. We have followed the approach of NAVDAT in adopting the 10-digit GeoRef Accession Number as the unique publication identifier. Furthermore, every sample must have an owner so that data use permissions can be properly evaluated. It is recognized that this will create some special issues in the long-term (e.g. the decease of a data owner), which will be dealt with as individual cases. Finally, published data will be labeled "public" and will not be editable (to preserve the published record) whereas unpublished data will be labeled "private" and will be editable by the owner.

Optional data include the country of origin (recognizing that the country may no longer exist), and the "region" from which the sample was collected. Multiple regions are allowed because it is recognized that petrologists will wish to locate samples using multiple geographic names. For example, a sample from the first author's collection comes from the "Eastern Alps", the "Tauern Window", the "Hohe Tauern" the "Großvenediger region", the "Froznitztal", and the "Tauern Eclogite Zone". Multiple regions might make searching more difficult initially, but it is planned to develop absolute geographic outlines for specific regions based on the database population that will facilitate searching.

A list of minerals and modes is optional data because it was discovered that not all published papers provide a list of minerals. Furthermore, it is specifically not required to list a mineral assemblage, because the term "assemblage" carries the implication of a stable equilibrium assemblage at some metamorphic conditions, and that is most definitely an interpretation. Additionally, the metamorphic facies (or grade) attribute (Table 3) can be multi-valued, due to the recognition that, for the same suite of rocks, one author may use the term "amphibolite facies" whereas another author may select "staurolite zone" as the descriptor of metamorphic grade. This will not complicate searches because it will be simple to choose both search criteria. Items such as collector and collection date are self-explanatory. The attribute "present sample location" is included so that individuals who wish to use a sample can find it.

Note that it is not necessary for chemical analyses to be available for a sample to be entered into the database. In this way, MetPetDB differs from other geochemical databases (e.g. PetDB, NAVDAT, and GeoRoc) in that it is, fundamentally, a database of metamorphic samples. This does not limit in any way the usefulness of MetPetDB and it provides the opportunity of recording the types of rocks found in a metamorphic terrane even when only a sample has been taken from the field.

Subsample

The data model recognizes that the main portion of analytical effort is expended on aliquots derived from the sample itself. Such aliquots may take the form of thin sections, polished thin sections, rock chips, or mineral separates. Such sample aliquots are conceptually accommodated by the data model in the form of a **subsample**. Required subsample attributes include sample name, subsample name, subsample type (thin section, polished thin section, rock chip, or mineral separate), and subsample owner. Subsample owner is listed because it is recognized that a different person may "own" a thin section or mineral separate than the original sample owner.

Images

It is recognized in the design of MetPetDB that most textural information in a metamorphic rock is visual. In addition, the location of analyses within minerals (e.g. core versus rim) is critical to their interpretation and a chemical analysis of a zoned mineral without this knowledge is significantly less valuable than one that is well-located.

Image attributes are shown in Table 5. Typically, an image will be associated with a specific subsample, but if the subsample number is blank, it will be assumed that the image is of the entire sample (e.g. a photograph of the sample). A number of different image types are envisioned, as listed in Table 6, including, but not limited to, photographs, photomicrographs, SEM images (BSE, SE, and CL), and X-ray maps. Additional image types will be added as needed. A scale attribute (image width in mm) is provided if this information is available. Image files always contain headers with information relating to the number of pixels in the X and Y dimensions, and this information can be extracted from the file. The image resolution (size of an individual pixel) can thus be calculated if the full width dimension is provided.

A key aspect of preserving the usefulness of metamorphic data is preservation of the spatial relationship among minerals. For example, is a particular muscovite crystal located in the shear plane or the Q-F domain of a crenulation cleavage? Is a particular monazite located in the core or near the rim of the garnet host? In other words, the spatial relationship of minerals with respect to each other (i.e. the texture) is critical for the interpretation of the chemistry of these phases. To preserve these relationships, MetPetDB will incorporate a subsample grid system whereby all images and analyses will be located with respect to a coordinate system for the individual subsample. It is perhaps simplest to think of this as a GIS (Geographic Information System) for thin sections where the data include images and analyses. The system will include an image viewer that will permit the user to register images with respect to one another and to locate and view analyses with respect to these images. Examples of the image viewer will be given below in the section on user interface.

Some images include more than a single sample as, for example, a photograph of an outcrop or a scanned image of a map on which sample locations are marked. These types of images are designated as *supersample images*. Required attributes for the supersample image (Table 7) include image file name, image type (map or photograph), and sample number (1-n). Optional X-Y coordinates for the sample on the image are provided so that sample positions can be linked directly to supersample images (e.g. a map or outcrop photograph).

Chemical analyses

Chemical analyses form the core of any geochemical database although, as noted above, it is not a requirement that chemical analyses exist for a sample to be entered into MetPetDB.

Required attributes for chemical analyses include (Table 8) (in addition to information about the sample and subsample): (1) mineral name (either identity of mineral or "bulk" for a whole-rock analysis); (2) analytical method; (3) element or species name; (4) units of concentration (e.g. weight percent oxide, ppm, etc.); (5) concentration value; (6) analysis owner. Optional attributes provide information about the analytical facility, analysis date and analyst, publication reference, and supporting analysis information (e.g., spot number, element precision, analysis total).

Two different types of data attributes are included to provide spatial information about an analysis. The "reference image name" refers to an image on which the analysis is located. In some cases this may be a scanned image (e.g. photograph or BSE image) with analysis spots located by hand. If the location of the analysis spot on this reference image is known, then the coordinates can be specified. Finally, it is common on many electron microprobes that the X-Y stage coordinates are saved with the analysis data. The reference frame for these coordinates change every time a sample is reloaded into the microprobe, but for a single session they can provide useful spatial information. For example, the X-Y stage coordinates for points taken on a line traverse provide information about the distance between analyses.

It is noted here the analytical method attribute could be expanded significantly to encompass a series of method-specific sub-attributes and metadata; this topic has been discussed at length (e.g., “Data Reporting in Geochemistry”, an EarthChem workshop held at Lamont-Doherty Earth Observatory in April, 2007). For example, the analytical method attribute = Electron Microprobe (EMP) generates a whole suite of subsidiary attributes relating to machine settings (e.g., diffraction crystal, calibration standard, detector gas), analytical conditions (accelerating voltage, analysis current, analysis time), and post-analysis processing (background fitting, ZAF correction routine). Instrument-specific analytical attributes are, at present, beyond the scope of this data model, but will be incorporated in future versions.

Data Ownership and Sharing

MetPetDB is being designed to fill three major functions: It is intended to be a repository of published data on metamorphic rocks (sample information, chemical analyses, etc.), it is intended to store supplemental, supportive data that compliments published data but is too voluminous to be published by traditional means, and it is intended to be a platform to facilitate the collaboration of research projects among geoscientists around the globe. This section discusses aspects of data permissions and sharing that will be implemented in MetPetDB.

Data Types and Data Ownership

All data housed in MetPetDB will tagged with an owner. Data that are published will be tagged as "public" ownership, which will allow any user to view it. Public data is *immutable*; its content in the database should be a mirror image of the original data in the publication, and, except under special circumstances, by a system administrator, cannot be added to or edited in any way, with the exception of addition of comments.

All other data will be considered "private" and will have a designated owner, which is typically the individual who collects and uploads the data. Private data will have restricted access unless the owner decides to make the data "public", in which case it will have the same access privileges as other public data. Sharing of data with collaborators will be coordinated through the *project manager* (see below).

Ownership of data will be prescribed at the levels of sample, subsample, images, and analyses (see Tables 1, 4, 5, 7, and 8). This level of ownership demarcation was considered necessary because it is easily envisioned that individual A might be the owner (collector) of a sample, whereas individual B is the owner of a thin section from that sample. Individual B might have done a considerable amount of work on this thin section, including publishing a subset of the data (images and analyses) collected. Permitting ownership at the level of subsamples, images, and analyses covers the common situation where only a subset of the data collected on a sample are published (and therefore "public" by default) but the data owner might wish to keep a large amount of unpublished data private.

Scientific Collaboration – the "Project Manager"

A major goal of MetPetDB is to facilitate research collaborations between individuals and research groups anywhere on the planet; such collaborations are likely to involve new and unpublished data that collaborators wish to share with each other, but not yet the general public.

This sharing of data, from one or more samples grouped under a collective research effort between individuals or groups, will be handled through the *project manager*. Collections of samples that pertain to a specific research effort will be called "projects". Projects will have a project chief plus participants, with the chief having responsibility for overseeing the organization of the effort. For example, the project chief will have responsibility for adding (or removing) participants from the project. Projects can contain both public and private data with read/write permissions being granted to project participants by the data owners and handled through the project manager. Data and individuals can belong to multiple projects, but each project can have only one project chief.

Database Implementation and Hardware

MetPetDB's implementation is based upon the Java servlet application programming interface (API); this choice gives us immediate access to the extensive library of open source Java software. In particular, we are interested in leveraging many of the web applications and geographic information systems frameworks that are typically available for the Java platform. PostgreSQL was chosen as the relational database due to its strong adherence to the SQL standards, and its relative ease of administration. The PostGIS extensions to PostgreSQL provided a compelling platform for our GIS needs.

Object metadata for samples, subsamples, images and chemical analyses are stored directly in a PostgreSQL database using a properly normalized relational schema. Most metadata constraints such as required values and valid value ranges are expressed and enforced by the database, with the MetPetDB application automatically mining the constraints out of the database and implementing the same logic within the Java application. This arrangement allows the database to provide complete enforcement of constraints, but also allows the application interface to offer immediate targeted feedback to the user when incorrect data is supplied.
We expect a fairly large collection of images in the system. Accordingly, image files are not stored within PostgeSQL due to the difficulties associated with backing up and restoring multi-gigabyte tablespaces (data files used by the database). Instead, the image files are stored in the UNIX filesystem in separate directories. The location of the file is determined by taking the file content and hashing it to filename using the SHA-1 algorithm. As SHA-1 provides a fairly uniform distribution over filenames, disk arrays can be divided into 256 partitions, using the first byte of the SHA-1 hash to automatically determine the partition that will hold the image file. On a 5 terabyte disk array each partition would be only 20 gigabytes. Backing up and restoring 20 gigabytes is a fairly simple process to manage.

Extensive use of AJAX (Asynchronous JavaScript and XML) through the Google Web Toolkit, (GWT) allows the end user's web browser to efficiently request only the raw data required to update each page instead of reloading the current page after a request. After simple gzip compression has been applied, the size of the transmitted data packet will average well under 1 kilobyte, facilitating easy server retrieval of the relevant object metadata and transfer to the client at wire speed. As the bulk of most per-page processing costs is generally the HTML production (and not the database query processing), pushing this load onto the client browsers reduces the server side costs, making the entire MetPetDB application more responsive to all users.

To ensure high-availability, fault tolerance, and adequate storage for images, we will house the system initially on multiple high-end database servers with adequate storage space (two QSQL Q524 5U 5Tb servers). The two servers will contain identical copies of the database at all times; user requests may then be answered from either server. This will improve system response time, enable large file transfers at a moderate time cost, and provide a back-up system in case of temporary failure of one of the servers. We plan to eventually store the servers in different locations to minimize the downtime caused by power outages and surges. A third 10Tb server will keep incremental updates of the two primary 5Tb servers for up to 30 days. All machines are currently housed in the Computer Science Department Laboratory at Rensselaer. An existing 300Gb server is currently being used as the development platform.

The User Experience

MetPetDB will be accessed through a web portal. The database home page (Fig. 2) has options for (1) login or registering; (2) viewing samples; (3) managing projects; and (4) searching the database. A user wishing to explore a suite of samples will see the requested information displayed in a number of different pages. The list of selected samples (these may be part of a project, a suite found from searching, or a user's entire sample collection) are displayed on the *sample list page*, a page that shows, in addition to the sample number, a list of other sample attributes (Fig. 3). For each page a default set of attributes will be displayed, and these will be fully configurable by each individual user. Selecting a sample brings up the *sample page* (Fig. 4), which displays general characteristics of a sample, as well as a sample image and a list of the subsamples. Selecting a subsample brings up the *subsample page* with configurable attributes about available images and analyses displayed (Fig. 5).

Chemical analyses may be of a mineral or the bulk rock. A list of the minerals for which analyses are available is provided and the user can choose to view the entire set of analyses for a particular mineral (Fig. 6), or to download the analyses to a spreadsheet format. Note that one attribute for chemical analyses is the reference image, which is the image on which the spot chemical analysis is located, and analysis locations can be viewed on these images as well.

Images for a subsample can be viewed in two different ways. The *image list* page provides details about every image available for a subsample, which can be viewed in either list format or as thumbnails (Fig. 7). Individual images can be examined in pop-up windows.

Alternatively, all images can be displayed on the *image viewer (*Figure 8). The image viewer is a module designed to display any or all of the images in a subsample in their correct spatial positions with respect to each other. The image viewer incorporates the concept of a subsample map (or image grid) onto which images can be placed. Each image can be scaled and rotated in order to locate them in their proper positions. Transparency can be adjusted to aid in registration and the entire image map can be viewed at a wide range of scales. The purpose of developing the image viewer is to preserve the textural setting of all data, which is an essential part of the interpretative process.

Figure 8 shows an example of an image map from a sub-sample (thin section) from sample V6B, a garnet-biotite-sillimanite migmatite from the Valhalla Complex, British

Columbia (Spear and Parrish, 1996). At wide zoom, one BSE image and three photomicrographs are located and scaled with respect to the thin section image (Fig. 8a). Figure 8b shows an enlargement of a part of the thin section image with a Ca X-Ray map overlay. Analysis spots are shown in their proper location on the sample in Figure 8c. In fully annotated samples, the analysis spots will be linked directly to the analyses in the database.

Searching the database

MetPetDB will be searchable by nearly all of the data attributes including words found in comment fields. Searchable items include (but are not limited to) rock type, minerals present, metamorphic grade, latitude and longitude, region name, collector, publication, availability of images (e.g. Ca X-ray map), availability of analyses (e.g. samples with scapolite analyses), and composition range of analyses (e.g. $2.5 < MgO < 7.5$).

Mineral-delimited searches will be hierarchical using a mineral classification tree. This approach will ensure that selecting "chain silicate" will search on all pyroxenes and amphiboles but that selecting "glaucophane" will find only those samples in which glaucophane has been listed. Users will be able to refine searches so that only the desired information is returned. Once a search is completed, results will be viewable by any of the viewing pages described above, or plotted on a map view such as Google Maps or Google Earth.

Managing Projects

The web portal will allow sample owners to create and manage projects. Managing a project includes selecting samples to be included in the project and collaborators to be a part of the project.

Evolution of and future directions for MPDB

The future of MetPetDB (or any other geological database) depends on its usefulness to the geologic community. We believe that a well-designed and comprehensive database for metamorphic geochemistry will be compelling to the geologic community, but a number of significant issues need to be addressed before this becomes a reality. In particular, the database must be populated with sufficient high quality data and searches return useful information. Second, there must be significant value added to compel individuals to put forth the effort that will be required to help populate the database. Third, there needs to be seamless integration with other databases so that individuals can find the information they seek without concern for where those data reside. Finally, there need to be demonstrated scientific problems that can only be solved by large collections of data and advanced search algorithms.

Populating the database

MetPetDB will only be as useful as the quantity and quality of data it contains. As mentioned earlier, it is estimated that less than one percent of metamorphic data are published, and the question must be addressed how the other >99% of data will be incorporated into the database.

In general, data will be added to MetPetDB in two ways. Individual samples can be added to the database via web pages designed for that purpose. Bulk uploading of sample and image attributes, images, and chemical analyses will also be implemented. It is anticipated that users will upload their own data, and it remains to be seen whether the large volume of data that currently resides on individual computers hard drives and in file cabinets will eventually be accessible to everyone. It is our hope that over time these data will find their way into the database as researchers and students find it desirable to do so. For example, an excellent place for a student to start a project in a new area is to become aware of all previous work that has been done in the region. If information on a region resides in the student's advisors file cabinets, this becomes an excellent opportunity to further populate the database. Once the backlog of existing data is uploaded, populating the database with newly collected data should become routine. Eventually, it is anticipated that direct uploading from analytical instruments (e.g. electron microprobes) may be possible so that users will need to expend only a minor amount of effort.

Published data will be digitized and incorporated into the database by the database working group during the initial project phase. This effort has already begun and, as of this writing, we have mined approximately 120 articles, dating between 1962-2005. Journals currently represented include American Journal of Science, The American Mineralogist, Contributions to Mineralogy and Petrology, Journal of Petrology, and Journal of Metamorphic Geology. Although the initial effort to add the wealth of published information will be large, it is hoped that at some point publication in an established journal will include addition of the relevant data to the database, thus ensuring the database is up to date with the most current published works.

MetPetDB user toolkit

A compelling reason to upload one's own data into MetPetDB, in addition to facilitating collaborations with colleagues, will be the ability to make use of the extensive toolkit that will be developed. It is planned to develop codes to do common tasks such as composition plotting and mineral formula recalculations, including ones requiring stoichiometric constraints to estimate ferric iron content (e.g. pyroxenes and amphiboles). The toolkit will contain a complete set of published thermobarometry routines (e.g. Spear et al., 1991) that can be used on any database analyses and it is planned to provide links to facilitate approaches such as the "average P-T calculation" approach (e.g. Powell and Holland, 1994; Berman, 1991). Additional analysis tools will include software for age calculations (from U, Th, Pb) measurements, pseudosection construction, and contour plotting of P-T diagrams.

Future plans for MetPetDB include the incorporation of derivative or interpretative data. These terms refer to the fact that some very useful information about metamorphic rocks is only obtained after significant recalculation, analysis, and interpretation of the basic composition and image data. Information such as P-T conditions, maximum temperature or pressure, the P-T path, the age of chemical domains of monazite, or the timing of crenulation development all fit the category of requiring significant interpretation of raw data. Such information is in a very real sense the desired result of many metamorphic studies, and would be of great use to researchers in a wide range of fields outside of metamorphic petrology. We do not yet have a satisfactory data model for these types of information, but it is planned to incorporate such derivative data into MetPetDB in the future.

Integration with other databases

Researchers searching for information are not concerned with where the retrieved information is stored, but they are concerned that they find all of the information relevant to their study. MetPetDB will be integrated with other databases of the EarthChem consortium (**www.EarthChem.org**), which includes PetDB, NAVDAT, and GeoRoc. Eventually, as other databases come on line, these will be integrated as well.

New Directions

We envision that MetPetDB will be useful for museum curators, university, and secondary school teachers. A significant potential for exploring the geologic world will exist through databases such as MetPetDB and it is our intention to include materials on the web portal that will assist guest users who may not be versed in geochemistry to find useful and interesting information.

The potential also exists for MetPetDB to alter basic research methodologies within the metamorphic petrology community. Hypothesis testing will be possible, in many cases, without the time and financial expense required to plan and execute field work, collect, process, and analyze samples. If the quality of the data in the database is maintained at a high level, the researcher essentially enters the project at the interpretation stage, which could, in many cases, dramatically lower the time and cost of research and allow significant new directions to be explored.

Additionally, future data mining algorithms may allow researchers to explore aspects of the global geochemistry that are currently impossible. For example, consider a hypothesis that a particular type of Ca zoning in garnet signifies a particular type of tectonic evolution. There might be several dozen Ca zoning maps of garnet in the entire metamorphic literature today, but there might be many thousands in a well-populated database. Examining these by hand would be prohibitive, but future searching might incorporate image recognition algorithms that would permit a query such as "find all of the samples with Ca zoning in garnet that demonstrates a core to rim increase of 20% or more". Sophisticated queries that require large databases are already being demonstrated by users of the igneous geochemistry databases (e.g., Spiegelman and Kelemen, 2003; Salters and Stracke, 2004; Hirschmann et al., 2003) and we anticipate similar applications of MetPetDB.

Acknowledgements

Members of the Metamorphic Petrology Database working group are thanked for their input and feedback over the course of this project. We would also like to thank Jennifer Rodger, Barbara Firebaugh, Samuel Broadaway, and Julia Cosgrove for archived data extraction and programming error checking routines. John Schumacher is thanked for his design of the MPDB logo. This work is funded by NSF grant EAR-0622345.

References Cited

Berman, R.G. (1991), Thermobarometry using multi-equilibrium calculations; a new technique, with petrological applications, *Can. Mineral.*, 29, 833-855.

Hirschmann, M.M., Kogiso, T., Baker, M.B., and Stolper, E.M. (2003), Alkalic Magmas generated by partial melting of garnet pyroxenite, *Geology*, 31, 481-484.

Lehnert, K., Su, Y., Langmuir, C.H., Sarbas, B., and Nohl, U. (2000), A global geochemical database structure for rocks, *G-cubed*, 1(5).

Powell, R., and Holland, T. (1994), Optimal geothermometry and geobarometry, *Amer. Mineral.*, 79, 120-133.

Salters, V.J., and Stracke, A. (2004), Use of the PetDB database for calculating the composition of the MORB-source, *Eos Trans. AGU*, 85 (47), Fall Meeting Supplement, Abstract SF32A-03.

Spear, F.S., Peacock, S.M., Kohn, M.J., Florence, F.P., and Menard, T. (1991), Computer programs for petrologic P-T-t path calculations. *Amer. Mineral.*, 76, 2009-2012.

Spear, F. S. and R. Parrish (1996), Petrology and petrologic cooling rates of the Valhalla Complex, British Columbia, *Can. Jour. Petrol.*, 37, 733-765.

Spiegelman, M., and Kelemen, P.B. (2003), Extreme chemical variability as a consequence of channelized melt transport, *G-cubed*, 4(7), 1-18.

Figure Captions

Figure 1. Schematic of data model for MetPetDB. Boxes highlight data attributes and lines show relationships between objects.

Figure 2. Mock-up of MetPetDB home page. Major user operations include managing sample collections, managing projects, and searching the database.

Figure 3. Mock-up of web page showing sample listing. Attributes shown in this and other pages are configurable by the user. Major options include sorting the listing, adding or editing sample information, and adding samples to projects.

Figure 4. Mock-up of Sample Detail page.

Figure 5. Mock-up of Subsample Detail page providing a summary of the information available on the specific subsample.

Figure 6. Mock-up of the Analysis Detail page. The information shown on this page can be user-configured. Analysis will also be made available for download.

Figure 7. Mock-up of Image Detail page for the specified subsample. Enlarged versions of each image can be viewed by clicking on the image.

Figure 8. Mock-up of Image Viewer. (A) Full view of a polished thin section. Background image is a thin section scan. Also shown are several photomicrographs and one BSE image. Boxes show location of (B) and (C). (B) Enlargement of part of the Image Viewer showing overlay of a Ca X-ray map. The transparency of each image can be adjusted. (C) Enlargement of a part of the BSE image showing location of analysis spots. Analysis spots for which a reference image and reference image coordinates are specified can be plotted automatically on the grid.

Table 1. Sample attributes

Sample attribute	Required (R) or optional (O)	Number of possible attributes	Comments
Sample number	R	1	
IGSN	O	1	International Geo Sample Number
Sample alias	O	0-n	Provides for multiple names for the same sample
Longitude (DD.DDD)	R	1	
Longitude error (DD.DDD)	O	1	
Latitude	R	1	
Latitude error (DD.DDD)	O	1	
Rock type	R	1	Rock names given in Table 2
Region	O	0-n	Geographic descriptor
Country	O	0-1	
Collector	O	0-1	
Collection date	O	0-1	
Present sample location	O	0-1	
Comment	O	0-n	Multiple comments possible
Reference	O	0-n	Archived data requires a reference
Metamorphic facies	O	0-n	Multiple facies allowed
Minerals present	O	0-n	
Mineral modes	O	0-n	
Sample owner	R	1	

Table 2. MetPetDB recognized rock names

Slate
Phyllite
Schist
Gneiss
Migmatite
Anatectite
Marble
Calc-silicate
Greenschist
Amphibolite
Blueschist
Eclogite
Granofels
Hornfels
Skarn
Quartzite
Jadeitite
Glaucophanite
Serpentinite
Garnetite
Pyroxenite
Mylonite
Cataclasite
Metapelite
Metaarkose
Metagreywacke
Metabasite
Metacarbonate
Metagranite

Table 3. MetPetDB recognized metamorphic grade names

Zeolite
Prehnite-pumpellyite
Greenschist
Amphibolite
Epidote amphibolite
Granulite
Blueschist
Eclogite
Hornfels

Chlorite zone
Biotite zone
Garnet zone
Staurolite zone
Staurolite-kyanite zone
Kyanite zone
Sillimanite zone
Andalusite zone
Sillimanite-K feldspar zone
Garnet-cordierite zone
Migmatite zone

Ultra high pressure
Ultra high temperature

Table 4. Subsample attributes

Attribute	Required (R) or optional (O)	Number of possible attributes	Comments
Sample number	R	1	IGSN, if available
Subsample number	R	1	e.s. SS01, SS02 etc.
Subsample type	R	1	Types include thin section, polished thin section, rock chip, mineral separate
Subsample owner	R	1	Subsamples may be owned by individuals different than the sample owner

Table 5. Image attributes

Image attribute	Required (R) or optional (O)	Number of possible attributes	Comments
Sample number	R	1	
Subsample number	O	0-1	e.g. SS02 or blank if it is a whole sample image
Image file name	R	1	
Image type	R	1	See Table 6
Image scale	O	1	Image width in mm
Owner	R	1	
Comments	O	0-n	

Table 6. Image types and attributes

Image type	Comments
Photograph	General sample imagery
Thin section scan	Special image of entire thin section
Photomicrograph-Transmitted Plane Polarized	
Photomicrograph-Transmitted Crossed Polars	
Photomicrograph-Reflected Plane Polarized	
Photomicrograph-Reflected Crossed Polars	
Secondary Electron Image	SE
Back-Scattered Electron Image	BSE
Cathodoluminescence Image	CL
X-ray Element Distribution Map	
Element name (R)	Element name is required for X-ray maps

Table 7. Supersample image attributes

Attribute	Required (R) or optional (O)	Number of possible attributes	Comments
Image file name	R	1	
Image type	R	1	Map, photograph
Collector	O	0-1	
Comments	O	0-n	Description of images, keywords, etc.
Sample number	R	1-n	Sample to which supersample image refers
Sample location X	O	1	Location of sample on image
Sample location Y	O	1	
Owner			

Table 8. Analysis attributes

Attribute	Required (R) or optional (O)	Number of possible attributes	Comments
Sample number	R	1	
Subsample number	R	1	
Mineral name	R	1	"Bulk" for a whole rock analysis
Analytical method	R	1	e.g. EMP, LA-ICPMS, SIMS
Location of analytical facility	O	0-1	
Reference	O	0-n	Published data require a reference
Analysis date	O	0-1	
Analyst	O	0-1	
Spot number	O	0-1	Unique spot number identifier
Reference image name	O	0-1	Image name or file name on which analysis spot is located
Image X coordinate	O	0-1	Location of spot on reference image
Image Y coordinate	O	0-1	Location of spot on reference image
X stage coordinate of analysis	O	1	Stage location of spot analysis
Y stage coordinate of analysis	O	1	Stage location of spot analysis
Element or species name	R	1-n	e.g. Si or SiO_2
Value	R		
Units	R		
Precision	O		
Analysis total	O	0-1	
Analysis owner	R	1	
Comment	O	0-n	

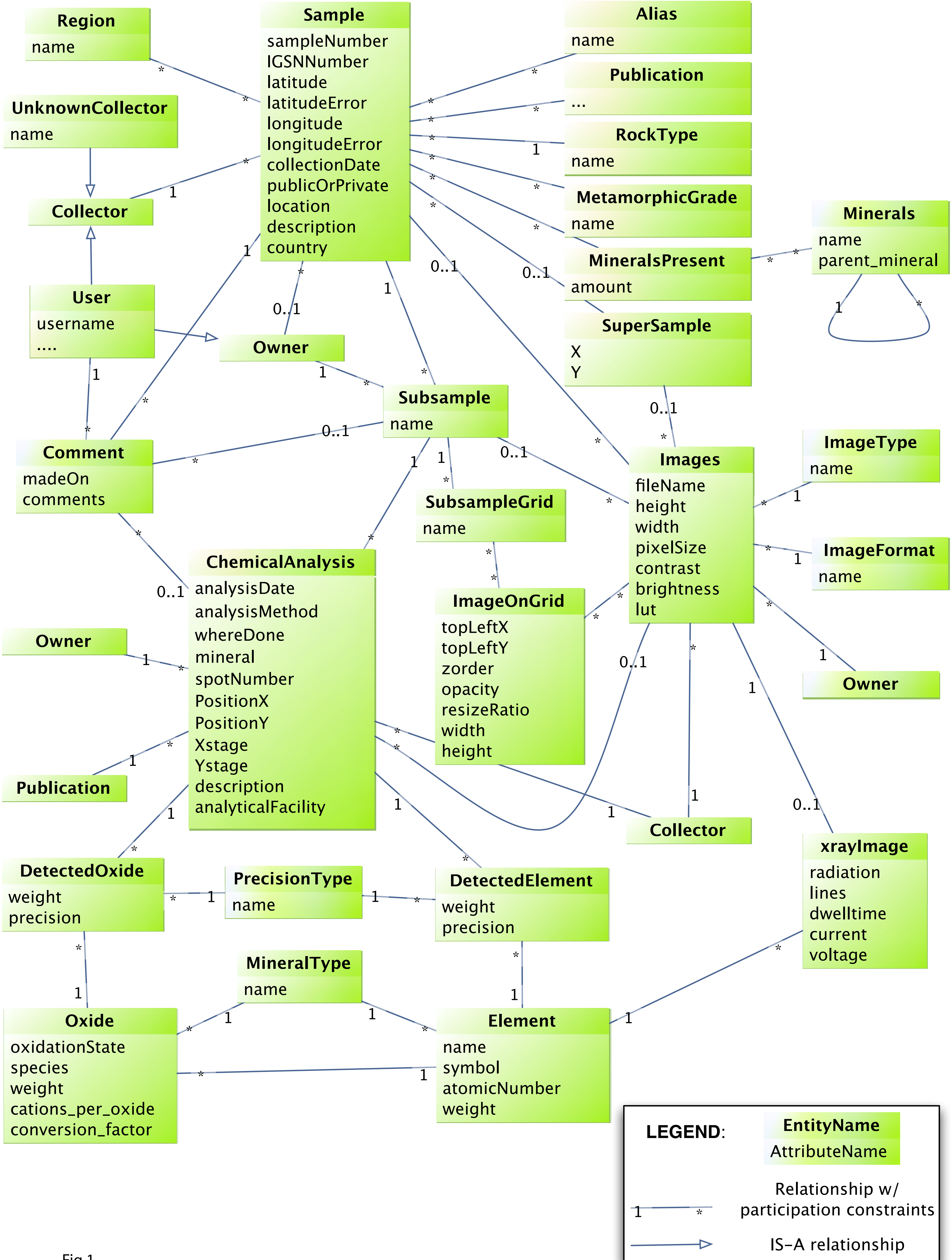

Region
name
Sample
sampleNumber
IGSNNumber
latitude
latitudeError
longitude
longitudeError
collectionDate
publicOrPrivate
location
description
country
Alias
name
Publication
...
RockType
name
MetamorphicGrade
name
Minerals
name
parent_mineral
MineralsPresent
amount
SuperSample
X
Y
UnknownCollector
name
Collector
User
username
....
Owner
Subsample
name
Comment
madeOn
comments
SubsampleGrid
name
Images
fileName
height
width
pixelSize
contrast
brightness
lut
ImageType
name
ImageFormat
name
ChemicalAnalysis
analysisDate
analysisMethod
whereDone
mineral
spotNumber
PositionX
PositionY
Xstage
Ystage
description
analyticalFacility
ImageOnGrid
topLeftX
topLeftY
zorder
opacity
resizeRatio
width
height
Owner
Owner
Publication
Collector
xrayImage
radiation
lines
dwelltime
current
voltage
DetectedOxide
weight
precision
PrecisionType
name
DetectedElement
weight
precision
MineralType
name
Oxide
oxidationState
species
weight
cations_per_oxide
conversion_factor
Element
name
symbol
atomicNumber
weight
LEGEND:
EntityName
AttributeName
Relationship w/ participation constraints
IS-A relationship

Fig 1

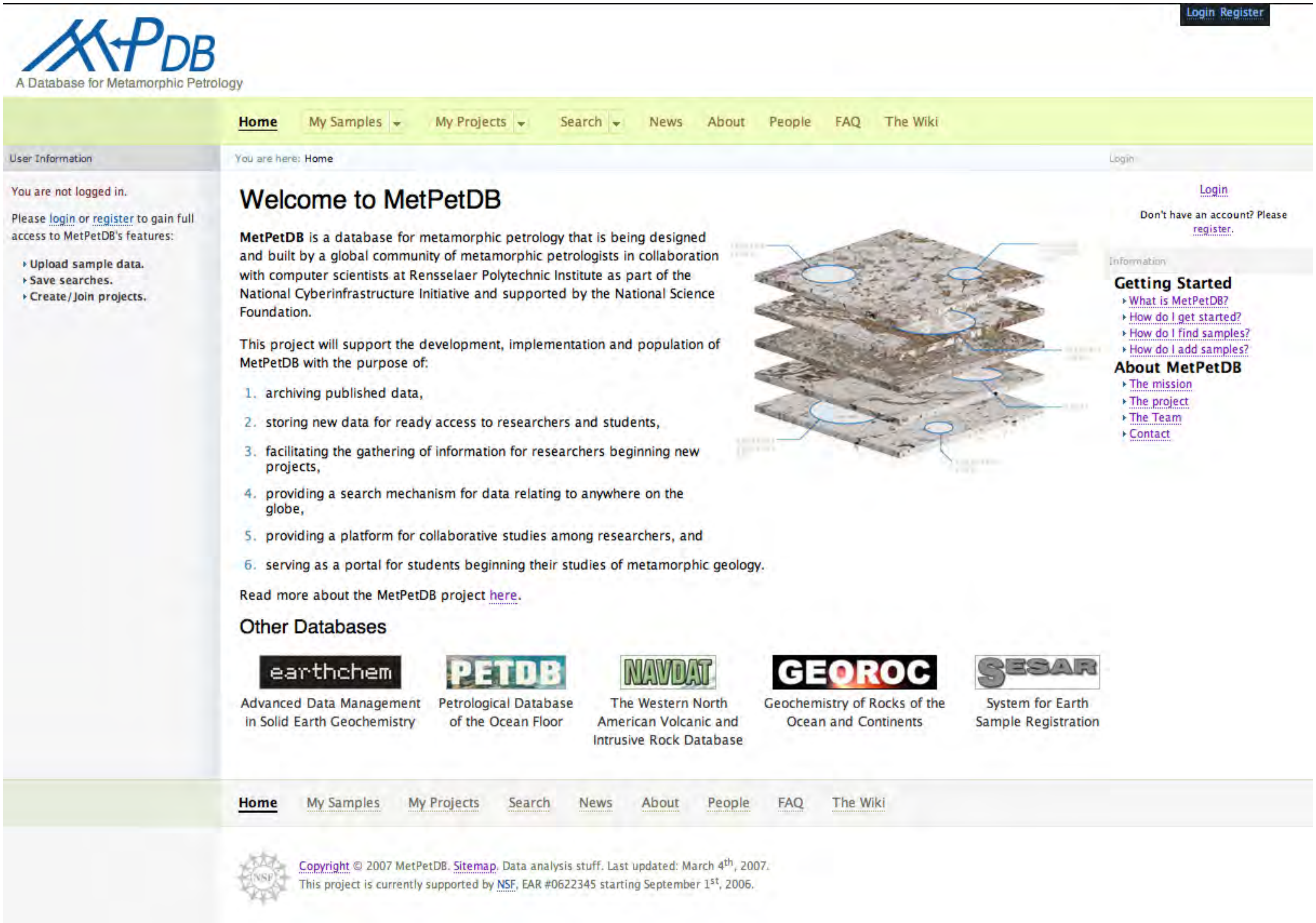

Login Register
MPDB
A Database for Metamorphic Petrology
Home
My Samples
My Projects
Search
News
About
People
FAQ
The Wiki
User Information
You are here: Home
Login
You are not logged in.
Please login or register to gain full access to MetPetDB's features:
Upload sample data.
Save searches.
Create/Join projects.
Welcome to MetPetDB
MetPetDB is a database for metamorphic petrology that is being designed and built by a global community of metamorphic petrologists in collaboration with computer scientists at Rensselaer Polytechnic Institute as part of the National Cyberinfrastructure Initiative and supported by the National Science Foundation.
This project will support the development, implementation and population of MetPetDB with the purpose of:
1. archiving published data,
2. storing new data for ready access to researchers and students,
3. facilitating the gathering of information for researchers beginning new projects,
4. providing a search mechanism for data relating to anywhere on the globe,
5. providing a platform for collaborative studies among researchers, and
6. serving as a portal for students beginning their studies of metamorphic geology.
Read more about the MetPetDB project here.
Login
Don't have an account? Please register.
Information
Getting Started
What is MetPetDB?
How do I get started?
How do I find samples?
How do I add samples?
About MetPetDB
The mission
The project
The Team
Contact
Other Databases
earthchem
Advanced Data Management in Solid Earth Geochemistry
PETDB
Petrological Database of the Ocean Floor
NAVDAT
The Western North American Volcanic and Intrusive Rock Database
GEOROC
Geochemistry of Rocks of the Ocean and Continents
SESAR
System for Earth Sample Registration
Home
My Samples
My Projects
Search
News
About
People
FAQ
The Wiki
NSF
Copyright © 2007 MetPetDB. Sitemap. Data analysis stuff. Last updated: March 4th, 2007.
This project is currently supported by NSF, EAR #0622345 starting September 1st, 2006.

Fig 2

You are here: Home > My Samples > All My Samples

My Samples

All Newest Favorites

Upload Sample Bulk Upload

Quick Filters: Recently Added in MyProj

Change View: Simple Detailed | Create New View

	Sample	Minerals Present	Region	Collector	Rock Type
☐	V6-B	Garnet, Biotite, Sillimanite, Plagioclase, Quartz, Ilmenite	Valhalla Complex Map	Frank S. Spear	Metapelite
☐	80-D	Garnet, Biotite, Muscovite, Plagioclase, Quartz	Mt. Moosilauke, NH Map	Joe Pyle	Metapelite
☐	90-A	Staurolite, Muscovite, Plagioclase, Biotite, Andalusite, Quartz	Mt. Moosilauke, NH Map	Sibel Adali	Metapelite
☐	92-D	Garnet, Biotite, Muscovite, Quartz, Plagioclase, Chloritoid	Mt. Moosilauke, NH Map	Boleslaw Szymanski	Metapelite
☐	33-X	Biotite, Chlorite, Hornblende, Plagioclase, Ilmenite	Buck Mtn., MT Map	Anthony Waters	Amphibolite

☐ Select All Add to 'My First Project' Apply to Selected

Fig 3

You are here: Home > My Samples > V6-B > Details

V6-B

Details

Edit

Sample Description
Sillimanite + garnet + biotite migmatite from Gwillam Creek shear zone. Large porphyroblast in photo is plagioclase. Leucosomes contain garnet crystals formed from dehydration melting reactions biotite + sillimanite = garnet + melt

IGSN	651435842
Alias	V6-B
Region	Valhalla Complex
Latitude	49.57320
Longitude	-117.62792
Collector	Frank S. Spear
Date of Collection	1991-09
Rock Type	Metapelite
Field Comments	Well-banded migmatites within the Gwillam Creek shear zone
Metamorphic Grade	Migmatite Zone
Minerals Present	Garnet, Biotite, Sillimanite, Plagioclase, Quartz, Ilmenite
Publication References	Spear & Parrish, 1996, JOP

Subsamples

Subsample	Type	Images	Image Map	Mineral Analyses	Added
SS001	Thin Section	4	Map	0	July 25, 2007
SS002	Thin Section	7	Map	0	July 20, 2007
SS003	Polished Thin Section	17	Map	16	July 14, 2007
SS004	Rock Chip	14	Map	10	July 11, 2007
SS005	Mineral Separate	0		1	

Comments

No comments yet. Leave a comment.

Fig 4

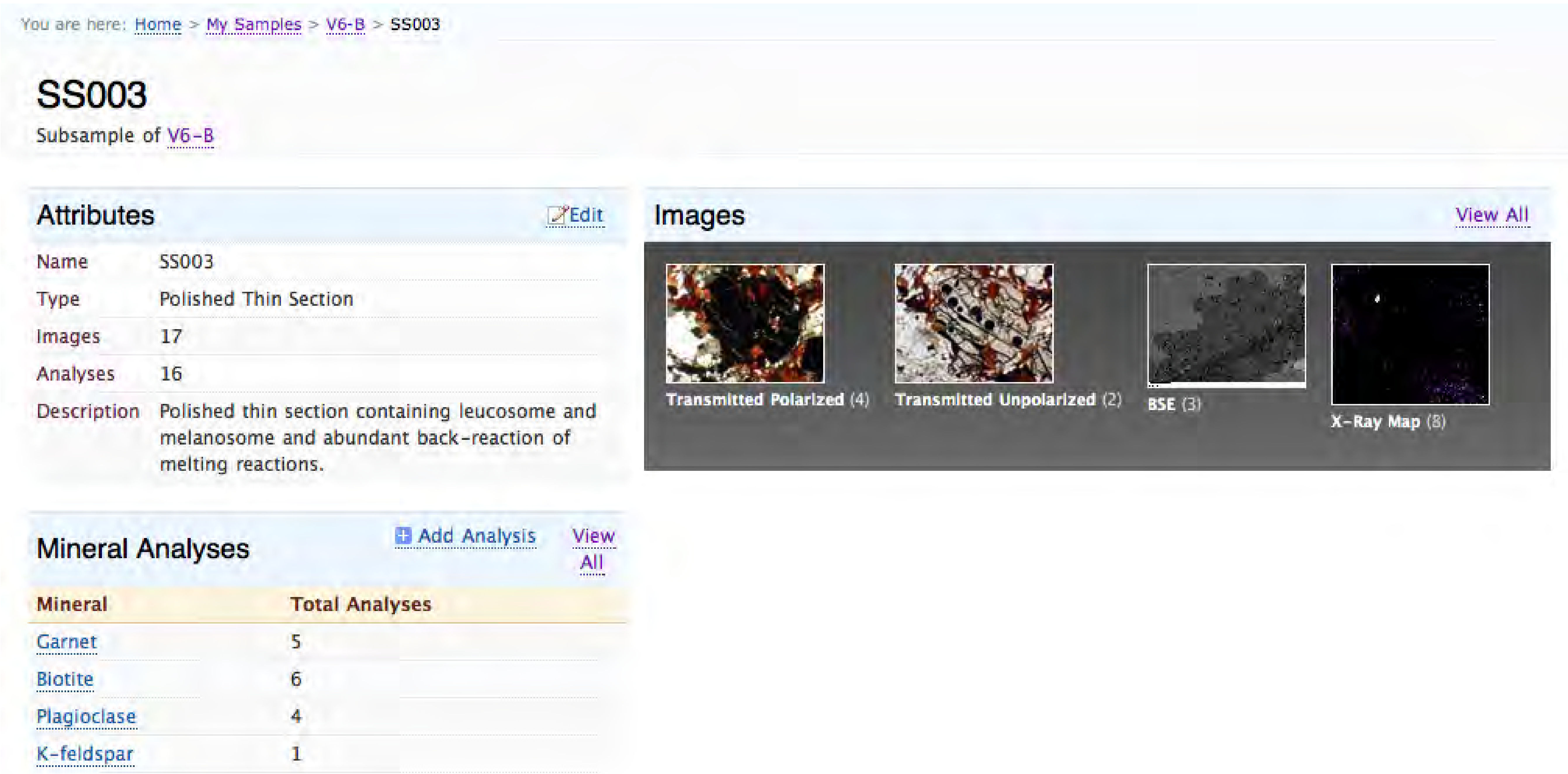

You are here: Home > My Samples > V6-B > SS003
SS003
Subsample of V6-B
Attributes
Edit
Name SS003
Type Polished Thin Section
Images 17
Analyses 16
Description Polished thin section containing leucosome and melanosome and abundant back-reaction of melting reactions.
Images
View All
Transmitted Polarized (4)
Transmitted Unpolarized (2)
BSE (3)
X-Ray Map (8)
Mineral Analyses
Add Analysis
View All
Mineral Total Analyses
Garnet 5
Biotite 6
Plagioclase 4
K-feldspar 1

Fig 5

You are here: Home > My Samples > V6-B > SS003 > Images

SS003 Mineral Analyses

Analyses of SS003

Add Analysis | Change View Options

Analyses

View All

Point	Mineral	Analyst Date	Method	Ref. Image	SiO_2	Al_2O_3	TiO_2	Fe_2O_3	MgO	FeO	MnO	CaO	Na_2O	K_2O	H_2O	Total
59	Garnet	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	38.77	22.18	0.00	0.00	6.11	31.71	0.67	2.16	0.00	0.00	0.00	101.60
60	Garnet	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	38.99	22.14	0.00	0.00	6.15	31.32	0.68	2.21	0.00	0.00	0.00	101.49
61	Garnet	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	38.76	22.12	0.00	0.00	6.14	31.47	0.68	2.21	0.00	0.00	0.00	101.38
63	Garnet	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	38.76	22.18	0.00	0.00	6.32	31.12	0.65	2.11	0.00	0.00	0.00	101.15
64	Garnet	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	37.48	27.05	0.00	0.00	6.67	28.04	0.54	1.84	0.00	0.00	0.00	101.64
71	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	35.88	20.15	1.60	0.00	9.69	20.13	0.06	0.02	0.59	5.36	0.00	93.50
72	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	37.12	20.78	2.10	0.00	10.07	19.14	0.05	0.00	0.19	5.42	0.00	94.88
73	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	38.21	20.64	3.05	0.00	12.28	16.05	0.03	0.00	0.22	5.56	0.00	96.05
78	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	36.75	20.42	3.10	0.00	9.63	18.60	0.05	0.00	0.19	5.45	0.00	94.18
79	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	37.06	19.68	4.22	0.00	9.55	19.14	0.05	0.00	0.16	5.58	0.00	95.43
80	Biotite	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	36.98	19.83	4.19	0.00	9.37	18.78	0.10	0.00	0.16	5.67	0.00	95.07
90	Plagioclase	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	61.21	25.88	0.00	0.00	0.00	0.27	0.00	6.99	5.88	0.13	0.00	100.35
91	Plagioclase	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	61.12	25.90	0.00	0.00	0.00	0.27	0.00	7.03	5.79	0.12	0.00	100.23
92	Plagioclase	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	60.97	25.45	0.00	0.00	0.00	0.08	0.00	6.57	5.91	0.14	0.00	99.11
93	Plagioclase	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	70.54	20.33	0.00	0.00	0.00	0.11	0.00	0.59	8.31	0.11	0.00	99.99
94	K-feldspar	F. S. Spear 01/11/94	EMP	V6b_melt_3.jpg	69.07	20.24	0.00	0.00	0.00	0.08	0.00	0.11	1.33	8.89	0.00	99.71

Fig 6

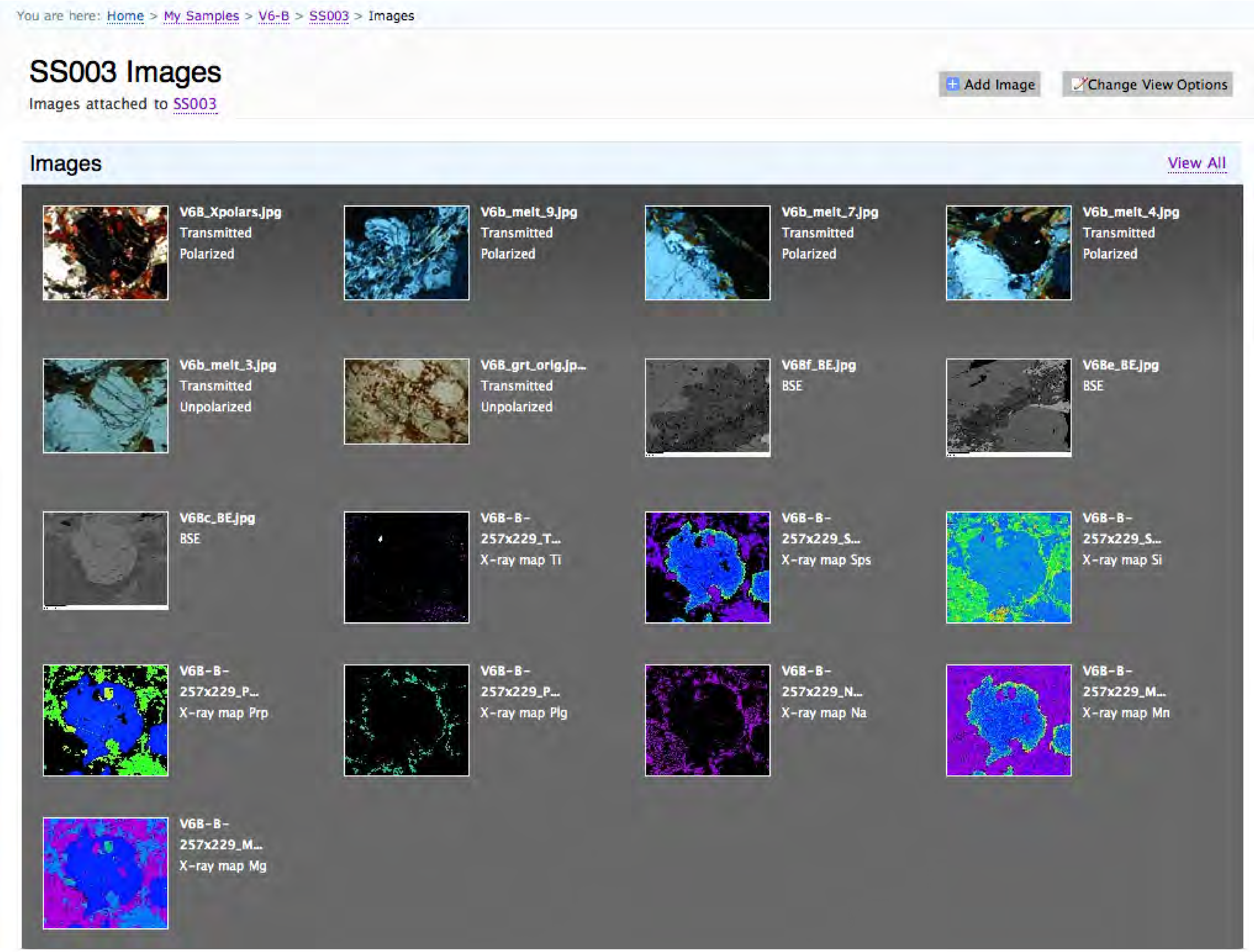
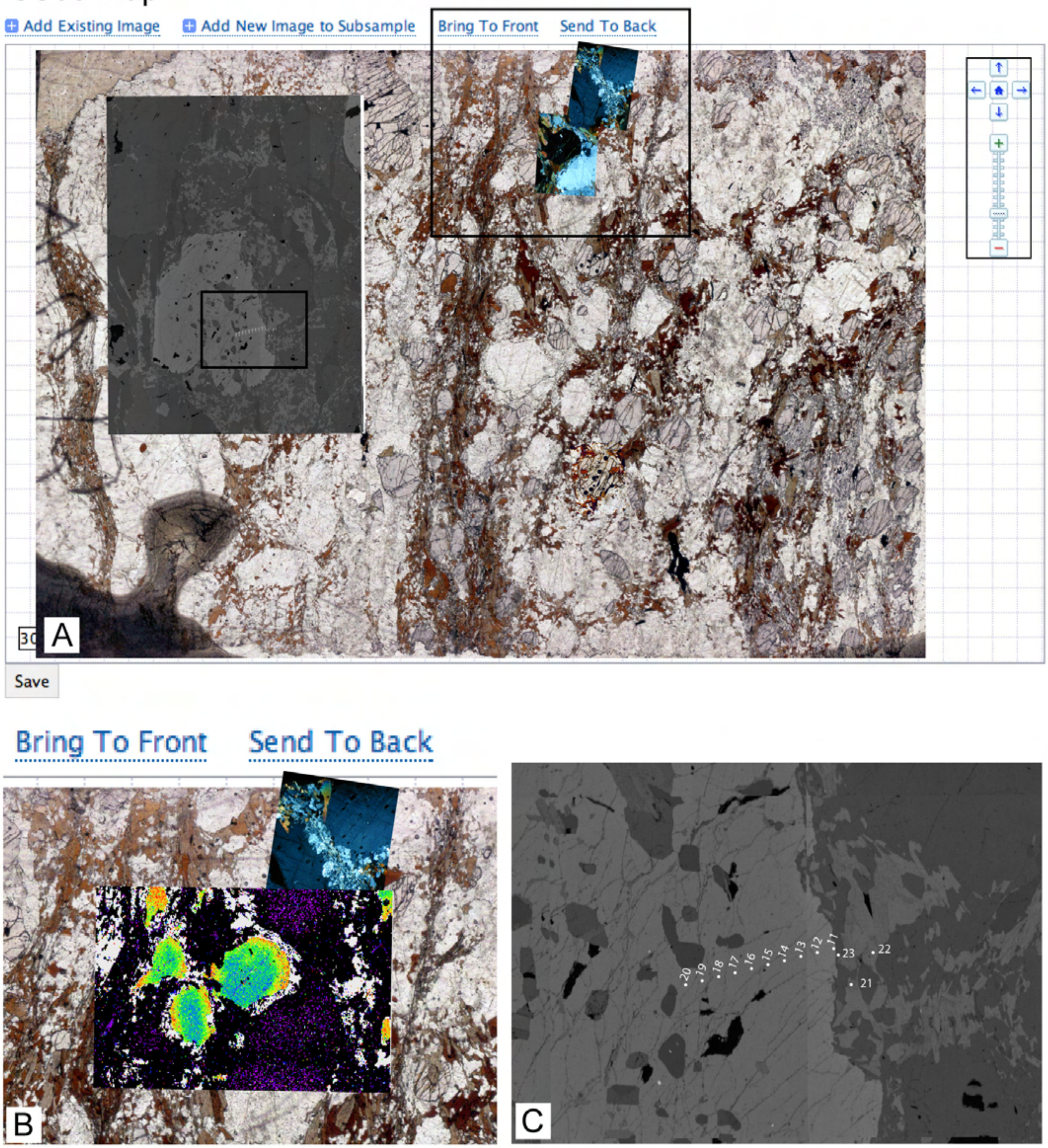

You are here: Home > My Samples > V6-B > SS003 > Images
SS003 Images
Images attached to SS003
Add Image
Change View Options
Images
View All
V6B_Xpolars.jpg
Transmitted
Polarized
V6b_melt_9.jpg
Transmitted
Polarized
V6b_melt_7.jpg
Transmitted
Polarized
V6b_melt_4.jpg
Transmitted
Polarized
V6b_melt_3.jpg
Transmitted
Unpolarized
V6B_grt_orig.jp...
Transmitted
Unpolarized
V6Bf_BE.jpg
BSE
V6Be_BE.jpg
BSE
V6Bc_BE.jpg
BSE
V6B-B-257x229_T...
X-ray map Ti
V6B-B-257x229_S...
X-ray map Sps
V6B-B-257x229_S...
X-ray map Si
V6B-B-257x229_P...
X-ray map Prp
V6B-B-257x229_P...
X-ray map Plg
V6B-B-257x229_N...
X-ray map Na
V6B-B-257x229_M...
X-ray map Mn
V6B-B-257x229_M...
X-ray map Mg

Fig 7

V6B

SS03 Map



Add Existing Image
Add New Image to Subsample
Bring To Front
Send To Back
A
Save
Bring To Front
Send To Back
B
C
20
19
18
17
16
15
14
13
12
11
23
22
21