HCV

Drug

Resistance

Advisory

Group

Sequence Analysis Working Group (SAWG)

Update

Ann Kwong

November 2, 2007

Overview of SAWG Issues

- *Dale's Bins*

1. *Issues where an immediate recommendation can be made to guide the process in the right direction*
2. *Issues where a recommendation can be made after discussion within the working group*
3. *Issues where there are multiple solutions and the SAWG come to a recommendation after discussion.*
4. *Issues where not enough information is known to make a recommendation*

Issue 1: what region to sequence- Bin #2

Issue 2: viral load requirements Bin #2

Issue 3: reference sequence Bin #2

Issue 4: nomenclature Bin #2, Bin #3 *(in progress)*

Issue 5: clonal vs population sequencing Bin #2,Bin #3

Issue 6: new technologies Bin #4

Validation: *The SWAG proposes that a separate section on Validation needs to be added to the whole HCV DRAG document which will include validation of phenotypic and sequencing assays.*

How the SAWG worked together

- Defined Major Issues in broad strokes
- Teleconference calls to discuss sub issues (questions) for each issue
- Draft Recommendation Team volunteered for each issue
- Recommendations compiled and edited by the entire team

What should be sequenced?

Team: Mike Otto, Rob Ralston, Isabel Najera, Hongmei Mo, Anita Howe, Jules O'Rear, Gaston Picchio, Ann Kwong (leader)

Questions

- Can we make recommendations for different drug classes/MOA?

What are the viral load requirements for sequence analysis?

Team: George Kukolj, Neil Parkin, Gaston Picchio, Tara Kieffer, Veronica Miller, Lieven Stuyver (leader)

Questions

- Sensitivity depends on the region being sequenced.
- 1000 IU/mL is a good starting target, but this should be a non-binding recommendation>> might be too hard to reach

Do we compare isolates to baseline or to consensus sequences?

Team: Mike Otto,Charles Boucher, Isabel Najera, Tara Kieffer, Lieven Stuyver, (Neil Parkin)

Questions

- 2 types of analysis
  - Isolates from patients failing treatment are compared to their baseline
  - The effect of baseline polymorphic differences on treatment outcome

Do we compare isolates to baseline or to consensus sequences?

Team: Mike Otto,Charles Boucher, Isabel Najera, Tara Kieffer, Lieven Stuyver, Neil Parkin (leader)

Questions- continued

- Need to define consensus sequence
  - All the baseline sequences in a clinical trial
  - Sequences in Gen Bank
  - A single representative sequence such as geno 1a HCV-h or geno 1b Con 1
  - Some genotypes are not well represented in the databases
- FDA would like a standardized approach

Nomenclature

Team: HongMei Mo, Charles Boucher, Neil Parkin, Kai Lin, Jean-Michel Pawlotsky, Ann Kwong

Questions:

- How do you distinguish between a drug-associated mutation and polymorphisms?
- What is a working definition of a resistance mutation?
- Specific amino acid substitutions at a particular amino acid location
- All changes from WT at a particular amino acid location
  - >10% in rebounders or <1-5% in naives (need databases)
  - Statistical definition (e.g. Poisson distribution)

Nomenclature

Team: HongMei Mo, Charles Boucher, Neil Parkin, Kai Lin, Jean-Michel Pawlotsky, Ann Kwong

Questions:

- What is best format for representing mutations? (genotype/gene/WT aa/position/mutant aa)
- Genotype vs. genotypic resistance; genotypic analysis vs. sequence analysis
- Mutant vs. variant
- Selected mutation vs. resistance mutation- should changes observed in the amino acid sequence in samples of patients rebounding on treatment be called a “resistance mutation” without phenotypic data?

Clonal vs population sequencing

Team: Jean-Michel Pawlotsky, Rob Ralston, Joe Fitzgibbon, Anita Howe, Tara Kieffer, Isabel Najera (leader)

Questions:

- Are they in competition or complementary?
- When best used?
- (POC studies, not phase III)
- Sampling: upon rebound, upon retreatment, after dosing, frequency
- Clonal analysis: How many clones should to be sequenced?

Clonal vs population sequencing

Team: Jean-Michel Pawlotsky, Rob Ralston, Joe Fitzgibbon, Anita Howe, Tara Kieffer, Isabel Najera (leader)

Questions:

- Effect of viral load on sampling/ amplification and representation
- Determination of mutation linkage
- Determination of in vivo replicative fitness (compare rates of growth post treatment)
- Sequence analysis: calling mixtures in population sequencing

New Sequence Analysis Technologies

Draft recommendation team:

- Robert Ralston, Anita Howe (leader)

Questions

- SAWG discussed adding a brief review the potential advantages/disadvantages and accessibility of alternate sequencing methods such as pyrosequencing, mass spec analysis, TaqMan/Primer mismatch analysis and SSLP analysis.
- This section would not be making a recommendation per say, but rather to provide information.

Validation & Implementation of Procedures Working Group (VIP WG)

The SAWG proposes that another working group be formed to write a set of detailed recommendations on:

1) validation of sequencing and phenotypic assays
2) procedures to implement appropriate quality controlled and assured protocols and processes

HCV DRAG *The creation of new resistant variants is a by product of an error-prone HCV replication process*

- The rapid replication rate of HCV, along with the poor fidelity of its polymerase, generates all possible genetic variants so that HCV exists as a quasispecies.
- This allows HCV to have great evolutionary flexibility and produce numerous viral variants with remarkable sequence variation.

Resistant variants pre-exist

- Because HCV has extremely high sequence diversity, the viral quasispecies present in subjects before treatment exist as a mixed population of master sequences (wild-type (WT)) and variants with varying levels of decreased sensitivity to a direct-acting antiviral drug. Thus, it is likely that drug resistant variants are constantly being generated in subjects every day. However, these variants often tend to be less fit than wild-type and so remain at a low frequency.
- Total body eradication requires elimination of both drug-sensitive and drug-resistant variants.
- In the absence of antiviral pressure, WT virus or similarly fit viral variants become the dominant species in a patient. In the presence of an antiviral drug, the WT virus is often the most sensitive and most rapidly inhibited. Inhibition of WT virus opens up the replication space for less fit, resistant viruses which are present before treatment at a very low level before treatment (which then become the most fit variants in the presence of drug).

Analysis of HCV sequences in patient samples is linked to viral load

- Patients infected with HCV typically have viral loads of 10^6 to 10^7 IU/mL in their plasma which translates into a total body burden of of ~10^{11} to 10^{12} IU.
- Currently, the limit of detection for HCV RNA in the most sensitive commercial assays is 10-15 IU/mL.
- Patients whose HCV load is undetectable may still have substantial amounts of virus in their body which needs to be cleared to achieve a sustained viral response (SVR).

Utility of performing sequence analysis

- There are many reasons for performing sequence analysis in HCV resistance testing, including
  - identification of mutations that confer resistance to the drug in vivo
  - analysis of the relative frequency of single and multiple mutations across patients receiving therapy
  - determination of whether these mutations confer cross-resistance to other antiviral compounds
- Important parameters that should be determined include
  - determining the frequency of variants present
  - the kinetics of their emergence
  - their persistence after dosing.
- In addition to such cross-sectional comparisons sequence analysis in longitudinal studies can be used to
  - identify variants present at baseline
  - define in vivo fitness of variants
  - characterize their relation to treatment outcome.

It is important to be able to characterize and understand drug failures

- If the sequence information from the targeted region does not explain the viral rebound observed based on previous phenotypic characterization of observed amino acid changes, additional sequences should be sequenced
- Sequence analysis of resistant variants may provide guidance for optimization of therapy, both in terms of drug combinations and duration of treatment

Focus on sequence analysis for direct-acting antivirals

- With the exception of cyclophilin inhibitors, the majority of new anti-HCV drugs which are being explored in the clinic are direct-acting antivirals (target a component of the virus or viral replication cycle); specifically NS5A binders, NS34A protease inhibitors, and NS5B polymerase inhibitors.
- At this point in time, the HCV DRAG sequence analysis working group has focused their efforts on developing recommendations for studying the resistance of HCV to viral-targeted inhibitors.
- As new information becomes available, recommendations for studying host-targeted inhibitors will be included in revised editions.

The selection and detection of new resistant variants during viral breakthrough is likely to be an indication of incomplete suppression of HCV replication

- The level of viral breakthrough and the genetic makeup of the resistant variants is a direct function of the effectiveness of the antiviral regimen (takes into account baseline resistance, PK, compliance, metabolism, plasma binding….) to suppress replication.
- In a perfect world, suppression is complete and all viral replication is stopped. In such a scenario, the only resistant variants present would be those already present before therapy at baseline.

The selection and detection of new resistant variants during viral breakthrough is likely to be an indication of incomplete suppression of HCV replication

- Less potent inhibitors will allow some degree of HCV viral replication to occur, resulting in the emergence of resistant variants via the HCV error-prone replication machinery. If treatment with an antiviral agent is not stopped once breakthrough is detected, additional resistant variants can be selected, potentially more fit and/or more resistant to drug than the original baseline variants.
- Thus, the timing of viral breakthrough and sequence analysis of the breakthrough variants can shed light on the potency of the antiviral regimen and whether new resistant variants are being created and selected in the presence of drug or whether variants already present at baseline account for the rebound.

- HCV sequence analysis has the potential to provide a window into what is happening to a patient's HCV population after they become "undetectable" by current technologies (~10 to 15 IU/mL).
- Sequence analysis of HCV variants present in a patient's virus population at baseline, on-treatment breakthrough or post-treatment relapse may provide insight into the ability of the antiviral regimen to clear virus (or not clear resistant variants) (and achieve an SVR) and the pathway(s) of development of resistance to an antiviral regimen.
- The timing of viral breakthrough and sequence analysis of the breakthrough variants also may shed light on the effectiveness of the antiviral regimen.

Sequencing of pre-treatment samples

- Baseline sequencing can reveal the potential for resistance present in the virus before treatment.
- This can be significant if the baseline quasispecies comprises a high proportion of potentially resistant variants or if a high degree of sequence diversity already exists in regions which encode domains that interact with the antiviral drug.

Sequencing of isolates obtained at virological breakthrough

- May identify the specific changes which confer or increase drug resistance.
- Prediction of the degree of drug resistance may also be possible if the phenotype conferred by specific changes has previously been established.
- Multiple resistant variants present early in breakthrough samples have the potential to increase our understanding of the different paths by which resistance can be initiated.

Sequencing of isolates obtained during relapse

- *Sequencing of isolates obtained during relapse* may reveal the presence of minor resistant variants which were not fully suppressed by the antiviral treatment regimen or the subject's immune response.

Sequencing of isolates obtained during viral breakthrough

- *Sequencing of isolates obtained during viral breakthrough* which occurs despite antiviral selection pressure should reveal how the quasispecies adapt and evolve.
- Persistance in post-treatment sequences implies that the new quasispecies is better adapted than the original; reversion of sequences to those present in pre-treatment samples implies that they are better adapted.

Sample collection and viral evolution

- The timing of sample collection for sequence analysis of drug-resistant virus from subjects is critical. In order to understand the underlying mechanisms and pathways for the generation and selection of resistant variants, it is crucial that sample collection be as close to the initial rise in HCV viral replication as possible while the virus is still under selective pressure (ie, while the subject is still on drug).
- If sample collection is delayed, sequence analysis is likely to reveal the evolution of the initial resistant variants in a subject to more fit virus such as WT virus if the patient has discontinued treatment, or to more resistant virus or more fit and resistant virus with additional genetic changes if a patient remains on treatment.

Selection of region to be sequenced

- The purpose of sequencing viral isolates is to understand the mechanism for loss of or lack of susceptibility in clinical trials of new anti-HCV drugs.
- At the end of the day, clinical virology should help characterize and understand patient failures.
- When there is a clearly defined target, e.g., NS3 protease or NS5b polymerase, sequencing of the target region should be the primary focus.

Selection of region to be sequenced

- As general rule, do the obvious things first by analyzing the suspect regions first.
- HOWEVER, if the sequence information from the targeted region (i) can not explain the viral rebound observed or (ii) the apparent loss of efficacy or (iii) if more mechanistic-like changes are suspected (such as protease cleavage sites), THEN additional regions of the genome may need to be sequenced.
- It is advisable with all new drug candidates, whether the precise mechanism or mode of action is known or not, that preclinical studies attempt to determine regions of the genome where resistance mutations are most likely to appear. For example, in cases where a clear molecular target is not known, replicon resistance studies should be performed to try to identify some of the potential routes of resistance.

Effect of VL on sensitivity

- Not all genes or gene fragments will have the same sensitivity, but all assays should share a common minimum target VL. Again, the key recommendation is for the user to establish and understand the limitations of their assay with respect to the actual VL, not to require a given VL minimum for all targets/assays.
- A key assay criterium which needs to be established is the amount of amplifiable genomes, which in turn is dependent on multiple factors including i) the genotype of the target; (ii) the length of the gene fragments of to be amplified (Table 2); (iii) to a lesser extent on the original viral load as established by quantitative assay; iv) the exact target- and genotype-specific primer sequences; v) the degree of homology between patient samples and assay conditions; and vi) the quality of the material being amplified.

Effect of VL on accuracy

- The accuracy of determination of variant frequency is dependent on the viral load of the sample.
- The standard deviation is much higher in samples with lower viral loads. This is because the actual number of RNA templates that go in to each reaction affects the estimation of frequency.
- When the number of templates in the assay is small, each variant template represents a larger percentage of the total, so larger discrete changes in the frequency estimate will occur and frequency resolution and accuracy are lower.

Population vs clonal sequencing

- The use of clonal and population sequencing methods are complementary and can answer different questions.
- To assess the development of selection of resistance upon treatment (in any phase of clinical studies), it is recommended that the initial genetic characterization be performed using population sequencing.
- If population sequencing provides clear data that can explain the “observed” rebound (or if the observed rebound can be explained by other reasons, i.e poor PK), no clonal analysis would be required.

Clonal sequencing

- Clonal sequencing is more laborious and time consuming, but can yield richer information. It should be judiciously applied early in clinical trial development of an inhibitor when sample sizes are more manageable.
- Clonal analysis should be performed to determine of the linkage of mutations in a complex resistance pathway and the contribution and role of each of the mutations. In this case, selection and sequencing of 10 to 20 clones is recommended.

Clonal sequencing

- In cases where population sequencing (or any other explanation, such as poor PK) cannot explain the viral load rebound, clonal analysis should be performed to look for minority variants before they become dominant.
- In addition, clonal analysis should be performed if population sequencing shows, and that replicon variant bearing single mutations cannot explain the magnitude of reduced drug susceptibility.

- In cases where it is important to investigate the potential existence of minority variants in the quasispecies at a frequency lower than the detection limit of the population sequence clonal sequencing methods can be used.
- The limit of detection for minority variant sequences can be increased using clonal analysis. Not withstanding important caveats such as the number of amplifiable genomes actually present in a sample, the lower limit of detection can be "dialed in" by selecting the number of clones to be

Reporting sequence analysis

- All mutations/polymorphisms present pre-therapy (at baseline) should be reported vs. the reference sequence for that genotype/subtype in table format using SAS transport files to the regulatory authorities.
- For on-therapy/longitudinal sample sequences, it is recommended that sequences be reported vs. both the reference and the patient's baseline sequence.
- Both approaches allow for the detection of mutations appearing (or disappearing) compared to pre-therapy, although the comparison within patient will be simpler.

Not all amino acid changes confer resistance

- Since HCV exists as a quasiespecies in each patient, when a patient's VL is suppressed from $10^{11\text{-}12}$ IU to close to undetectable per patient the genetic complexity of the patient's quasiespecies is reduced and undergoes an eclipse.
- When such a patient experiences a viral breakthrough or relapse, a subset of their quasie species containing changes conferring resistance to therapy is amplified in the presence of selective pressure.

Not all amino acid changes confer resistance

- That patient's HCV genome will contain both selected mutations conferring resistance as well as mutations or polymorphisms which do not confer resistance but were "carried along for the ride" with the drug-selected change.
- Thus, it is important to apply criteria to determine which mutations are relevant and confer a change in susceptibility to drug pressure or replicative fitness.

Phenotypic analysis

- Although not always available, the most powerful criteria is determination of the phenotype of a change in sensitivity to drug (cell- or enzyme-based assay) or replicative fitness (replicon or infectious-virus based assay).
- Alternatively, when phenotypic data does not exist, statistical methods can be used for active site inhibitors to identify amino acid changes which track with clinical resistance and are not simply polymorphisms in the larger baseline population.
- It should be noted that for allosteric inhibitors, it is not unusual for “baseline drug-resistant variants” to be fairly fit and pre-exist as a significant proportion of the population.

Statistical determination of potential resistance mutation using a Poisson distribution

- Sequences should be aligned and analyzed for substitutions. Potential resistance substitutions should be analyzed using criteria that consider the prevalence of the mutation in subjects treated with a drug, compared to the frequency of these mutations seen in the population (treatment-naïve).
- One method to identify mutations selected by drug treatment uses a Poisson distribution with the threshold based on a Bonferroni correction

Polymorphisms and baseline resistant variants

- Due to the high replication rate and error prone polymerase it has been calculated that all drug-selected mutations are pre-existing at some level.
- Polymorphisms can be any variant, whereas a drug-resistance mutation is a change in the population from one form (variant) to another in response to antiviral selective pressure.
- At baseline, all changes are polymorphic; whereas drug associated mutations can be broken down into those affecting resistance or fitness.

Polymorphisms and baseline resistant variants

- Drug-associated or treatment-associated mutations should not be called resistance mutations or fitness mutations unless there is phenotypic data supporting this description.
- This does not mean that treatment-associated (and selected) mutations are not important and should not be tracked, rather that one should be careful with respect to nomenclature.
- In the same vein, polymorphisms can have an impact on response without having a detectable shift in susceptibility and it was agreed that all selected mutations, regardless of whether a phenotype is assigned will be followed.
- This emphasizes the importance of defining the treatment populations's baseline spectrum so correlates of treatment outcome can be analyzed.

- There are excellent reasons for a nomenclature which is a finely detailed description for every mutation which could include virus, genotype, subtype, protein name, amino acid number, and WT vs mutant amino acid.
- On the other hand, there are also excellent reasons for a less cumbersome more fluid description including at the minimum, WT and mutant amino acid and position.
- Having a uniform nomenclature for uniform situations is critical to facillitate and enable scientific communication.
- An attempt has been made to detail situations where one type of nomenclature might be more suitable than another and on the whole, to find a balance between those who would like mutations to be exhaustively detailed and those who want ease of use

Recommendation for detailed nomenclature

- In general the recommendation for detailed nomenclature is to adapt the nomenclature recommended by the three European databases (Los Alamos, euhcvdb and Japanese) described by Kuiken et al. Kuiken, C. et al. A comprehensive system for consistent numbering of HCV sequences, proteins and epitopes. *Hepatology* **44**, 1355-61 (2006).

Example of detailed nomenclature (won vote)

HCV1a_NS3_R155K

Recommendation for abbreviated nomenclature

Example of abbreviated nomenclature:

- WT amino acid/position/mutant amino acid, eg. R155K for a HCV protease variant
  - When describing only HCV (eg not a HIV/HCV coinfection situation), it is permissible to drop “HCV”, eg. 1a_NS3_R155K
  - When describing resistant variants in a single gene in a focused research paper or when listing all changes in a protein (for example is a SAS transport file for the FDA), use the table title and headings to delinate the protein and genotype and drop the protein prefix, eg. R155K