

RDML – neuer Standard für qPCR-Daten

Dr. Andreas Untergasser, Universität Wageningen, Niederlande; Steve Lefever, Dr. Filip Pattyn, Dr. Jan Hellemans und Prof. Dr. Jo Vandesompele, Universität Gent, Belgien

Das RDML-Konsortium wurde mit dem Ziel gegründet, ein universelles Format für real-time PCR-Daten zu entwickeln: die Real-time PCR Data Markup Language, kurz: RDML. Den Anstoß dazu hatte die Erfahrung gegeben, dass es schwierig ist, aus verschiedenen Laboren, Softwarepaketen sowie Analysewerkzeugen stammende qPCR-Daten auszutauschen. Das RDML-Konsortium hat RDML als einen universellen Standard entwickelt, der genügend Informationen enthält, um den Entwurf eines Experiments zu verstehen, die Daten erneut zu analysieren und die Ergebnisse zu deuten. Aufgrund der XML-Basis des Standards sind alle Daten in einer selbstdokumentierenden Form gespeichert, was ihre zukünftige Verwertbarkeit garantiert und eine einfache Integration in existierende Software ermöglicht. Wir sind sicher, dass RDML die Entwicklung innovativer Software und Analysemethoden anstoßen wird.

Die quantitative PCR (qPCR) ist heute eine weitverbreitete Methode zur Messung der Genexpression und zur Quantifizierung von Genkopien. Sie ist in medizinischen Diagnoseverfahren zur Bestimmung viraler oder bakterieller Belastungen nicht zu ersetzen. Doch die Fülle der Auswahl an PCR-Maschinen unterschiedlicher Hersteller und Anwendungsoptionen ist nicht unproblematisch. Denn die Vielfalt der Maschinen spiegelt sich auch in den Softwarepaketen wider, die mit den Geräten geliefert werden. Jeder Hersteller vertreibt seine eigene Version und verwendet ein jeweils proprietäres Format, um die Daten zu speichern. Diese Datenformate sind miteinander weitestgehend inkompatibel. Daher ist es derzeit kaum möglich, Daten austauschen, wenn nicht auf der gleichen Instrumentenplattform gearbeitet wird. Da nur eine kleine Zahl von Wissenschaftlern die Originaldaten lesen kann, macht es keinen Sinn, sie zu publizieren. RDML wurde entwickelt, um ein echtes offenes Format zu etablieren. Die von Jan Hellemans und Jo Vandesompele auf dem

2nd International qPCR-Symposium (September 5-9, 2005, Freising, Deutschland) ins Leben gerufene RDML-Initiative war auch unlängst auf dem 3rd International qPCR-Symposium in Weihenstephan ein viel diskutiertes Thema.

Das RDML-Konsortium

Jan Hellemans und Jo Vandesompele von der Universität Gent, Belgien gründeten im vergangenen Jahr die auf Real-time PCR spezialisierte Datenanalyse-Firma Biogazelle, (www.biogazelle.com) aus. Filip Pattyn (Universität Gent, Belgien; Entwickler und Administrator der Real-time PCR-Primer and Probe Database <http://medgen.ugent.be/rtprimerdb/>) sowie Andreas Untergasser (Wageningen University, Niederlande; Entwickler von primer3plus und beteiligt in der Weiterentwicklung von primer3) stießen 2007 dazu. Im vergangenen Jahr wurde schließlich das internationale RDML-Konsortium gegründet mit dem Ziel, das RDML-Datenformat weiterzuentwickeln.

Interessierte Personen oder Firmen können der Organisation jederzeit ohne Kosten beitreten (www.rdml.org). Das RDML-Konsortium ist in eine Kernentwicklergruppe, eine Teilnehmergruppe sowie Unterstützer aufgeteilt. Aufgabe der Kernentwicklergruppe ist es, Kommentare und Anfragen innerhalb des Konsortiums zu diskutieren und sie in neuen RDML-Empfehlungen zu vereinen. Diese Empfehlungen durchlaufen vier Reifegrade: Arbeitskonzept, Empfehlungskandidat, vorgeschlagene Empfehlung und schließlich Empfehlung.

Das RDML-Format

Nur die RDML-Empfehlung ist bindend. Bis sie veröffentlicht ist, können sich alle Details des vorgeschlagenen Standards noch ändern. Das RDML-Format erlaubt einfachen Austausch annotierter Rohdaten zwischen verschiedenen Laboratorien. Es ist ein Textfile in Extensible Markup Language (XML), genannt RDML mit der Datei-Erweiterung *.rdml oder *.rdm. Das Format ist unabhängig von Computerhardware, Betriebssystem oder dem Softwarepaket und kann falls nötig künftig erweitert werden. Die Details des Standards werden in Kürze publiziert werden, die endgültige RDML-Empfehlung wird spätestens im Sommer 2008 verfügbar sein.

RDML wird es ermöglichen, qPCR-Daten in wissenschaftliche Veröffentlichungen einzufügen und es Lesern sowie Gutachtern erlauben, Originaldaten unabhängig zu beurteilen, ähnlich den MIAME-Richtlinien für Microarray Experimente². Das RDML-Konsortium möchte zudem die Entwicklung von Software unterstützen. Steve Lefever arbeitet momentan an einem Konvertierungswerkzeug, welches es erlauben wird, RDML-Dateien zu erstellen. Des Weiteren ist ein Validierungswerkzeug verfügbar, das es gestattet zu überprüfen, ob die RDML-Dateien mit dem Standard kompatibel sind. Weitere Werkzeuge werden entwickelt werden und unter Open Source-Lizenzen verfügbar sein, inklusive einer Bibliothek, um das Dateiformat lesen/schreiben zu können und einem Programm für einfache Datenanalyse.

Literatur

- [1] submitted
- [2] Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M., Nat Genet. 2001 Dec;29(4):365-71.

Korrespondenzadresse

Dr. Andreas Untergasser
Universität Wageningen
Tel.: +31 6 23073797
Fax.: +31 317 483584
andreas@untergasser.de

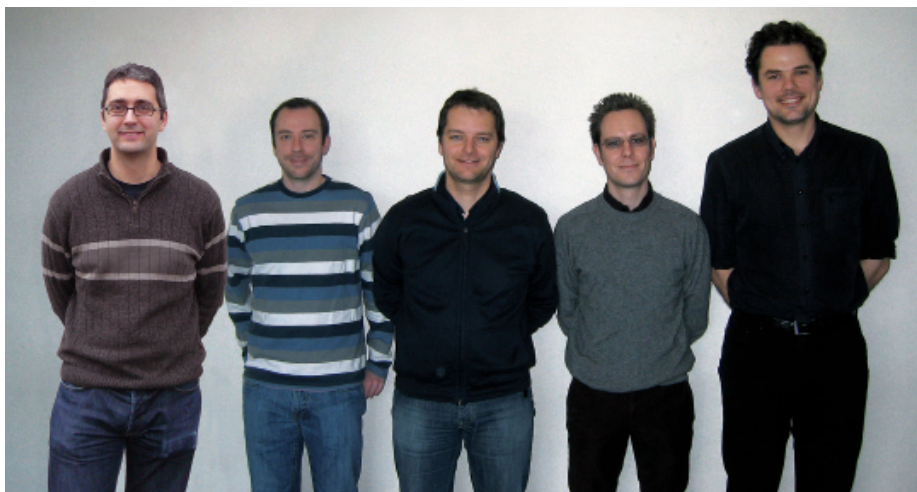


Abb: Die RDML-Kernentwicklergruppe: Andreas Untergasser, Steve Lefever, Jo Vandesompele, Jan Hellemans and Filip Pattyn