

PREP database: Extension to Genomic Form

Anahit Nazari-Ghadikolaei¹

¹ Interbull Centre, Department of Animal Biosciences (HBIO), SLU- Box 7023, S-75007 Uppsala, Sweden

E-mail address: anahit.nazari@slu.se

Abstract

The Performance Recording, Evaluation and Publication (PREP) database is an online platform developed by Interbull Centre under the umbrella of the Centre's activities as the European Union Reference Centre (EURC) for Zootechnics, specifically designed for breed associations and National Genetic Evaluation Centres in order to submit and share descriptive information for a broad range of breeds and traits for both dairy and beef cattle. The main purposes of PREP include transparency, comparison and harmonization of information collected and used at the national level. To date, there are different electronic forms available to collect descriptive information for both dairy and beef cattle regarding *conventional* genetic evaluation for several trait groups. Next steps would be to expand PREP, so to be able to also collect descriptive information for *genomic* information. To do this, different approaches, and their level of pros and cons, were considered leading to identify a slightly different form's structure, compared to the one currently applied for conventional evaluations, as the most convenient for recording genomic information. The new form's structure will have all breeds and trait groups in one form, rather than having different forms per trait groups, as specific general traits' information regarding, for example, trait definition, method of recording, heritability etc. will be already available in the dedicated conventional forms. Moreover, the selected approach will allow to copy/paste information across different breeds/trait groups, which often do appear to be very similar (if not identical) when dealing with genomic evaluations. In summary, the new electronic genomic form will be considered as the general form for users to provide descriptive information related to their national genomic evaluations.'

Key words: PREP, harmonization, genomic form, national genomic evaluation, EURC

Introduction

The Performance Recording, Evaluation and Publication database (PREP) was developed, and is hosted, at the Interbull Centre within the context of the Centre's function as the European Union Reference Centre (EURC) for Zootechnics (Bovine breeding) with the aim of promoting transparency of methodologies applied, and easier comparisons of best practices (via a standardized set of options available for each question), all in all leading to harmonization and/or improvement of the methods of performance testing and/or genetic evaluation applied.

Regarding traits' harmonization, trait correlations play an important role in the quality

of the estimations for international evaluations. On the other hand, harmonization of traits helps to improve correlations in order to achieve more accurate international and national evaluations. In this regard, extracting and reviewing information for each individual trait from PREP can lead the given correlated traits to International Committee for Animal Recording (ICAR)-Interbull guidelines in order to improve the across-country compatibility of traits. To date, the ICAR-Interbull guideline for trait harmonization has been published for Calving traits in 2022 (https://interbull.org/ib/eu_detailed_technical_reports), and Fertility traits harmonization has been approved in 2024 and will be included in the ICAR-Interbull guideline in the near future.

PREP is an online platform available for all breed societies and National Genetic Evaluation Centres (NGECs) beyond their direct involvement with any of the international evaluations currently offered by Interbull Centre.

Benefits of PREP

PREP has several benefits over document-based records:

a) Provides easier harmonization and standardization of information by comparing evaluation methods, trait definitions etc. across countries, breeds and traits in an easier and more efficient way

b) Represents a common database accessible to all cattle breeding organisations and third parties (both inside and outside Europe) including NGECs, researchers and competent authorities regardless of their direct involvement in any of the Interbull Centre's International evaluations for dairy (Interbull) and/or beef (Interbeef)

c) Collected information is freely available to the world wide web, either by directly accessing of specific country-breed-trait forms or through ad-hoc query options that can be easily defined by the user under “*Submission*” and “*Data queries*” tabs.

Current and upcoming forms on PREP

Several conventional genetic forms are at the moment available within PREP for different trait groups, both for dairy and beef cattle (<http://prep.interbull.org>). Available forms cover trait groups such as production, calving, female fertility and conformation traits for dairy cattle and adjusted weaning weight, calving and carcass traits for beef cattle. Throughout the course of this year, PREP will be expanded with new forms for

udder health, longevity and workability traits for dairy cattle.

In addition to the above-mentioned forms, PREP also includes an “Other traits” form, aimed at collecting information regarding a wide range of traits and breeds recorded at the national level but not yet evaluated internationally. The relevance of such information is twofold: 1) Provide an overview of novel traits recorded in different countries and the status of their evaluation: implemented or still in a research phase; evaluated via a conventional or a genomic model (single step or two-step approach). 2) Explore the opportunity to identify new traits suitable for international evaluations, and the related challenges (i.e. new phenotype recording) for such traits. Examples of traits currently collected through the “Other traits” form are claw health, metabolic diseases (such as milk fever, clinical and sub-clinical ketosis), gestation length, and feed efficiency.

Overview of the structure of the current forms

Each of the conventional electronic forms currently available in PREP covers a specific trait group. Each form is made up by four different sections, with the first (and main) section allowing to select the different breeds/traits (within the specific trait group) the form will deal with, as shown in Figure 1.

The second section collects all general information regarding each trait selected. This section includes information referring to, for example, trait definition, method of recording, trait heritability, genetic variance, and data edits, as shown in Figure 2. Each set of questions is equipped with a list of standardized options that can be chosen from.

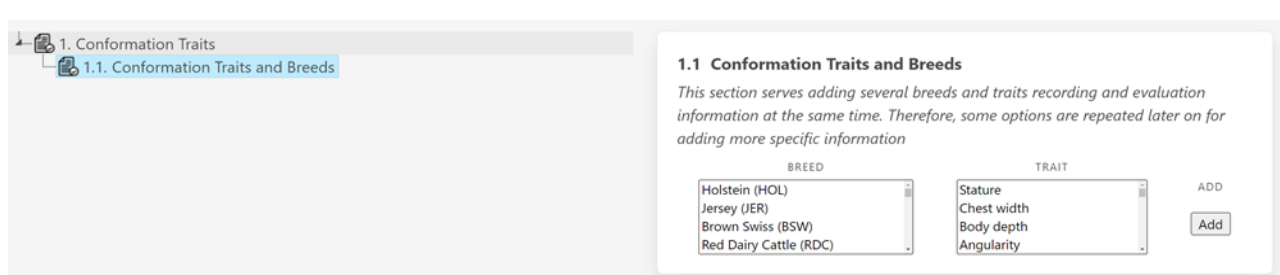


Figure 1: Section related to breed trait combinations of the current conventional form on PREP

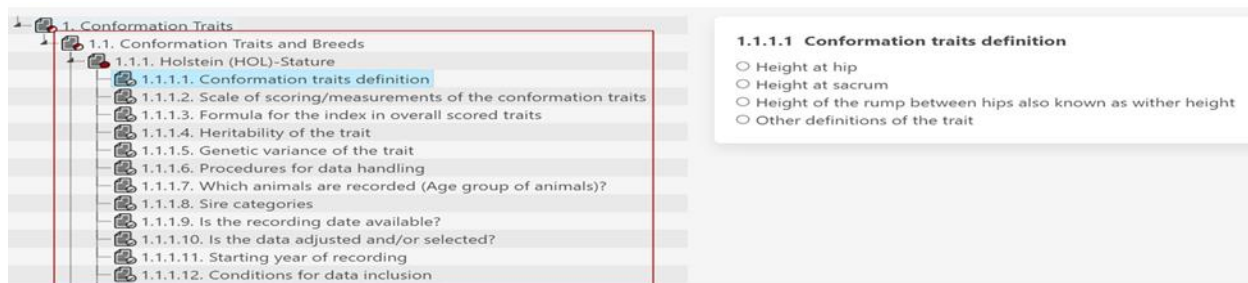


Figure 2: Section related to the general information for each individual trait of the current conventional form on PREP

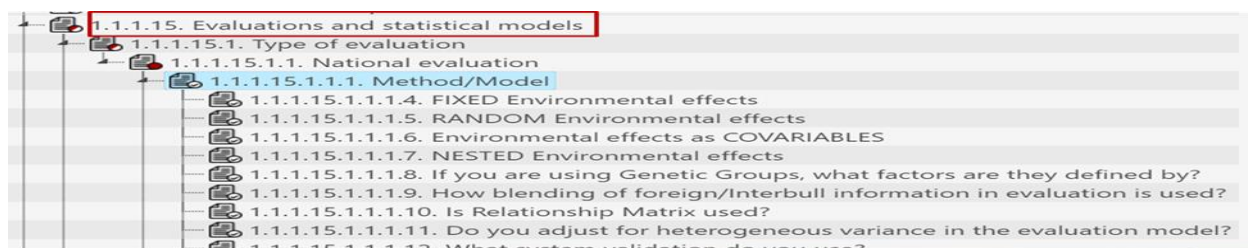


Figure 3: Section related to the evaluations and statistical models of the current conventional forms on PREP

The third section collects information regarding the evaluation methods and statistical models applied. After choosing the type of evaluation (national or international), questions regarding this section will focus for example on the model applied (i.e. Multi-breed (MB) or Multi-trait (MT), BLUP or Animal model), fixed and random effects used and publication criteria. An example of this section is shown in Figure 3. Database's tools will allow the user to copy and paste (and modify) the answers provided to one type of evaluation/breed to the others available within the same form.

The fourth and last section collects information regarding the scientific base, such as the scientific references that have been used for reliability or validation methodologies, as shown in Figure 4.

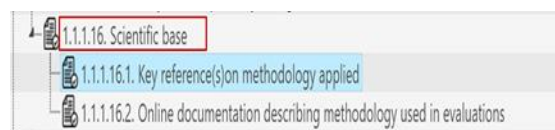


Figure 4: Section related to the scientific references of the current conventional form on PREP

Collecting genomic information via PREP

To complete the collection of descriptive information, specific forms need to be created in order to also collect information regarding genomic evaluations. Several approaches have been considered by the Interbull Centre. The approach described in this paper was the one found as the most efficient for collecting general genomic information in PREP.

Structure of the general genomic form

The structure of the genomic form remains similar to the current conventional form, but with questions and type of answers being more genomic-oriented. The new genomic form will consist of three main sections:

a) *Genomic information*: such as SNP-chip used, method for imputation and reference population.

b) *Genomic methods and models*: for example single-step genomic evaluation (ssGBLUP), bayesian or polygenic models.

c) *Genomic reliability and system validation*: including publication criteria, genomic reference base, scientific base etc. will also be part of the genomic form.

The new genomic form will therefore have one section less compared to the conventional form: the section reporting general information regarding the trait definition, method of performance recording, data edits, heritability etc., in fact, will not be part of the new genomic form as such information will already have been provided within the related conventional (genetic) form and will therefore already be available in PREP.

Another main difference will be within section 1: While in the conventional form section 1 will allow users to select breeds and individual traits,

within the trait group the form was referring to (i.e. production, conformation, calving), section 1 of the genomic form will allow selecting *breed-trait group* combinations. This means there will be only one form for all available breed-trait group combinations (Figure 5).

Advantages of the general genomic form

The chosen approach has several advantages: It would be more efficient and easier for users to fill in only one form; as the trait-specific information has previously been provided in the PREP conventional forms, there is no need for repeating the same information for each trait or trait group. All information provided for one breed-trait group can be easily and rapidly copied across the remaining breed-trait groups selected. For example, information provided for production traits could be copied to calving or female fertility traits, or from Holstein to Jersey breed (https://interbull.org/ib/prep_user_manual), as shown in Figure 6. Should the copied information need some adjustments to make them fit properly the specific breed-trait groups they have been pasted into, such information can be easily modified. For example, if the definition of the reference population differs between the Holstein and Jersey breed, the information can be edited as shown in Figure 7.

The screenshot shows a web interface for '1.1 Trait groups and Breeds'. It features two dropdown menus: 'BREED' and 'TRAIT GROUPS'. The 'BREED' menu lists: Holstein (HOL), Jersey (JER), Brown Swiss (BSW), and Red Dairy Cattle (RDC). The 'TRAIT GROUPS' menu lists: Production, Calving, Conformation, and Female fertility. An 'ADD' button is located to the right of the 'TRAIT GROUPS' menu. Below the menus, there is a text description: 'This section serves adding several breeds and traits group genomic information and evaluation at the same time.'

Figure 5: Structure of the proposed genomic form on PREP

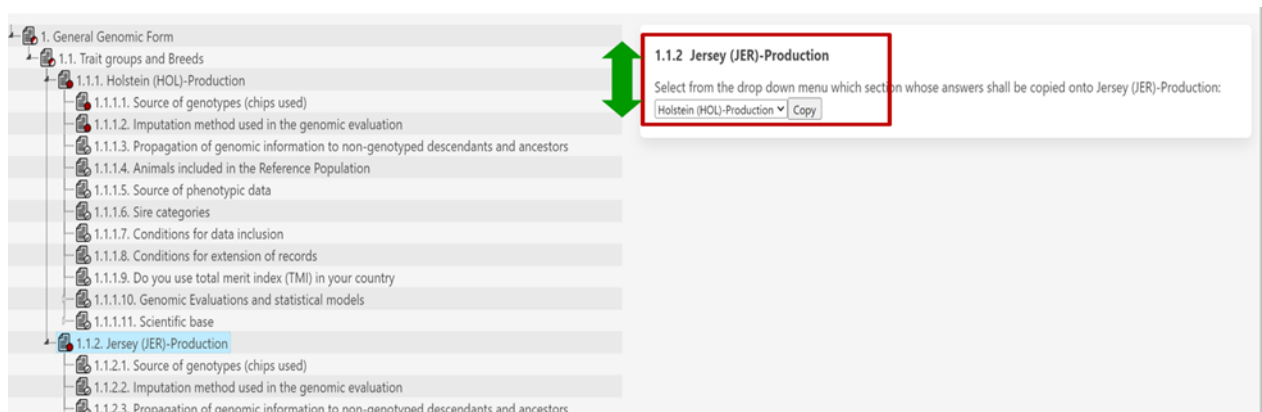


Figure 6: Feature for copying information among different breed-trait-groups in the proposed general genomic form.

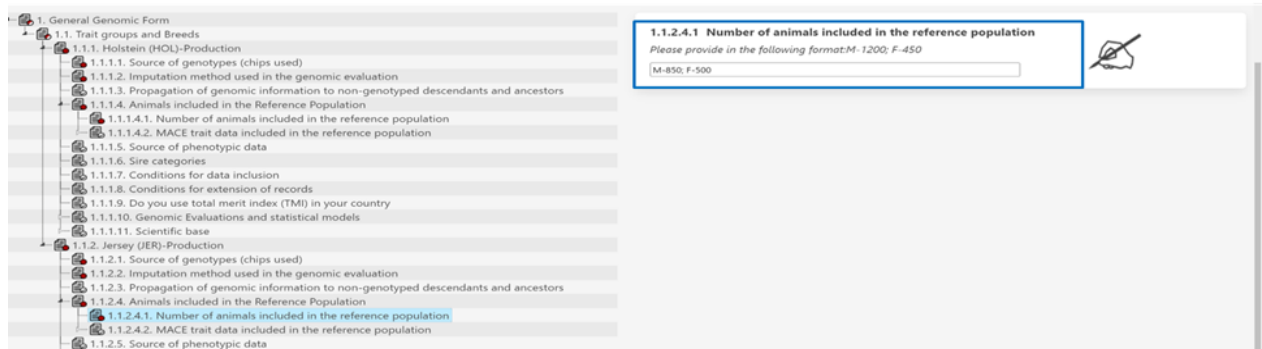


Figure 7: Feature for editing information for different breed-trait-groups in the proposed general genomic form.

Conclusion

PREP is an online platform with many benefits not only for Interbull users but also for breed organizations, NGECS, universities, students and other third parties, regardless of their involvement in international evaluations. Free access to the submitted data via *Submission* and *Data query* tabs makes it efficient and user-friendly to browse and look up information for different traits and breeds for both dairy and beef cattle. The development of a genomic information form is the next milestone in the development of the PREP database. A *general genomic form* with all trait groups in one form will make it easier for users and organizations to provide genomic information in one go, without the need to repeat the information already provided for each trait within the genetic form. The feature for copying and editing the information across different breeds and trait groups makes the genomic form

more efficient. In order to fully benefit from all advantages of the new genomic form, users would still require to fill in all trait-specific information available within the conventional forms.

Acknowledgements

Interbull Centre would like to thank Dr. Paul van Raden (United States Department of Agriculture, USDA) for his ideas and feedback during the development of the genomic form.

References

<http://prep.interbull.org>
https://interbull.org/ib/prep_user_manual
https://interbull.org/ib/eu_detailed_technical_reports