Statistical-based software for the prediction of mutagenicity
Introduction to Sarah

Sarah Nexus is a statistical-based methodology for the prediction of mutagenicity.

Drawing on over 30 years of experience, Lhasa scientists developed Sarah with input from the Food and Drug Administration (FDA) under a Research Collaboration Agreement. Sarah uses a unique machine learning methodology to build a statistical model for Ames mutagenicity. This facilitates the identification of potentially mutagenic impurities in line with the ICH M7 guideline. The learning algorithm uses a novel, self-organising hypothesis network (SOHN) (Hanser et al. 2014). This novel system for the fragmentation of query structures enables greater transparency and adds interpretability of predictions in order to assist the expert review required by the ICH M7 guideline.

Sarah and Derek Nexus (the preferred expert rule-based system for the prediction of toxicity (Dobo et al. 2012, Sutter et al. 2013)), in combination, provide you with the means to meet the computational, toxicological assessment requirements of the ICH M7 guideline from one intuitive interface. Assessing your potentially mutagenic impurities quickly and easily and submitting those results to regulators, reduces the need for time consuming and expensive in vitro tests. Both Derek and Sarah have been designed independently to meet the requirements of ICH M7, and both systems can be run from within the same Nexus interface to help simplify your workflow.

WHY CHOOSE LHASA?

• Over 30 years of experience in developing state-of-the-art in silico tools, including Meteor Nexus, Derek Nexus and Wit Nexus, improving workflow efficiency.

• Transparency of Lhasa systems allows trust and confidence in the science presented.

• Software is easy to use and well supported.

• Over 30 years of experience in developing state-of-the-art in silico prediction and database systems.

• All science is developed in-house, providing the opportunity to discuss directly with Lhasa expert scientists.

Accuracy of predictions is maintained across the entire chemical space thus minimising problematic “Out of Domain” predictions pose a significant problem when assessing impurities under the ICH M7 guideline as two negative predictions from complementary methodologies are required to conclude that an impurity is of no mutagenic concern. “Out of Domain” predictions are of little value when conducting expert review and add additional complexity to this process. Sarah’s large, high quality training set ensures maximum coverage of chemical space thus minimising problematic “Out of Domains”.

Sarah Nexus has been extensively evaluated using pharmaceutical proprietary data sets and delivers high performance (Barber et al. 2019).

Features of the Model Builder:

• Enhanced Predictivity
  Sarah’s intuitive model builder enables you to rapidly expand the model to deliver predictions tailored to your proprietary chemical space.

• Maximise Model Accuracy
  Custom models can be tested and subsequently fine-tuned using the in-built external validation functionality.

• Reduced False Strengths
  New data added to a model, or used to create a brand new model is processed and all structures standardised. Conflicting or duplicated data, which may affect signal strengths, is then removed.

• Reduced Bias in Small Datasets
  Integrating the Lhasa developed Sarah model by adding your own data to create a new model reduces the bias arising in smaller datasets.

The model builder feature means users can duplicate the Lhasa model and supplement it with their own data, or build an entirely new model. These custom models can then be exported and shared with other people.

Features of the Model Builder:

• Auto Classification
  A user editable ICH M7 classification is automatically derived from the predictions provided by Derek and Sarah and any relevant experimental information from the Carcinogenicity Potency Database and Lhasa certified Ames data.

• One Interface, Multiple Predictions
  Using Sarah within the Nexus platform can give you direct access to other Lhasa in silico tools, including Meteor Nexus, Derek Nexus and Wit Nexus, improving workflow efficiency.

• User Control, Multiple Predictions
  Sarah has default prediction settings that have been designed to ensure the best balance between sensitivity and specificity for use under the ICH M7 guideline. However, it is recognised that users may want to change this approach for their particular need and Sarah’s intuitive prediction setup options allow you to tailor the prediction accordingly.

• Transparent Predictions
  Predictions are clearly represented and supported by a measure of confidence, the fragments on which the predictions are based and relevant examples from the Sarah training set ordered by structural similarity to the query compound. This high level of transparency facilitates the expert review process.

• Rapid Mutagenicity Assessment
  Sarah can swiftly provide single or batch predictions for the mutagenicity of query compounds.

• Extensive Coverage of Chemical Space
  Sarah predictions are accepted by regulators under the ICH M7 guideline. The M7 prediction functionality allows for simultaneous compound processing in Sarah and Derek against the mutagenicity in vitro endpoint, fulfilling both the expert rule-based and statistical-based predictions required under ICH M7.

• Expert ICH M7 Support
  Sarah predictions are accepted by regulators under the ICH M7 guideline. The M7 prediction functionality allows for simultaneous compound processing in Sarah and Derek against the mutagenicity in vitro endpoint, fulfilling both the expert rule-based and statistical-based predictions required under ICH M7.

• Customisable Reports
  Sarah incorporates a reporting framework that allows (.doc .pdf .xlsx and .sdf) file export. Report templates are fully customisable by the end user ensuring that you can provide the right information at the right time.
Who are we?
Lhasa Limited is an active, scientific research organisation working in chemistry and the life sciences. Our areas of specialisation include the development of software for toxicology and metabolism prediction and data management.
Working closely with Lhasa Limited members and the broader scientific community, Lhasa continues to draw on over 30 years of experience to create user-friendly, state-of-the-art in silico prediction and database systems.

What makes us special?
Lhasa Limited was founded on the basis of data and knowledge sharing. Building on its reputation as the eponymous ‘honest broker’, Lhasa Limited has continued to work on the basis of ‘shared knowledge, shared progress’ for more than 30 years. Over this time, Lhasa has regularly been trusted with proprietary data and this can be seen in the many successful consortia that we continue to be involved in. The sustained success of such consortia is demonstrative of how working with Lhasa has a positive impact on the research and development process of its members.

What do we offer?

**Derek Nexus™**
An expert rule-based system or the prediction of toxicology.

**Meteor Nexus™**
An expert rule-based system for the prediction of metabolic fate.

**Mirabilis™**
A tool for assessing the relative purging of mutagenic impurities.

**Vitic Nexus™**
A chemical database and information management system.

**Zeneth™**
An expert rule-based system for the prediction of degradation pathways.

Other Activities
Lhasa Limited has various collaborations including elemental impurities, intermediates and Derek in StarDrop™. Visit our website to find out more.

References