



J. Dürr, H. Jorjani, R. Reents
SLU, Sweden & vit, Germany
2014 ICAR/Interbull Conference, Berlin, May 2014





International platform for genomic data exchange (Dürr & Philipsson, 2012)

- reducing costs and optimizing investments on genotyping
- improving reference populations
- screen large populations for recessive alleles detection
- worldwide parentage verification data base
- study diversity within the bovine populations

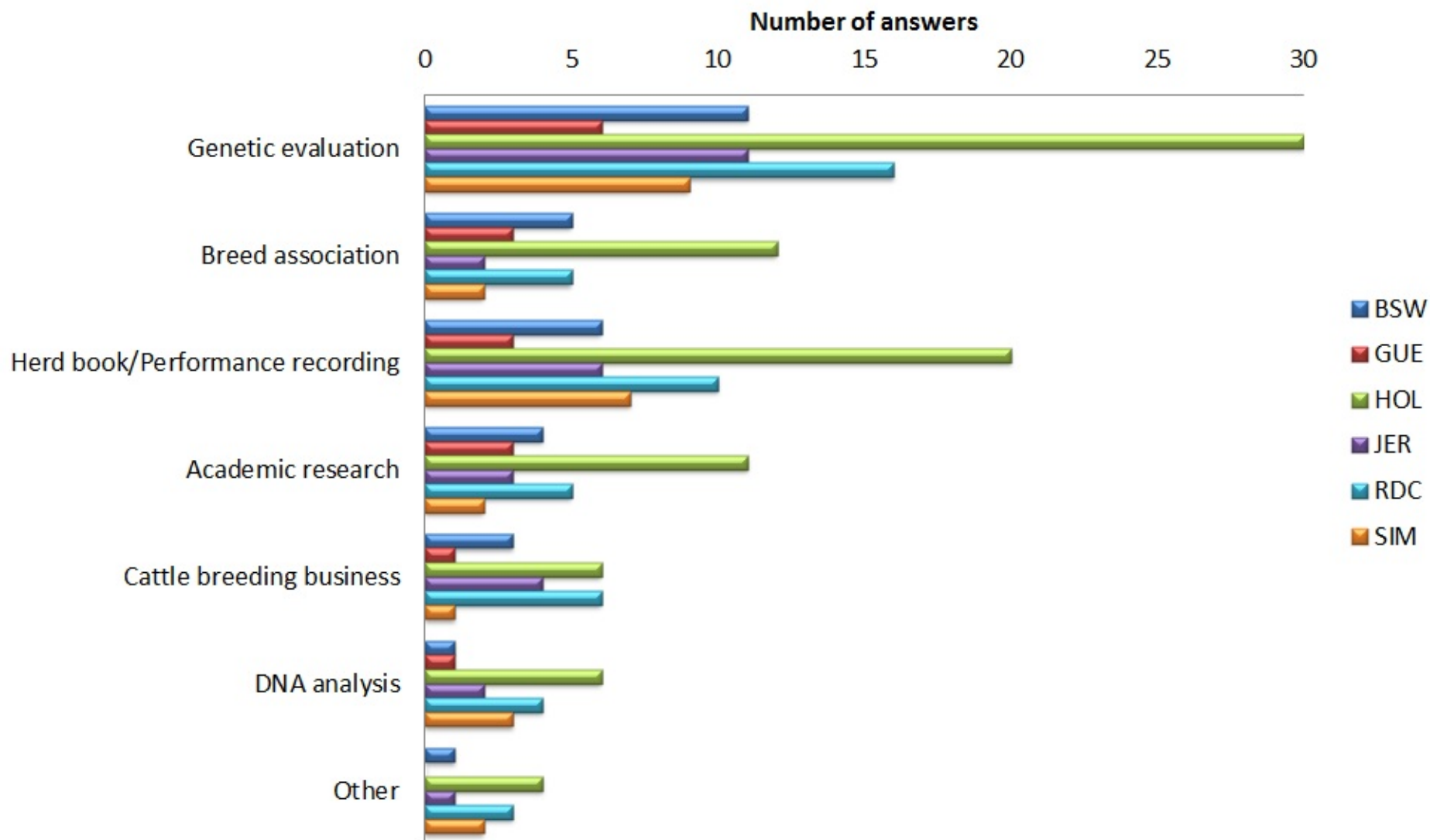


48 responses from 30 countries

INTERBULL SURVEY

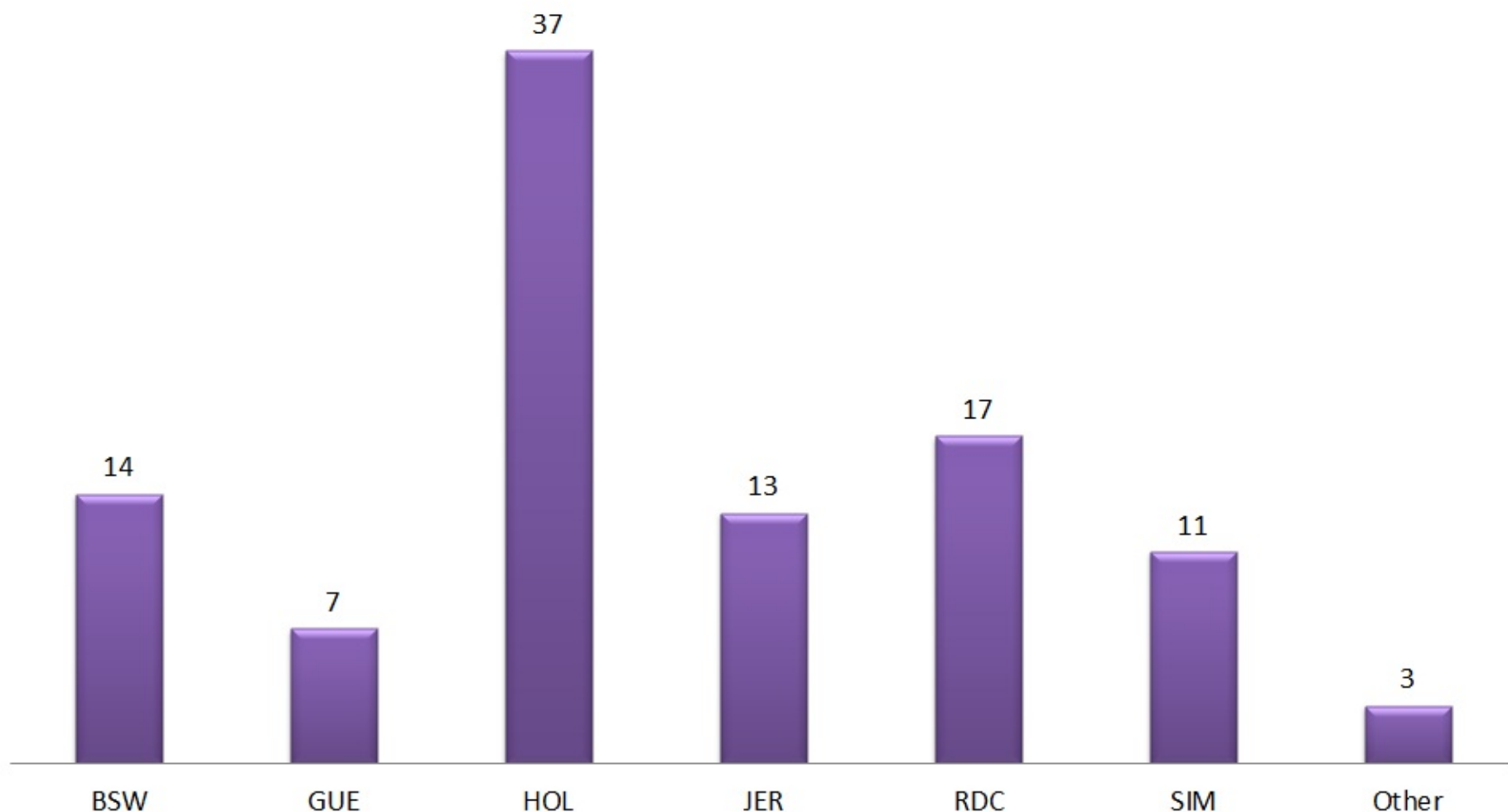
(AUGUST 2012)

Which is the main operation of your organization?



Figure

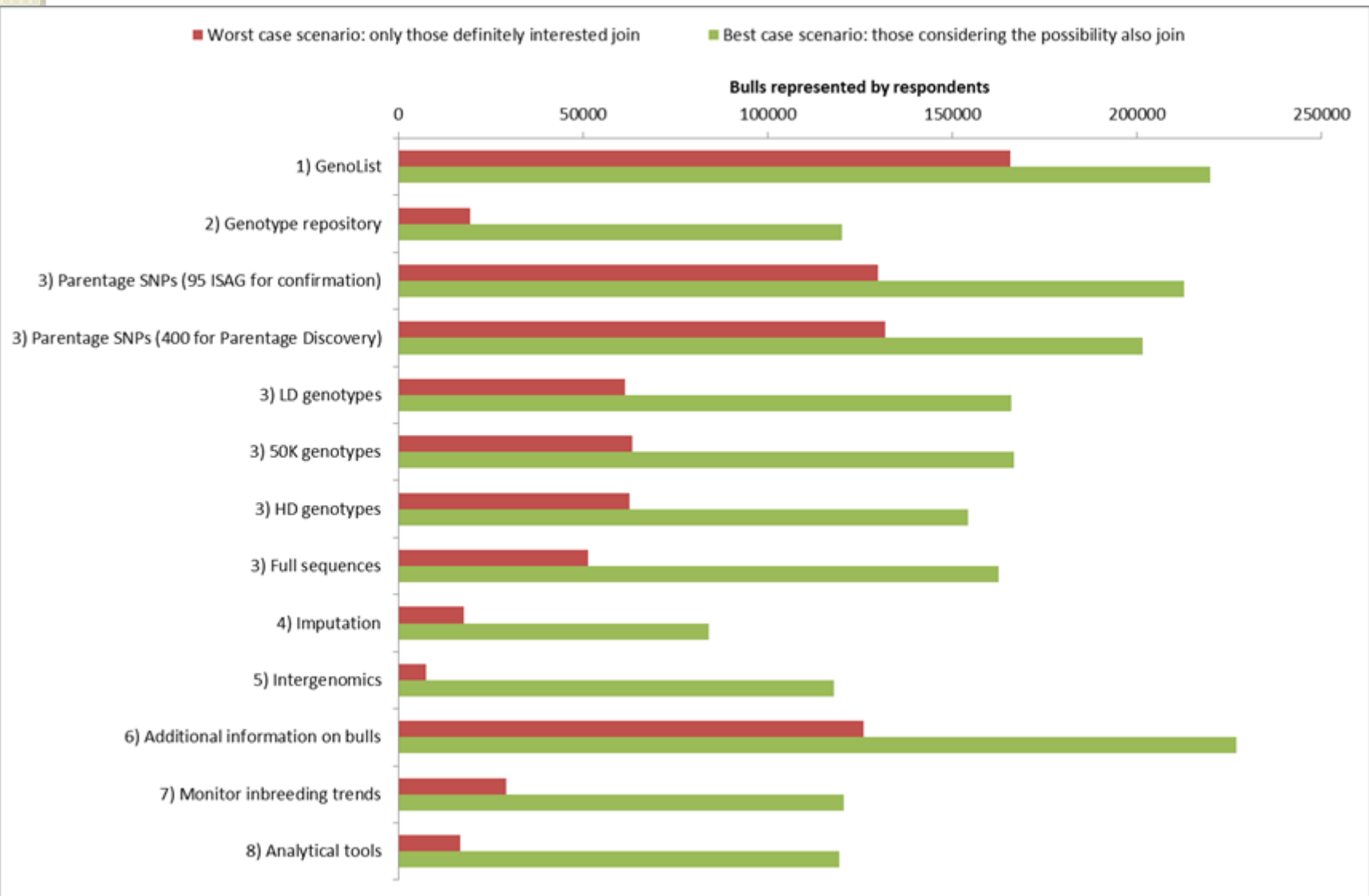
Please, indicate which breed(s) your answers in this section refer to. If your organization has different policies by breed, we suggest you send us one reply to the survey for each case (n=48).



Figure



Genoex – potential demand from survey





Survey conclusions (I)

- Clear interest (immediate establishment):
 - Common list of genotyped animals
 - Additional information on AI bulls
 - Exchange platform for parentage SNPs
- Sufficient interest (worth to invest):
 - Exchange platform for SNP arrays
 - Inbreeding monitoring service
- Customized demand:
 - National genomic data storage
 - Exchange platform for full sequences



Survey conclusions (2)

- Depending on the genomic data base (2nd phase):
 - Imputations
 - Multi-country evaluations
 - Genomic tools
- Different groups demanded different types of services
- Countries with less advanced national genomic programs expect more from Interbull
- The pricing policy must reflect the multitude of services
- Legal implications need to be clearly examined
- Include the participation of organizations that are not among the usual Interbull service users



GENOEX PROPOSAL



VALUE PROPOSITION

- Establish the INFRASTRUCTURE necessary for international cooperation based on SNP data
- Optimize customer investments in genotyping by AVOIDING DUPLICATION
- Establish STANDARD PROTOCOLS FOR GENOMIC DATA EXCHANGE
- Become the international source of BOVINE PARENTAGE SNPS
- Facilitate MULTILATERAL SNP DATA EXCHANGE by establishing a common repository and customer driven access rules
- (Provide affordable GENOMIC DATA STORAGE for small populations)



Choice of tools - In house development

Pros

- Lower initial cost
- Built in integration with Interbull Data Echange Area (IDEA)
- Customized solution
- Possible to adopt open source solutions, such as PostgreSQL

Cons

- Complexity of the task
- Development time
- Lack of in house resources specialized in bioinformatics
- Large number of other projects competing for resources



Choice of tools - IGenoP

Pros

- Lower initial cost
- Already operational
- Developed based on similar needs

Cons

- Uses Oracle (not available at SLU, and license adds a significant cost)
- Transposing the Oracle tables/codes into PostgreSQL would require a significant investment
- ICBF is not a software company and further development would need to be in house
- ICBF manifested no interest to host the services on behalf of Interbull



Choice of tools - BC|SNPmax, by Biocomputing Platforms¹

Pros

- Strong technical background
- Already in use for bovine genomics (Interbull customers)
- Offers solutions/tools highly sophisticated
- Tailored for multiple users with customized settings
- Faster to implement

Cons

- Higher initial cost
- Uses IBM DB2, while IDEA uses PostgreSQL, but the license can be obtained at low cost
- Liability of depending on a commercial software for service operations

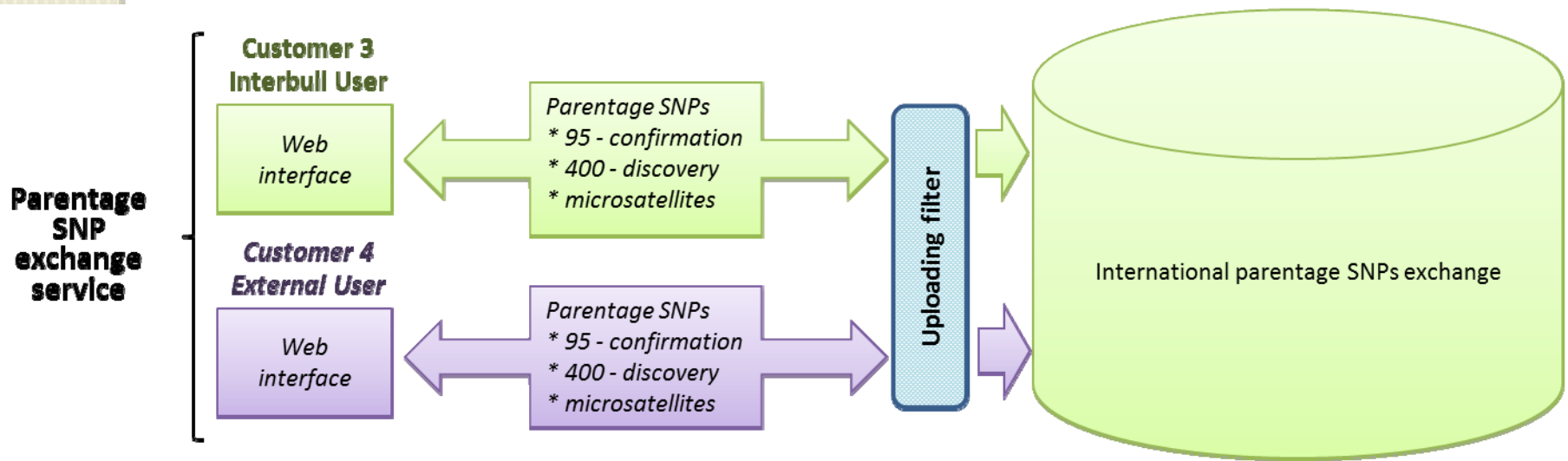


Proposed service categories

- Parentage SNP exchange service (PSE)
- Genomic data exchange service (GDE)
- Customized genomic repository service (CGR)



Parentage SNP exchange service (PSE)





PSE User categories

- Interbull user
- ICAR member user
 - Organization that is an associate or full member of ICAR or represents a full member of ICAR and that is responsible for official parentage verification services in its own coverage area



PSE Data types

- Parentage confirmation SNPs (~100)
- Parentage discovery SNPs (~400)
- Parentage confirmation microsatellite markers

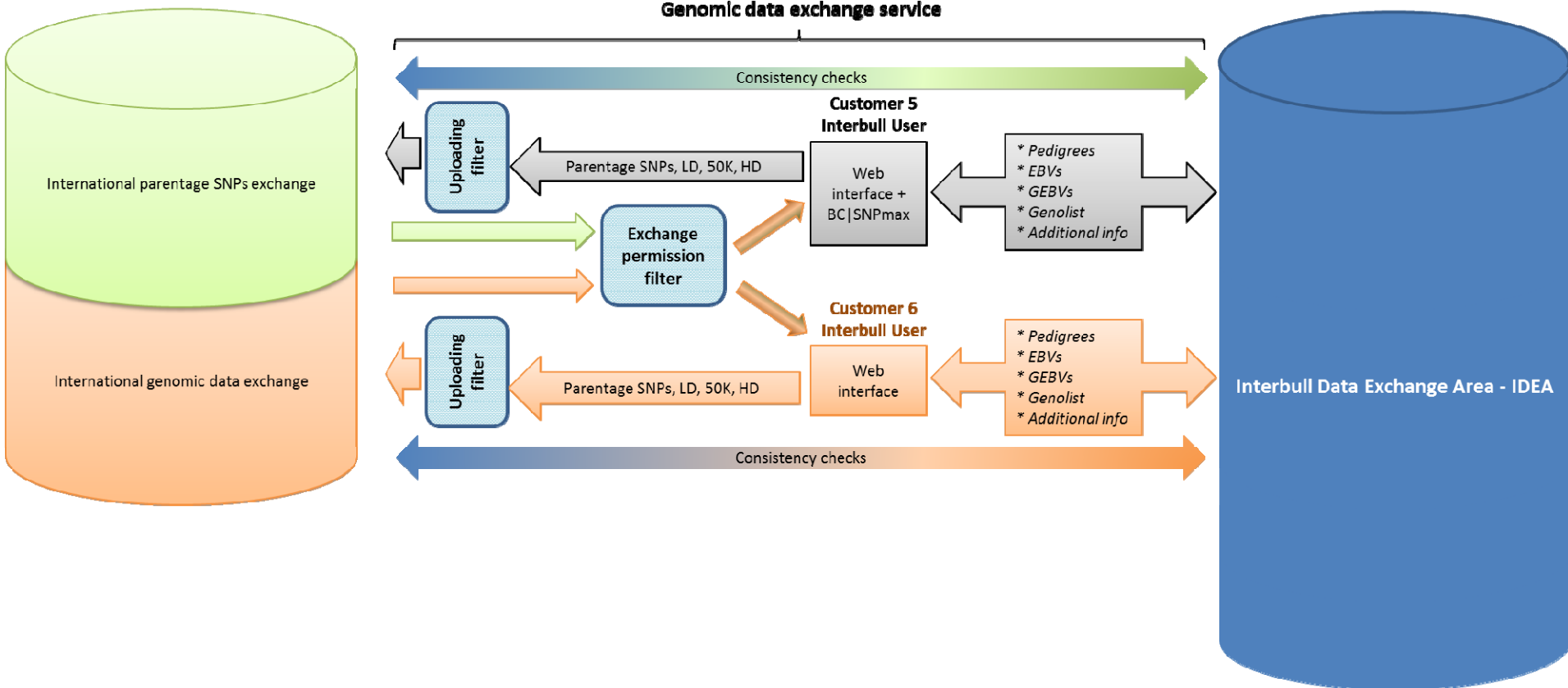


PSE Interface

Type of rule	Example
Data accessibility	Only official national parentage verification service suppliers will be allowed to join the service
Uploading consistency checks	Correct file formats?
	Animal already in the Interbull pedigree (IDEA)?
	Animal parentage has been confirmed/verified previously?
	Animal already has parentage SNPs in the Genoex DB?



Genomic data exchange service (GDE)





GDE User categories

- Interbull users only



GDE Data types

- Parentage confirmation SNPs (~100)
- Parentage discovery SNPs (~400)
- Parentage confirmation microsatellite markers
- Low density SNP arrays
- Middle density SNP arrays
- High density SNP arrays



GDE principles

- Ownership of genotypes belongs to the organization supplying the genotype in the first place and consequently the control of the access rules for users that may have access to the information
- Users are charged proportionally to the respective downloading activity
- Users are granted credit proportionally to the respective uploading activity

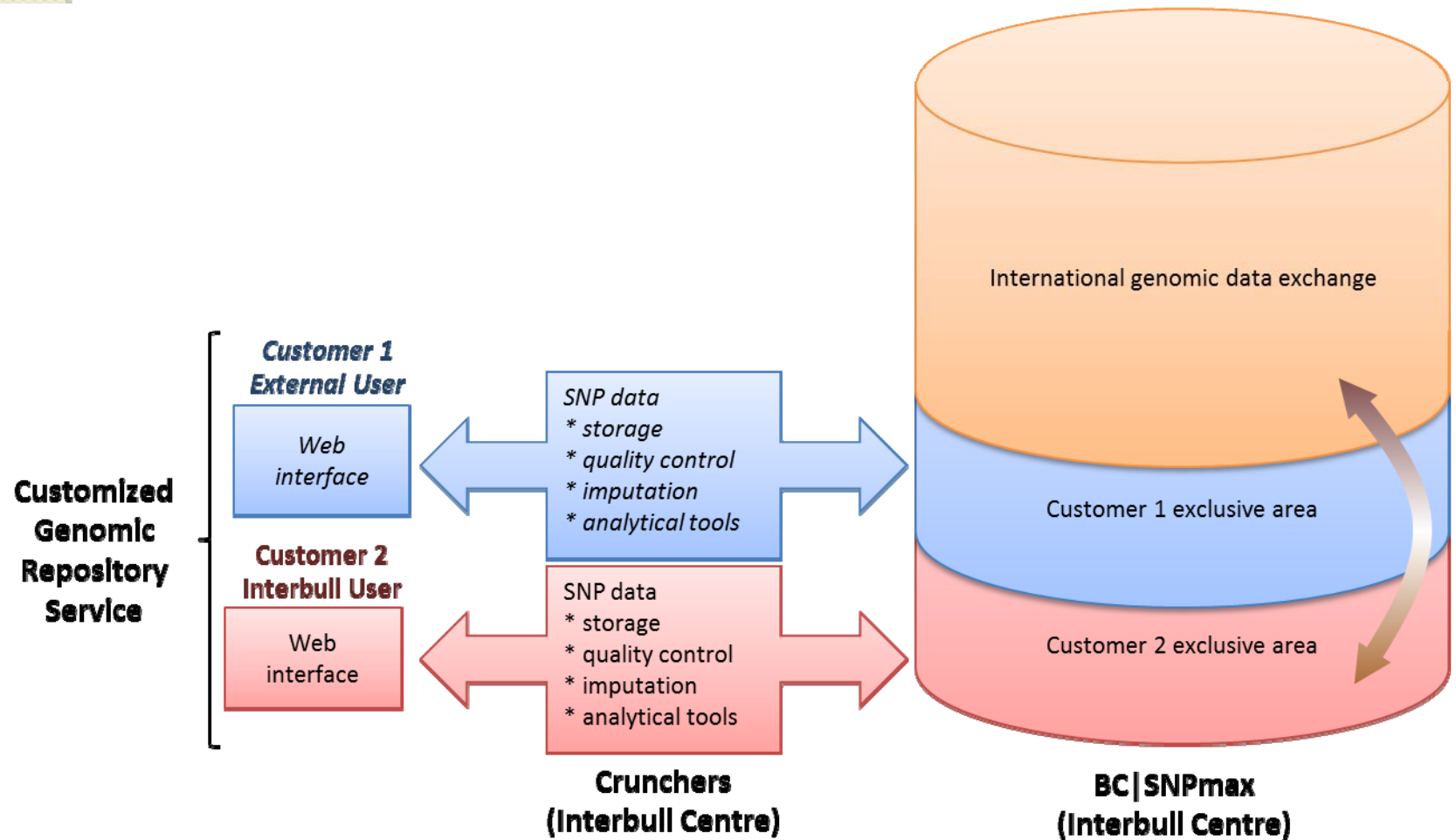


GDE Interface

Type of rule	Example
Data accessibility	Only Interbull users can subscribe to the services
	Permission from owner granted?
Uploading consistency checks	Correct file formats?
	Animal already in the Interbull pedigree (IDEA)?
	Animal already has marker genotype(s) in the Genoex DB?



Customized genomic repository service (CGR)





CGR User categories

- Interbull user
 - Already participating in Interbull services and therefore with access to IDEA
 - Enrollment in the CGR service does not include access to the other service categories described in this proposal
- External user
 - Associate or full member of ICAR or represents a full member of ICAR. No access to IDEA will be granted to external users.



CGR Data types

- Parentage confirmation SNPs (~100)
- Parentage discovery SNPs (~400)
- Parentage confirmation microsatellite markers
- Low density SNP arrays
- Middle density SNP arrays
- High density SNP arrays
- Sequencing data



CGR principles

- Storage for national genomic data in an exclusive area of the Genoex DB servers
- Processing capacity on Interbull crunchers for the users
- Also use the BC|SNPmax data management system

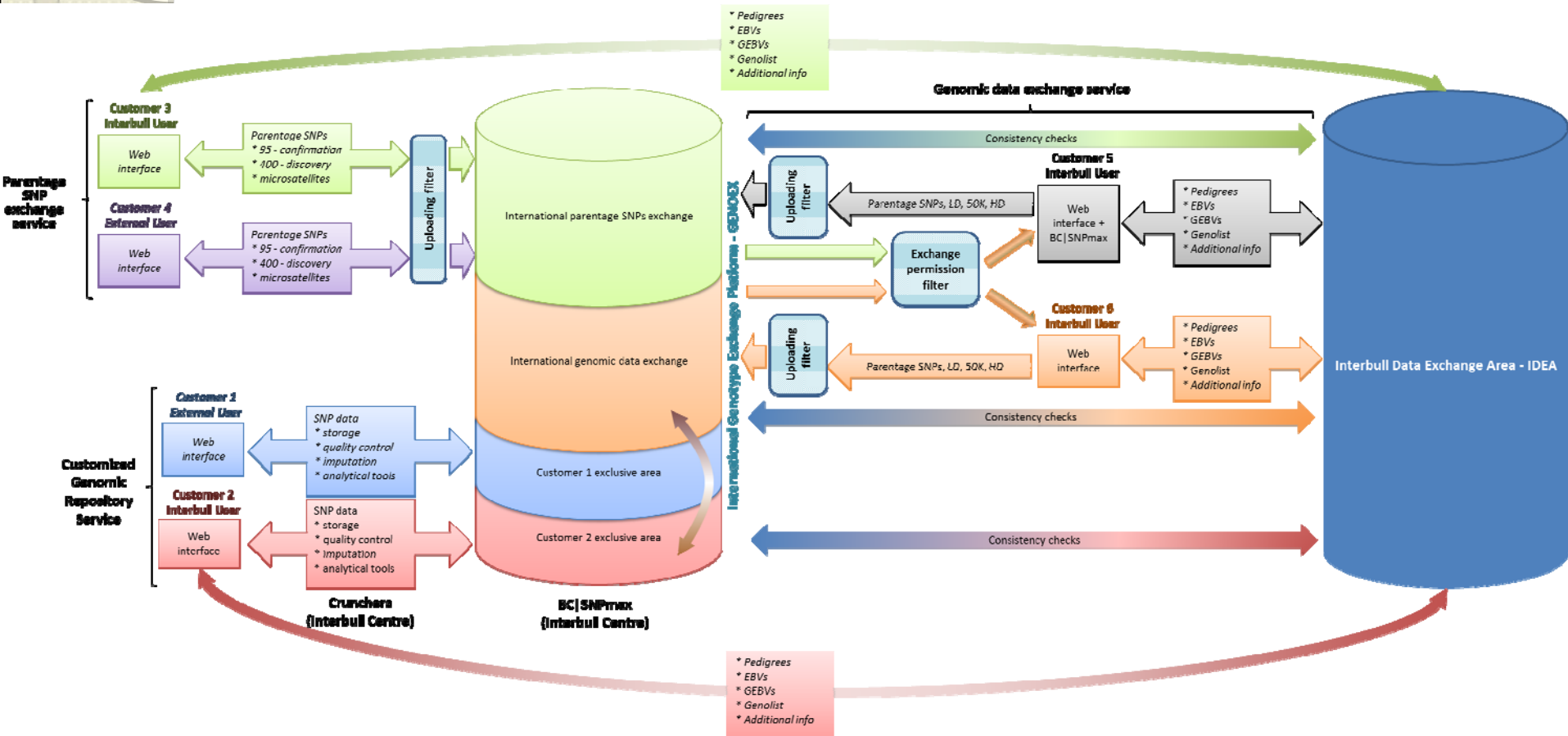


CGR Interface

Type of rule	Example
Data accessibility	Only Interbull users can subscribe to the services
	Permission from owner granted?
Uploading consistency checks	Correct file formats?
	Animal already in the Interbull pedigree (IDEA)?
	Animal already has marker genotype(s) in the Genoex DB?



Genoex – service categories





Service fees

- Parentage SNP exchange service
 - Participation fee I, irrespective of usage rate
- Genomic data exchange service
 - Participation fee II + usage fee proportional to genotype downloads
 - Credit based on genotype uploads
- Customized genomic repository service
 - Participation fee IV + usage fee proportional to disk allocation
- Multiusers
 - Single participation fee + respective usage fees



Implementation calendar

Period	Activity
Summer 2014	BC SNPmax license purchase + installation
Summer – Fall 2014	Development of the parentage SNP exchange (PSE) service interface
Winter 2014/15	Launching of the PSE service
PENDING DECISION	
Spring – Summer 2015	Development of the genotype data exchange (GDE) service interface and integration with IDEA
Winter 2015/16	Launching of the GDE service
2016	Development of the customized genomic repository (CGR) service



Final remarks

- The GENOEX proposal represents a unique opportunity for ICAR/Interbull to provide a much needed international platform for genomic information exchange.
- Demand has been identified and the means presented here show that it is not only feasible, but also affordable.
- The projected service fees are remarkably lower than isolated investments necessary for genotyping and genomic data storage/handling.



Final remarks

- As a cooperative effort that would involve only those customers interested in benefiting from the proposed platform, it does not prevent other Interbull customers to continue developing their own strategies and infrastructure.
- As it is the case for all Interbull activities, there is no intention to replace or compete with national organizations; on the contrary, GENOEX is conceived to offer auxiliary tools for the national expertise to develop their own programs more efficiently.



J. Dürr, H. Jorjani, R. Reents
2014 ICAR/Interbull Conference, Berlin, May 2014

THANK YOU!

WWW.INTERBULL.ORG

GEN  EX
International Genotype Exchange Platform