NSW GOVERNMENT DATA QUALITY STATEMENT: 08 OCTOBER 2026

Name of dataset or data source:	Genomic data and DNA sequences used for genetic mapping of Stb19 in an Australian bread wheat diversity panel
Custodian of the dataset or data source:	TBD
Description:	This dataset contains genotyping data for wheat, including 11,200 SNPs generated via DArT and KASP markers. It supports genetic mapping of Stb19, a newly identified resistance gene against Zymoseptoria tritici. The data enable reproducible analysis of Stb19 inheritance, marker-trait associations, and breeding applications for disease resistance. The associated publication for these data is "Yang, N. N., McDonald, M. C., Solomon, P. S., & Milgate, A. W. (2018). Genetic mapping of Stb19, a new resistance gene to Zymoseptoria tritici in wheat. Theoretical and Applied Genetics, 131(12), 2765-2773. https://doi.org/10.1007/s00122-018-3189-0"
Data quality rating:	★Institutional Environment - 5 ★Accuracy - 5 ★Coherence - 5 ★Interpretability - 5 ☆Accessibility - 2

INSTITUTIONAL ENVIRONMENT — Excellent ★

✔ DPI is the recognised custodian for this dataset
✔ The data is collected and managed according to a data quality framework

i.e. are there rules, processes and checks in place to ensure high data quality?(Policy, Information asset governance, standards)

✔ Data collection is authorised by law, regulation or agreement
✔ The Custodial agency has no commercial interest or conflict of interest in the data

ACCURACY — Excellent ★

✔ Data has been subject to a quality assurance process (e.g. checking for errors at each stage of data collection and processing, or verifying data entry and making corrections if necessary.)
✔ A revision policy exists for the dataset (i.e if errors are identified in data)
✔ There are no known gaps in the data or if there are gaps (for example: non-responses, missing records, data not collected), they have been identified in caveats attached to the dataset.
✔ No changes have been made or other factors identified that could impact the validity of data (e.g weighting, rounding, de-identification of data, changes or flaws in data collection or verification methods).

The adjustments, changes/factors have been identified in caveats and are attached to the asset.

✔ The data collection meets the objectives of the primary user. i.e.The data correctly represents what it was designed to measure, monitor or report.

ℹ Find out more about the quality assurance processes from the NSW Government Standard for Data Quality Reporting. https://www.finance.nsw.gov.au/ict/resources/data-quality-standard

COHERENCE — Excellent ★

✔ Standard definitions, common concepts, classifications and data recording practices have been used.
✔ Elements within the data can be meaningfully compared.
✔ This data is generally consistent with similar or related data sources from the same discipline
✔ The data can be analysed over time (for example, there have not been any significant changes in the way items are defined, classified or counted over time).
✔ The data does not form part of a collection or, if it is the latest in a series of data releases, there have not been any changes in methodology or external impacts since the last data release.

INTERPRETABILITY — Excellent ★

✔ A data dictionary is available to explain the meaning of data elements, their origin, format and relationships
✔ Information is available about the primary data sources and methods of data collection (e.g. instruments, forms, instructions).
✔ Information is available to help users evaluate the accuracy of the data and any level of error
✔ Information is available to explain concepts, help users correctly interpret the data and understand how it can be used
✔ Information is available to explain ambiguous or technical terms used in the data

ℹ Find out more about the data dictionary from the Custodian (contact details below).
ℹ Find out more about the primary data sources and methods of data collection from the Custodian (contact details below).
ℹ Find out more about concepts used in this dataset and how to understand or interpret the data from the Custodian (contact details below).
ℹ Find out more about ambiguous or technical terms used in the data from the Custodian (contact details below).

ACCESSIBILITY — Fair ☆

✔ Data is available in machine-processable, structured form (e.g. CSV format instead of an image scan of a table)
✔ Data is available in a non-proprietary format (e.g. CSV, XML)

✗ Data is available online with an open licence
✗ Data is described using open standards that facilitates interoperability & data exchange (e.g. RDF, SPARQL) and persistent identifiers (a long lasting reference like Digital Object Identifiers)
✗ Data is linked to other data, to provide context (e.g. employee ID is linked to employee name or species name is linked to genus)

DATA DISCLAIMER

You must check and comply with the licensing conditions for the information you wish to use. This may require you to contact the Office of DPI, or other custodial agency, or the third party copyright owner for permission to use the material. You may also use any material in accordance with rights you may have under the Copyright Act 1968 (Cth), for example under the fair dealing provisions or statutory licences. Use of material in a way not permitted by this copyright notice may be an infringement of copyright. Infringing copyright may expose you to legal action by, and liability to, the copyright owner. Wherever a third party holds copyright in material, the copyright remains with that party. Their permission may be required to use the material and you should contact that party directly. As far as practicable, material for which the copyright is owned by a third party will be clearly labelled. Excluded material can only be used under the specific terms of use attached to that material. If you want to use this material in a manner that is not covered by those specific terms of use, you must request permission from the copyright owner of the material.

DPI endeavours to make sure that information provided is correct at the time of its publication. However, as necessary you should obtain independent advice before making any decision based on the information. The information is made available on the understanding that custodial agencies and the State of NSW accept no responsibility for any damage, cost, loss or expense incurred by you as a result of:

- any error, omission or misrepresentation in the information provided
- without limiting the above, any delay, failure or error in recording, displaying or updating information, including but not limited to, data relating to credit holdings.

Custodial agencies and the State of New South Wales disclaim all responsibility and all liability (including without limitation, liability in negligence) for all expenses, losses, damages and costs you might incur as a result of the information being inaccurate or incomplete in any way, and for any reason.

For more information about this dataset or data source, contact:	DPI

Understanding the Data Quality Statement

The data quality statement aims to help you understand how a particular dataset could be used and whether it can be compared with other, similar datasets. It provides a description of the characteristics of the data to help you decide whether the data will be fit for your specific purpose.

About the quality rating:

The reporting questionnaire asks five questions for each of these data quality dimensions:

- Institutional Environment
- Accuracy
- Coherence
- Interpretability
- Accessibility

For each question: “yes” = 1 point; “no” = 0 points
The number of points determines the Quality Level for each dimension (high, medium, low).
Only dimensions with four or five points receive a star.

Points	Quality Level	Star / No Star
0	Poor	No Star
1	Poor	No Star
2	Fair	No Star
3	Good	No Star
4	Very Good	Star
5	Excellent	Star

Evaluating data quality

Quality relates to the data’s “fitness for purpose”. Users can make different assessments about the dataquality of the same data, depending on their “purpose” or the way they plan to use the data.
The following questions may help you evaluate data quality for your requirements. This list is not exhaustive.Generate your own questions to assess data quality according to your specific needs and environment.

- What was the primary purpose or aim for collecting the data?
- How well does the coverage (and exclusions) match your needs?
- How useful are these data at small levels of geography?
- Does the population presented by the data match your needs?
- To what extent does the method of data collection seem appropriate for the information being gathered?
- Have standard classifications (eg industry or occupation classifications) been used in the collection of the data?If not, why? Does this affect the ability to compare or bring together data from different sources?
- Have rates and percentages been calculated consistently throughout the data?
- Is there a time difference between your reference period, and the reference period of the data?
- What is the gap of time between the reference period (when the data were collected) and the release date of thedata?
- Will there be subsequent surveys or data collection exercises for this topic?
- Are there likely to be updates or revisions to the data after official release?