Antibody Characterization Standards

Characterization of an antibody for a transcriptional regulator (including TFs, chromatin remodelers, and chromatin modifiers)

Before ChIP-seq data can be released for any transcriptional regulator (including TFs, RNA-binding proteins and chromatin remodelers), the specific lot number of the antibody used to collect the data must be approved for use in that cell type by the antibody characterization review committee of the ENCODE DCC. A primary and secondary characterization must both be submitted for review and deemed *compliant*, in order for an antibody lot to be considered validated and *eligible for use* in new binding assays.

1. In all cases, a primary characterization of the antibody must be provided for each lot number and each specific cell type. This can be either an immunoblot (Western blot) or an immunoprecipitation (IP-Western). [slide 1]

a. If the major band is within 20% of the size predicted by the size of the coding region of the protein and corresponds to >50% of all bands on the gel (excluding the antibody bands in the case of an immunoprecipitation), then the antibody passes this initial characterization. Each immunoblot should indicate which cells types were tested and include size markers. For immunoprecipitations, a control IgG precipitation should also be included.

b. If the antibody does not pass this initial characterization (e.g. due to multiple splice variants or multiple modifications that alter the electrophoretic properties of the protein), it can be rescued by a secondary characterization that supports the conclusion that the band(s) detected correspond to the correct protein (e.g. all bands are reduced upon treatment with siRNA to that protein).

c. If the antibody does not pass this initial characterization either due to no signal (or only signals at the wrong size) on the gel, it is possible that the antibody does not recognize the denatured protein. In this case, the antibody can be rescued by Mass Spec (see below) as the sole characterization method. However, the failed Western blot must be provided as part of the antibody characterization document. If other methods are used to characterize an antibody that doesn't recognize the denatured protein (and thus does not work on a Western blot), the antibody characterization document should be submitted as a special request (see Note 3).

2. Use of the same lot number of a previously characterized antibody in a new cell type. If a specific lot number for an antibody has previously passed both primary and secondary characterization in another cell type and if the banding pattern on the immunoblot or immunoprecipitation is the same in the new cell type as in the characterized cell type, then no further characterization is needed for the antibody in that new cell type. If the banding pattern is different in the new cell type, a secondary characterization must be performed in the new cell type. [slide 2]

3. Use of a different lot number of a previously characterized antibody. If this is the first time that a new lot number has been used for a previously characterized antibody, a primary characterization must be performed using at least one of the same cell types used to characterize the previous lot number (indicate which specific previously characterized lot should be used for comparison by its ENCODE antibody accession) and including the cell type for which ChIP-seq data will be deposited for the new lot number. If the patterns for the new lot number are the same in the previously characterized cell type and in the cell type for which

ChIP-seq data will be deposited as shown in the characterization of the original lot number of that antibody, then no further characterization is required. If the banding patterns are different, a secondary characterization must be performed. [Slide 3]

Allowable Secondary Characterization Methods:

1. Secondary characterization using siRNA or shRNA against the mRNA of the target protein. For siRNA or shRNA characterization, the band(s) detected by the antibody should be reduced to no more than 50% of the original signal. The sequence or vendor and catalog number of the oligonucleotide(s) should be provided. A control knockdown should also be performed. Cell types should be labeled and size markers should be included on the immunoblot. A brief description of the transfection protocol should also be provided.

2. Secondary characterization using ChIP-seq data obtained using a previously characterized antibody for that factor. If ChIP-seq data for a different lot number of a previously characterized antibody or a previously characterized, but different, antibody for a given transcriptional regulator is available, this ChIP-seq data can be used for characterization of a new antibody or new lot number. In addition, ChIP-seq data obtained using an epitope-tagged version of the target protein and an antibody that recognizes the tag can be used for comparison. The ChIP-seq data from the new antibody or new lot number should be compared to the previous ChIP-seq data. If the two datasets pass the ENCODE IDR cuts-offs for narrow peak ChIP-seq reproducibility (see below for current IDR standards), then the secondary characterization of the new antibody/lot number is compliant with this defined standard. When submitting documents, indicate the specific antibody lot and ChIP-seq data used in the comparison by their ENCODE antibody and experiment accessions respectively.

3. Secondary characterization using overexpressed, tagged proteins. For proteins that are resistant to knockdown using siRNAs (e.g. very stable proteins), comparison to overexpressed, tagged proteins can be used for characterization. In this case, the primary characterization of the antibody must first show the appropriate specificity. Then, two side-by-side immunoblots should be performed using control cells and cells overexpressing the factor. The first immunoblot should employ the antibody to the tag to show the position of the exogenous factor and the second immunoblot should employ the antibody to the endogenous factor to show that the band in the control cells that is identified by the endogenous antibody is the same size as the exogeneous protein.

4. Secondary characterization using Mass Spectrometry.

<u>What should be reported:</u>

- IP-western blot of gel image with outline of gel slices.
- All peptides (with peptide counts) from all immunoreactive bands.
- Fold enrichment of all peptides in the immunoreactive bands vs either mock IP or a set of proteins that have been immunoprecipitated from the same cell type using a collection of other antibodies from the same host species (the list of proteins used as the set of IP contaminants list must be provided).
- Indication as to which proteins above the target protein on the ranked list are TFs and which TFs are members of the same TF family as the target protein.

<u>What should be required to pass validation:</u>

- The target should be enriched in the IP when compared to a mock IP or to a set of proteins that have been immunoprecipitated from the same cell type using a collection of other antibodies from the same host species.
- The target should be in the top 20 ranked proteins and the top most enriched TF in the immunoreactive band, unless the higher ranked TFs are known interacting partners.
- The target should be the top ranked member of that family of TFs (exceptions will be allowed if a publication is provided that demonstrates that a higher ranked family member is known to dimerize with the target protein).
- If an antibody doesn't meet these characteristics, the antibody characterization document can be submitted for consideration as a special request (see Note 3).

Notes:

1. These methods refer to characterization of antibodies that recognize endogenously expressed proteins. The requirements for characterization of epitope-tagged proteins are described elsewhere.

2. Current IDR standards for a narrow-peak ChIP-seq dataset are:
   Rescue Ratio RR_new = | Np U Nt | / | Np ^ Nt |
   Self consistency ratio SR_new = | N1 U N2 | / | N1 ^ N2 |
   where
   ^ = intersection (common) of 2 peak sets
   U = union (merge) of 2 peaks sets

   If (RR_new > 2) AND (SR_new > 2) then the replicates are proclaimed to have low reproducibility (failed) and flagged with -1 quality score
   If (RR_new > 2) OR (SR_new > 2) but not both, then the replicates are proclaimed to have moderate reproducibility (passed) and flagged with a 0 quality score
   If (RR_new <= 2) AND (SR_new <= 2) then the replicates are proclaimed to have high reproducibility (passed) and flagged with +1 quality score

3. We realize that, in rare cases, situations may arise in which antibodies do not pass these standards but the data producers feel that the datasets should be submitted. Therefore, exceptions to these characterization standards will be considered for special cases. The antibody characterization review committee of the DCC, along with the Production PIs, will consider these special requests. If an exception is granted, the datasets using these "exempt" antibodies will be flagged in the ENCODE datasets.

Characterization of an antibody for an RNA Binding Protein

Before binding data can be released for any RNA binding protein (RBP), the specific lot number of the antibody used to collect the data must be approved for use in that cell type by the antibody characterization review committee of the ENCODE DCC. A primary and secondary characterization must both be submitted for review and deemed *compliant,* in order for an antibody lot to be *eligible for use in new binding assays*.

1. In all cases, a primary characterization of the antibody must be provided for each lot number and each specific cell type. This can be either an immunoblot (Western blot) or an immunoprecipitation (IP-Western). [slide 1]

a. If the major band is within 20% of the size predicted by the size of the coding region of the protein and corresponds to >50% of all bands on the gel (excluding the antibody bands in the case of an immunoprecipitation), then the antibody passes this initial characterization. Each immunoblot should indicate which cells types were tested and include size markers. For immunoprecipitations, a control IgG precipitation should also be included.

b. If the antibody does not pass this initial characterization (e.g. due to multiple splice variants or multiple modifications that alter the electrophoretic properties of the protein), it can be rescued by a secondary characterization that supports the conclusion that the band(s) detected correspond to the correct protein (e.g. all bands are reduced upon treatment with siRNA to that protein).

c. If the antibody does not pass this initial characterization either due to no signal (or only signals at the wrong size) on the gel, it is possible that the antibody does not recognize the denatured protein. In this case, the antibody can be rescued by Mass Spec (see below) as the only characterization method. However, the failed Western blot must be provided as part of the antibody characterization document. If other methods are used to characterize an antibody that doesn't recognize the denatured protein (and thus does not work on a Western blot), the antibody characterization document should be submitted as a special request (see Note 3).

2. Use of the same lot number of a previously characterized antibody in a new cell type. If a specific lot number for an antibody has previously passed both primary and secondary characterization in another cell type and if the banding pattern on the immunoblot or immunoprecipitation is the same in the new cell type as in the characterized cell type, then no further characterization is needed for the antibody in that new cell type. If the banding pattern is different in the new cell type, a secondary characterization must be performed in the new cell type. [slide 2]

3. Use of a different lot number of a previously characterized antibody. If this is the first time that a new lot number has been used for a previously characterized antibody, a primary characterization must be performed using at least one of the same cell types used to characterize the previous lot number and including the cell type for which ChIP-seq data will be deposited for the new lot number. If the patterns for the new lot number are the same in the previously characterized cell type and in the cell type for which ChIP-seq data will be deposited as shown in the characterization of the original lot number of that antibody, then no further characterization is required. Indicate which specific previously characterized lot should be used for comparison by its ENCODE antibody accession. If the banding patterns are different, a secondary characterization must be performed. [Slide 3]

Allowable Secondary Characterization Methods:

1) **Secondary characterization using siRNA or shRNA against the mRNA of the target protein**. For siRNA or shRNA characterization, the band(s) detected by the antibody should be reduced to no more than 50% of the original signal. The sequence or vendor and catalog number of the oligonucleotide(s) should be provided. A control knockdown should also be performed. Cell types should be labeled and size markers should be included on the immunoblot. A brief description of the transfection protocol should also be provided.

2) **Secondary characterization using overexpressed, tagged proteins.** For proteins that are resistant to knockdown using siRNAs (e.g. very stable proteins), comparison to overexpressed, tagged proteins can be used for characterization. In this case, the primary characterization of the antibody must first show the appropriate specificity. Then, two side-by-side immunoblots should be performed using control cells and cells overexpressing the factor. The first immunoblot should employ the antibody to the tag to show the position of the exogenous factor and the second immunoblot should employ the antibody to the endogenous factor to show that the band in the control cells that is identified by the endogenous antibody is the same size as the exogeneous protein.

3) **Secondary characterization using Mass Spectrometry.**

What should be reported:

- IP-western blot of gel image (with outline of gel slices if individual slices are analyzed).
- All peptides (with peptide counts) from all immunoreactive bands.
- Fold enrichment of all peptides in the immunoreactive bands vs either mock IP or a set of proteins that have been immunoprecipitated from the same cell type using a collection of other antibodies from the same host species (the list of proteins used as the set of IP contaminants must be provided).
- Indication as to which proteins above the target protein on the ranked list are RBPs and which RBPs are members of the same family as the target protein.

What should be required to pass validation:

- The target should be enriched in the IP when compared to a mock IP or to a set of proteins that have been immunoprecipitated from the same cell type using a collection of other antibodies from the same host species.
- The target should be in the top 20 ranked proteins and the top most enriched RBP in the immunoreactive band, unless the higher ranked RBPs are known interacting partners.
- The target should be the top ranked member of that family of RBPs (exceptions will be allowed if a publication is provided that demonstrates that a higher ranked family member is known to dimerize with the target protein).
- If an antibody doesn't meet these characteristics, the antibody characterization document can be submitted for consideration as a special request (see Note 3).

Notes:

1. These methods refer to characterization of antibodies that recognize endogenously expressed proteins. The requirements for characterization of epitope-tagged proteins are described elsewhere.
2. We realize that, in rare cases, situations may arise in which antibodies do not pass these standards but the data producers feel that the datasets should be submitted. Therefore, exceptions to these characterization standards will be considered for special cases. The antibody characterization review committee of the DCC, along with the Production PIs, will consider these special requests. If an exception is granted, the datasets using these "exempt" antibodies will be flagged in the ENCODE datasets.

Characterization of a Histone Antibody

Before ChIP-seq data can be released for any antibody that detects a modified histone, the specific lot number of the antibody used to collect the data must be approved by the antibody characterization review committee of the ENCODE DCC. A primary and secondary characterization must both be submitted for review and deemed *compliant*, in order for an antibody lot to be *eligible for use* in new binding assays. [slide 4]

1. **In all cases, a primary characterization of each lot number of an antibody for a modified histone must be performed**. Primary characterization of a specific lot number for a histone antibody should include an immunoblot using nuclear or whole cell extracts from the species in which the antibody will be used for ChIP-seq. The specific histone band detected in the extracts should constitute at least 50% of the protein signal and show at least 10-fold enrichment relative to any other single band. In addition, this signal should be at least 10-fold enriched relative to that detected using the unmodified recombinant histone (if this is included on the western blot). Although it is preferable that a recombinant unmodified histone be included on the western blot, if unmodified relevant peptides are included in the dot blots, peptide arrays or peptide competitions that are used in the secondary characterization step, then it is not required that recombinant histones be included in the primary characterization step. Each immunoblot should label the cells and include size markers.

2. **Use of the same lot number of a previously characterized antibody in a new cell type.** Once a specific lot number for an antibody to a modified histone has passed both the primary and secondary characterizations, it is eligible for use in all cell types of that same species.

3**. Use of a different lot number of a previously characterized antibody.** If this is the first time that a specific lot number has been used for a previously characterized antibody to a modified histone (same catalog number but different lot number), both a primary and a secondary characterization must still be performed for this lot number.

Allowable Secondary Characterization Methods:

1. Secondary characterization using dot blots, peptide arrays or peptide competitions. Commercial arrays are acceptable and the vendor and catalog number should be provided. If arrays that include small numbers of peptides are used, it is critical that the array contain the most relevant peptides for the antibody being tested. If peptide competition is used to validate a histone antibody, it is also important that the most relevant peptides for the antibody being tested are used in the competition series. A 10-fold enriched binding signal for the modification

of interest relative to other modifications is required. An annotated map of the peptides present on the blot or array should be provided.

2. It is allowable to use similarity of ChIP-seq data of a different lot number of the same histone antibody, a different antibody for the same histone modification, or a component of the histone modifying complex that puts on the histone mark as a characterization method. If the two datasets pass the ENCODE IDR cuts-offs for ChIP-seq reproducibility (see below for current IDR standards), then the secondary characterization of the new antibody/lot number is compliant with this defined standard. When submitting documents, indicate the specific antibody lot and ChIP-seq data used in the comparison by their ENCODE antibody and experiment accessions respectively.

Notes:

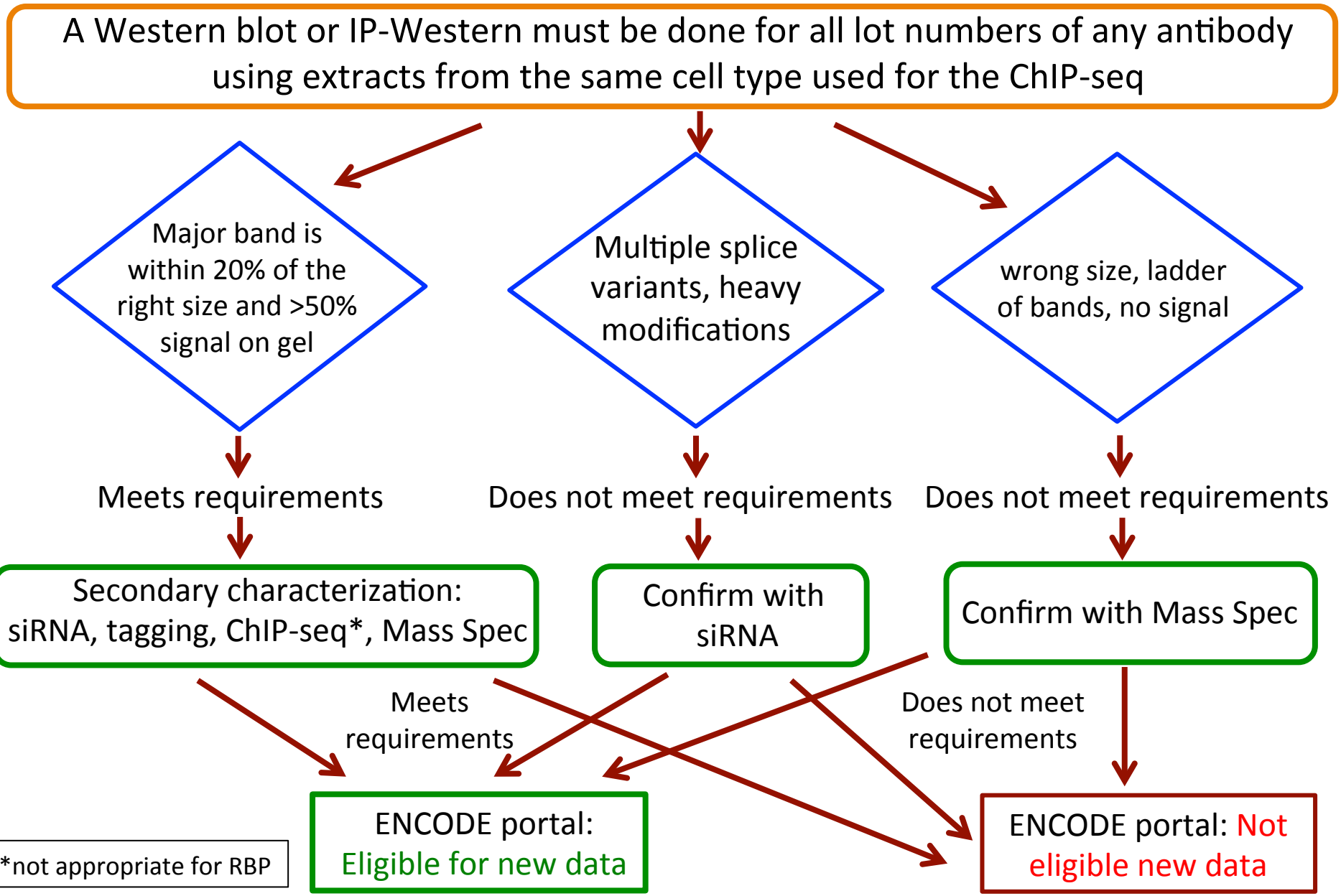
1. At the present, the only approved IDR standards for ChIP-seq datasets are for narrow-peak datasets. The current IDR standards for a narrow-peak ChIP-seq dataset are:
   Rescue Ratio RR_new = | Np U Nt | / | Np ^ Nt |
   Self consistency ratio SR_new = | N1 U N2 | / | N1 ^ N2 |
   where
   ^ = intersection (common) of 2 peak sets
   U = union (merge) of 2 peaks sets

   If (RR_new > 2) AND (SR_new > 2) then the replicates are proclaimed to have low reproducibility (failed) and flagged with -1 quality score
   If (RR_new > 2) OR (SR_new > 2) but not both, then the replicates are proclaimed to have moderate reproducibility (passed) and flagged with a 0 quality score
   If (RR_new <= 2) AND (SR_new <= 2) then the replicates are proclaimed to have high reproducibility (passed) and flagged with +1 quality score

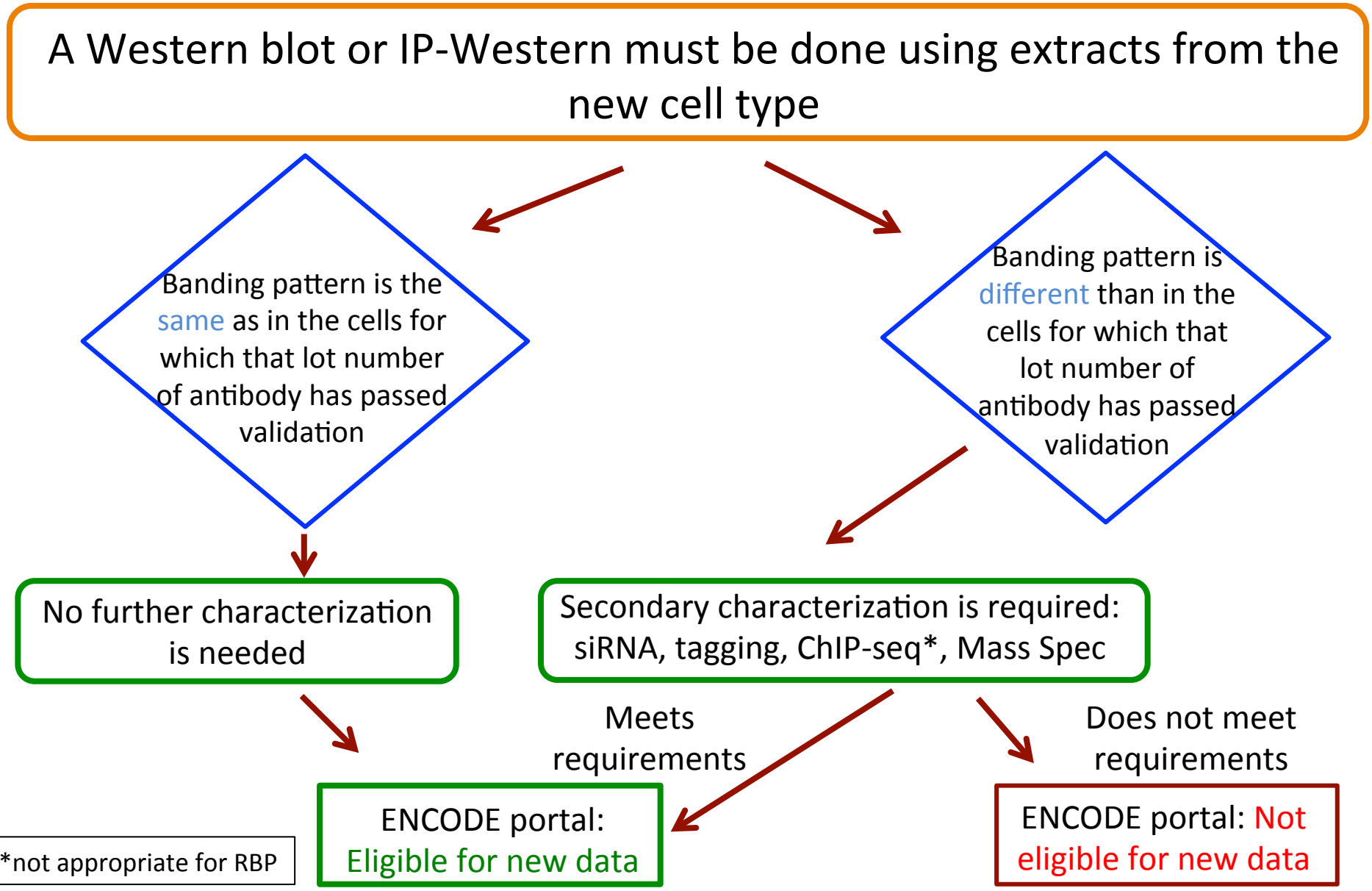
2. Additional IDR standards for histone ChIP-seq datasets (e.g. for broad peaks) will be incorporated into future Antibody Characterization standards.

3. We realize that, in rare cases, situations may arise in which antibodies do not pass these standards but the data producers feel that the datasets should be submitted. Therefore, exceptions to these characterization standards will be considered for special cases. The antibody characterization review committee of the DCC, along with the Production PIs, will consider these special requests. If an exception is granted, the datasets using these “exempt” antibodies will be flagged in the ENCODE datasets.

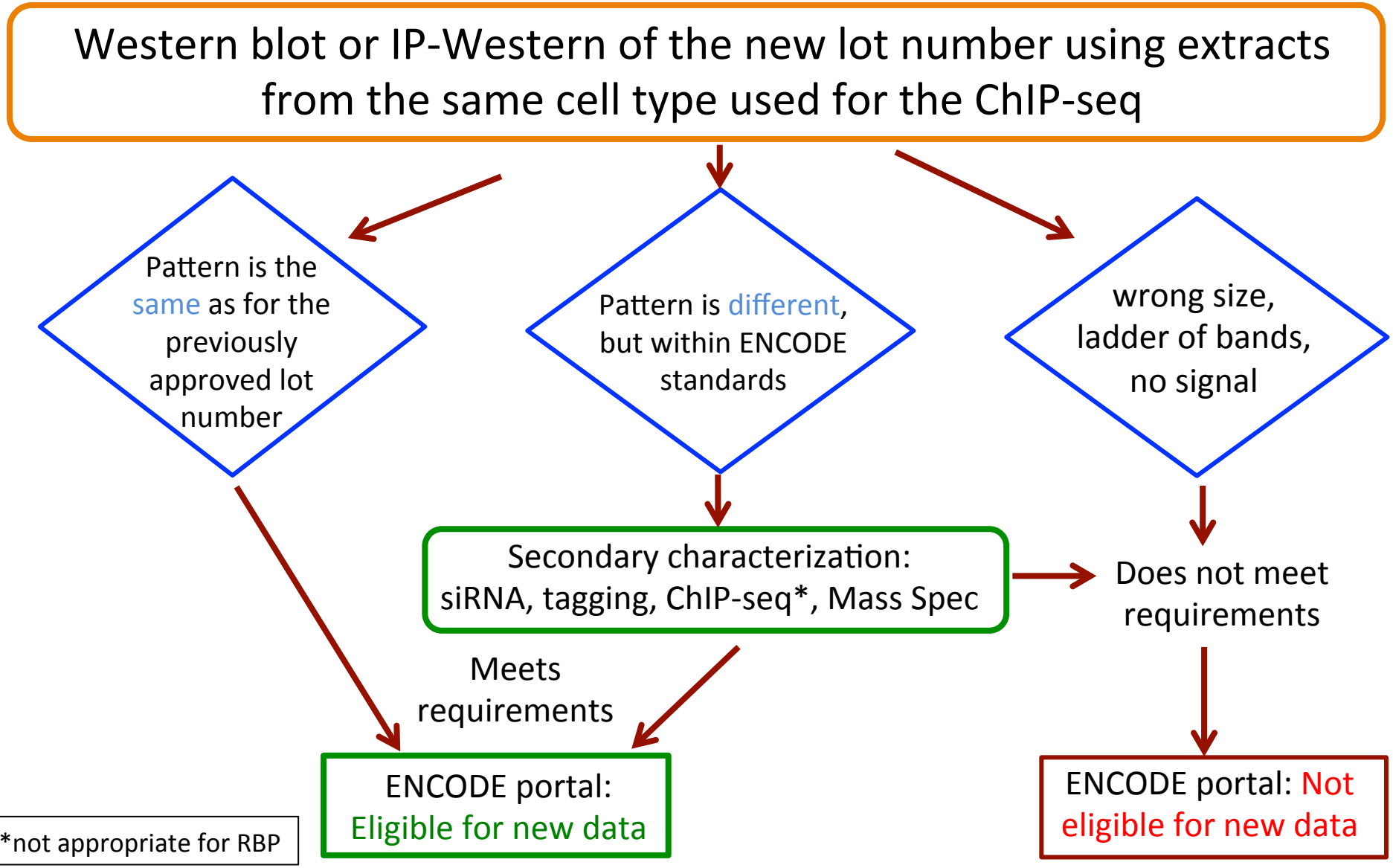
Initial characterization of an antibody (for TFs, chromatin remodelers, chromatin modifiers, RBPs)



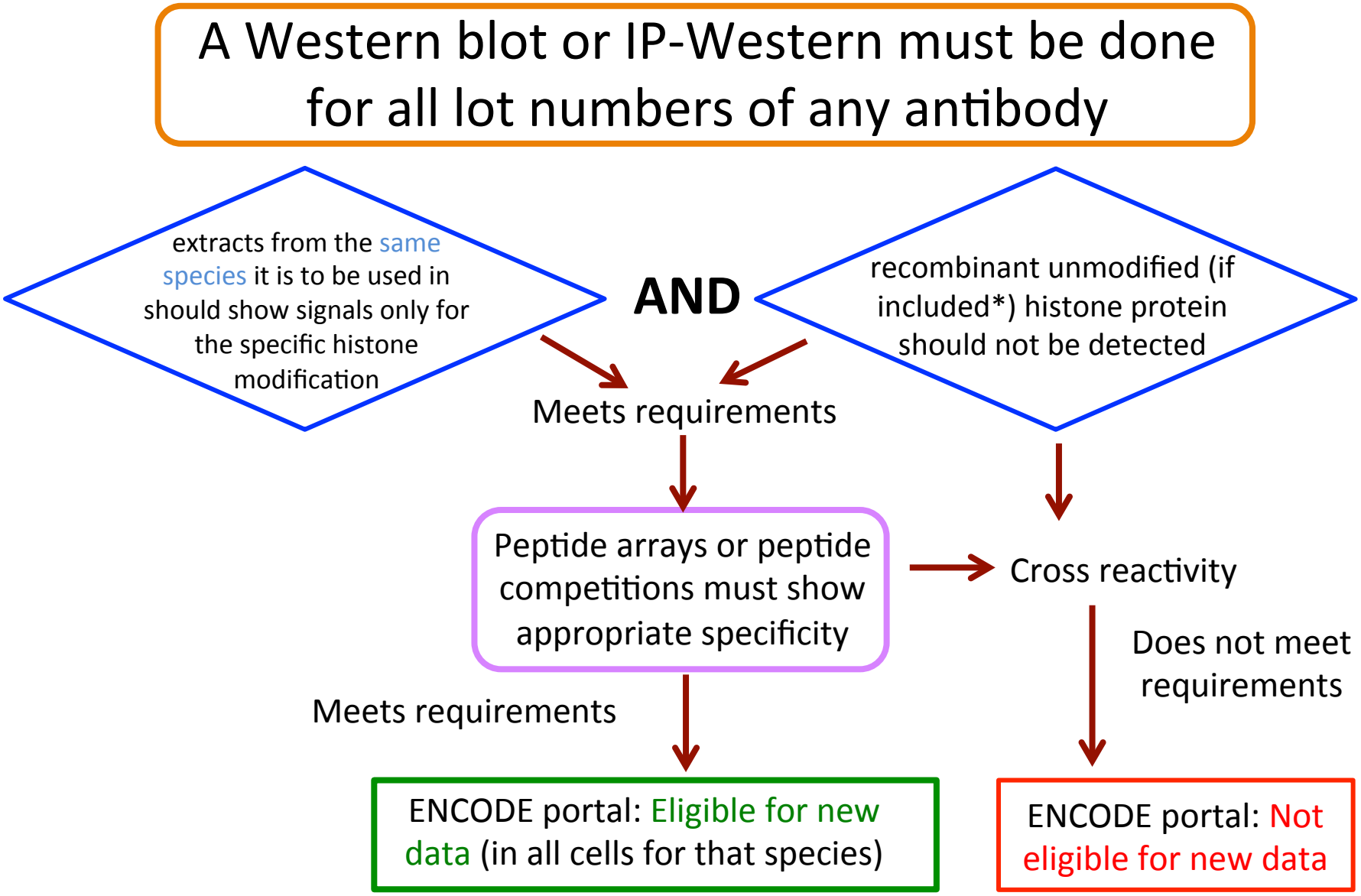
Use of a previously characterized lot number of an antibody in a new cell type (for TFs, chromatin remodelers, chromatin modifiers, RBPs)



Use of a new lot # of a previously characterized antibody (for TFs, chromatin remodelers, chromatin modifiers)



Characterization of a modified histone antibody



*not required on blot if relevant unmodified peptide is included in the secondary characterization