Reviewer #1 (Comments for the Author):  
  
Summary  
This works describes the use of transcriptomics for the identification of several dehydrogenases from the genetically tractable organism, Novoshingobium aromaticivorans DSM12444, with the potential to initiate the metabolism of G-diketone. In vitro characterization supported the conclusion that several dehydrogenases tested could reduce the C2 ketone from G-diketone and GP-1. These chemicals can be produced during the chemical depolymerization of lignin, thereby expanding the potential for the microbial valorization of this renewable resource. As such, this work is of significant interest to groups that study bacterial physiology and metabolism for biotechnological applications.

We thank the reviewer for recognizing the significance of our work and the contributions of the observations described in the paper.  
  
Nonetheless, as described below, I found that there were parts of this work that were not described in enough detail and some data that should have been presented (in vitro data) were absent.   
  
We appreciate these specific suggestions to improve the clarity and impact of the work. Below we describe how we addressed these comments in the revised version of the manuscript.  
  
General Points  
1. The data from Figure 1 shows that it take ~150 hours to reach what I imagine is full density (though a clearly defined lag wasn't observed), and it takes about 75 hours for the full consumption of ~2 mM G-diketone in the presence of glucose. Doesn't this organism grow this slowly using only glucose as a sole carbon source? This seems to compare unfavorably to other genetically tractable organisms that are frequently used for the valorization of lignin-derived substrates (P. putida, E. coli, R. opacus, etc.). Are there particular stress tolerance or catabolic advantages that N. aromaticivorans has that these other organisms do not that would overshadow its slower growth rate?

The data provided in this and a previous paper (<https://doi.org/10.1074/jbc.RA117.001268>), plus others cited in the text of the original submission show that *Novoshingobium aromaticivorans* DSM12444 is not a slow growing organism. For example, Supplementary Figure S6 (formerly S5) shows growth curves of cultures grown in the presence of glucose as a sole organic carbon source and glucose + G-diketone. We have revised this supplementary figure to show both growth curves on the same time and density scales to illustrate that *N. aromaticivorans* grows rapidly on glucose. In this and previously published work, the organism grows slower in the presence of G-diketone compared to growth on glucose, which is not surprising. In addition, the growth and metabolism of aromatic diketone by *N. aromaticivorans* could be improved by subsequent genetic or other modifications. Thus, we respectfully disagree with the reviewer in the assessment that this is a slow-growing organism compared to others being used for lignin valorization.

The focus of this study was to identify *N. aromaticivorans* enzymes that metabolize aromatic di-ketones. To our knowledge, there are no other published reports on the ability of any microbe (including SYK-6 or the others mentioned) to metabolize aromatic diketones produced by some lignin deconstruction methods. Indeed, not all of the microbes mentioned by this reviewer are known to contain homologues of the aromatic dehydrogenases that we report on in this paper. We agree with both reviewers that our observations improve the possibility for microbial valorization of lignin. Our findings may enable a comparative analysis of strains or enzymes that metabolize aromatic diketones, but we feel that identifying other organisms or enzymes that metabolize aromatic diketones is an important separate study, one that is possibly enabled by our findings, and beyond the scope of this paper.

Is N. aromaticivorans unique in its ability to catabolize G-diketone or can SYK-6 do this as well? Do SYK-6 and N. aromaticivorans grow somewhat comparably or does one grow better than the other?

We focused on metabolism of aromatic diketones by *N. aromaticivorans*. We have not analyzed SYK-6 for several reasons. First, we have been unable to obtain the published SYK-6 strain for research purposes (we have not found it in public strain repositories and the material transfer agreement presented by the academic source of this strain is too onerous for acceptance by our research office and other institutions). Thus, we do not know whether SYK-6 can degrade G-diketone. To emphasize that this study provides the first analysis of G-diketone metabolism in the literature, we have added to the revised manuscript the following sentences:

L83: However, the enzymes that initiate G-diketone metabolism have yet to be identified, and to our knowledge, degradation of these aromatic diketones in other potential lignin-valorizing bacteria has not been described.

L275: *N. aromaticivorans* is the only bacterium tested to date with the ability to catabolize these diketones.

Perhaps a brief statement or two regarding the unique advantages of using N. aromaticivorans over these other organisms in the manuscript would be useful to frame the importance of this particular bug.

This is an excellent suggestion. The introduction of the revised paper reviews data from previous papers which led us to select *N. aromaticivorans* for analysis (<https://doi.org/10.1074/jbc.RA117.001268>).

L74: *N. aromaticivorans* has several advantages for studying degradation of depolymerized lignin as this bacterium can natively catabolize all three lignin monomer types (G, S, and H) (10) and can completely metabolize aromatic dimers with β-O-4 linkages (11).

2. The authors state that they use a Benjamini-Hochberg false discovery rate corrected p-value. However, in Figure 2, the authors state that significance was called using different p-value thresholds. I imagine that this is a typo and these are p-values corrected for multiple testing by the Benjamini-Hochberg method? If so, please change these to 'q-values'. Additionally, I prefer to use 'q-value' here, as opposed to 'FDR', as used in legend of Figure S3, since the q-value was defined by Storey in 2003 during description of the positive FDR method (pFDR) as the analog to the p-value, where p-value <0.05 defines a false positive rate of 5%, a q-value < 0.05 is meant to represent the similar false discovery rate (FDR) of 5%. Based upon how you use the q-value cutoff of 0.05, I take it that you are actually using the pFDR method, where q<0.05 represents a 5% False discovery rate. Please cite Storey JD, 2003 "The Positive False Discovery Rate: A Bayesian Interpretation and the q-Value" along with the original Benjamini-Hochberg paper, if the authors are not using this method, they need to provide p-values in the supplemental dataset (and it would be nice if they provided it even if they used the pFDR method). The original Benjamini-Hochberg FDR procedure involves comparing the p-value to the critical value - (r/t)Q--, determined from the p-value rank (lowest to highest) (r), total number of tests (t), and false discovery rate (Q) and then selecting the largest p-value that is smaller than the critical value as the significant call cutoff.

We thank the reviewer for correcting our terminology and pointing us towards the appropriate citations. We used the Benjamini-Hochberg FDR correction as implemented in the R package edgeR. Following the reviewer’s suggestion, we now use the q-value terminology in the revised manuscript instead of p-values or FDR, have added the Storey citation to the methods section, and now include p-values and q-values in the Supplemental Dataset S1. The following sentence was added to the methods section of the revised manuscript:

L471: “The software edgeR version 3.26.8 (35) was used to identify significantly differentially expressed genes from pairwise analyses based on q-values, using a Benjamini and Hochberg false discovery rate (FDR) less than 0.05 as a significance threshold (35–37).”  
  
3. Figure 4 shows GC-MS identification of products and substrates from LigD, LigN, and LigL reaction with either G-diketone or GP-1. However, Figure 3 only shows catalytic data for G-diketone substrate. Can the authors please add to figure 3 the catalytic data for GP-1 as substrate? Additionally, can the authors please add to the supplemental materials the reaction velocity vs. substrate concentration graphs for each enzyme on each substrate (GGE, G-diketone, threo-GD) that was used to make Km and kcat calculations? Also, please report all Km measurements in millimolar as opposed to M.

We have added data from using GP-1 as a substrate to Figure 3 in the revised manuscript. The text was modified to refer to the udpdated information from these new experiments:

L203: GP-1 was not stable in the 24-hr incubations with the Lig dehydrogenases in the presence or absence of NADH (Figure 3B), but when the products of GP-1 transformation were assayed with GC-MS, both threo- and erythro-GD were detected only in experimnts with LigL (Figure 4).

We have added the primary data used to generate apparent Km and *k*cat calculations to the revised manuscript as a supplementary figure (new Supplementary Figure S5).

In the revised manuscript, we report the apparent Km values in millimolar.

The revised legend for Figure 3 is the following:

**Figure 3. Time-dependent loss of G-diketone and GP-1 *in vitro* when incubated with recombinant LigL, LigN, and LigD,** **with and without NADH.** 0.5 mmol/L of G-diketone (Panel A) or GP-1 (Panel B) was incubated with each in enzyme with and without 2 mmol/L NADH. Addition of NADH initiated the reaction. Samples were incubated in the dark at 30°C. Concentrations of G-diketone and GP-1 was measured using HPLC-MS.

The legend for the new Figure S5 is the following:

Figure S5: Reaction velocity vs. substrate concentration used to calculate Km and kcat values for LigL, LigN, and LigD on GGE, G-diketone, and GD.  
  
4. The authors looked at growth effects from knocking out each gene individually and didn't see a significant result (though WT in the G-diketone and Glucose condition showed a large amount of variation between replicates). However, what I would be more interested in is whether the authors saw decreased G-diketone consumption of each individual mutant. Do the authors have these data? Also, have the authors tried any combinations of double or triple mutants, etc. yet?

We do not have data on G-diketone consumption of each individual mutant, nor have we made combinations of strains that lack multiple aromatic dehydrogenases. G-diketone is not commercially available, so it must be custom synthesized. This makes performing the experiment requested prohibitively time-consuming to perform, especially when working under local COVID-19 guidelines. The requested study plus a detailed analysis of G-diketone metabolism in double or triple mutants would best be part of a separate study. No modifications were made to the manuscript in response to this comment.

Specific Points  
Line 131-I am unclear on what is meant by 'similar' here... Does this mean to say that the spectra match, or how different were the calculated masses? The authors provide an MS spectrum for this unknown species at 14.81 min in Figure S2. Could the authors perhaps also include the previously published spectrum for GP-2 alongside their spectrum? In such situations, I think it best to tone down use of concrete words like 'confirming' in place of something like 'supporting the conclusion that this isomer of GP-1 also accumulated in the media of G-diketone grown cells'

We thank the reviewer for this comment, and have modified the text to more accurately report our findings. We also added the published spectrum for GP-2 from Mitchell (2014) to Figure S2 and cited this reference as the source of this data in the revised materials. We defer to the journal and ASM publications to determine if this is allowed under existing policy on including images from other publications in our manuscript. We have also modified the following to use the language suggested by the reviewer:

L133: The GC chromatograms showed the presence of multiple species, and MS spectra comparison suggested the presence of GP-1 and threo-GD in these samples (Figure S2). Furthermore, an additional GC species had an MS spectrum matching the previously published spectrum of GP-2 (1), potentially indicating that this isomer of GP-1 also accumulated in the media of G-diketone grown cells.

Figure S2:We have also included the published spectrum for GP-2 from Mitchell (2014) (1) for reference to our proposed identification of GP-2.

Line 143-If performed, I am curious what the measured amounts of vanillin/vanillate were in the SMB + Glucose + G-diketone/GP-1 medium when not inoculated with cells but subjected to the same growth and sampling protocol (no-cell control).

For reasons cited in our response to comment #4 of this reviewer, we did not perform this experiment.  
  
Line164-Please correct to 'these data'

Corrected, thanks for catching this typographical error.   
  
Lines 171-173-What is meant here by 'transcription unit'. Do they mean to say complex between RNA pol, its substrate, and associated factors or do they mean to focus on distinct transcriptional factors and their targets (regulon)? If the latter, I prefer the term regulon. Also, if the latter, what evidence from their organism do the authors have that co-localization (synteny) dictates the likelihood that genes are regulated via a common transcription factor or not? I think that it is one thing to say that genes are not expressed as part of an operon or that they are not co-localized and quite a different thing to say that they therefore are likely regulated as part of several discrete regulons.

The term “transcription unit” is traditionally used to describe the number of genes (one or more) under control of a promoter. We used recently published transcriptional start site data from this bacterium (<https://doi.org/10.1128/MRA.00880-20>) to place co-transcribed genes in a transcription unit.

The term “regulon” is traditionally used to describe the number of “transcription units” that are under direct control by a given DNA binding protein.

To prevent confusion, we revised the manuscript to clarify this point. We thank the reviewer for stimulating us to clarify this point. The edited sentence is the following:

L180: Most of these *lig* genes are not co-localized on the genome (Figure S4), so the increase in abundance of these transcripts likely results from the activity of multiple promoters when cells are grown in the presence of either G-diketone or GP-1.  
  
Line206-Do these enzymes have more optimal or worse Kms or kcats than their counterparts in SYK-6? If so, this could be a potential point to bring up in the Discussion

The focus of this study was to identify *N. aromaticivorans* enzymes that can process aromatic di-ketones. For reasons outlined above, the objective of this work was not to perform a comparison with other organisms. Therefore, the text has not been changed in response to this comment.  
  
Line 369-How many biological replicate cultures were grown?

We have provided the requested information in the following, edited sentence:

L403: To assay metabolism of G-diketone, triplicate cultures of *N. aromaticivorans* were grown overnight with 1 g/L D-glucose (Sigma-Aldrich, St. Louis, MO, USA)before adding an equal volume of SMB containing 1 g/L glucose and incubating for one hour.

Line 386-Can the authors provide what the theoretical COD equivalent to 0.5g/L ended up being for each aromatic substrate?

We have provided the requested information in the following edited sentence:

L418: Aromatic substrates used were PCA, vanillic acid, vanillin, ferulic acid (Sigma-Aldrich, St. Louis, MO, USA)*,* GP-1 (Key Organics, Camelford, UK), and G-diketone, synthesized as previously described (10). Amounts of aromatic substrate in the culture were normalized to have a theoretical COD equivalent to 0.5 g/L (amount added to each culture: ferulic acid - 0.29 g/L, vanillin - 0.28 g/L, vanillic acid – 0.33 g/L, G-diketone – 0.29 g/L, GP-1 – 0.27 g/L, PCA – 0.37 g/L (Supplemental Dataset S1)).

Line 469- How much NADH was used? What temperature were the reactions run at?

We have provided the requested information in the following edited sentences:

L517: These purified enzymes were used for *in vitro* enzymeassays using 0.5 mM G-diketone as a potential aromatic substrate along with 1μM of enzyme with or without 2mM NADH (Sigma-Aldrich, St. Louis, MO, USA) as a cofactor, in a buffer containing 25 mM Tris-HCl (pH 8) and 25 mM NaCl (Sigma-Aldrich, St. Louis, MO, USA). An additional control was run with no enzyme added to assess spontaneous degradation of G-diketone. Assays were incubated at 30°C in the dark. In some cases, concentrations of G-diketone were measured at 0, 0.5, 1.5, 2.5, and 24 hours using HPLC-MS. We tested for the presence of additional reaction products by GC-MS using material from the 24-hour timepoint.

Spectrophotometric NADH oxidation/NAD+ reduction assays were performed using LigL, LigN, and LigD in the presence of G-diketone, GP-1, GGE, or threo-GD (BioCrick Co. Ltd., Chengdu, Sichuan, China). Assay conditions were set as described above at lower concentrations to remain within accurate range of the spectrophotometer (0.025-0.2mM substrate, 0.1mM co-factor, and 0.1μM enzyme (0.01μM for LigL with G-diketone due to increased reaction velocity)), with NADH or NAD+ added immediately prior to measurement of optical density at 340 nm on an Olis DW-2000 spectrophotometer (OLIS, Inc., Athens, GA, USA)

Figure 1- I prefer that figure legends contain enough information for me to understand the experiment. Please list the growth medium (SMB), the concentration of glucose (0.5 g/L) and G-diketone (0.418 g/L) used, how many biological replicates there were (something I don't think is stated anywhere in the manuscript for this experiment), and what the error bars represent (SEM, SD, 95% CI?)

We have provided the requested information in the following updated figure legend:

Figure 1 legend: Triplicate cultures were grown in SMB with 0.5 g/L glucose and 0.418g/L G-diketone. In each panel, error bars represent the standard deviation. (Panel A) Increases in *N. aromaticivorans* cell density as monitored by Klett colorimeter units. (Panel B) Extracellular concentrations of G-diketone, GP-1, and threo-GD identified and quantified via HPLC-MS and HPLC-UV (Figure S1, see text).

Figure 3- Should the y-axis be labelled 'mmol/L G-diketone' instead of just diketone? I would also like a bit more information pertaining to the experiment listed in the figure legend. E.g. how much substrate was added? How much enzyme? How was the reaction initiated? How was mmol/L G-diketone measured? You don't need to go into the detail you do in the Methods section, but a couple sentences of detail for a reader to understand the experiment by just looking at the Figure, without needing to dig through the methods would be very appreciated.

In the revised manuscript, the y-axis title has been changed to “G-diketone (mmol/L)”. We have also modified the Figure 3 legend as follows:

Figure 3: 0.5 mmol/L of G-diketone (Panel A) or GP-1 (Panel B) was incubated with each in enzyme with and without 2 mmol/L NADH. Addition of NADH initiated the reaction. Samples were incubated in the dark at 30°C. Concentrations of G-diketone and GP-1 was measured using HPLC-MS.

Figure 4- again, you don't need to go crazy here, but I would like the description to include 'incubated 24 hr at <what> temperature with G-diketone and NADH'

We have provided the requested information in the modified figure legend:

Figure 4: GC-MS analysis of derivatized aromatic substrates and enzyme reaction products after indicated Lig dehydrogenases were incubated for 24 hours at 30°C with the G-diketone and NADH (Panel A, C) or GP-1 and NADH (Panel B, D).

Figure S2-Wouldn't trimethylchlorosilane also be expected to derivatize the aromatic hydroxyl group from GP-2, erythro-GD, threo-GD, and GP-1? If so, can the authors please change the depicted structure accordingly.

That is correct. We thank the reviewer for this question and we modified Figure S2 in the revised manuscript.

Table S2-I won't go so far as to say that if the authors don't have the data that they need to perform the experiment, but a thorough control for this experiment would have been the quantification of vanillic acid and vanillin in the absence of G-diketone and GP-1 (presumably showing none), to show that these only appear in the presence of either aromatic. If the authors have the data from such a control, I will just ask that they add it to Table S2. As I wrote above, quantification of vanillin/vanillate in G-diketone/GP-1 medium without cells would also be a helpful control to add here, since you say that the pure preps of G-diketone contained slight amounts of contamination. Also, would it make sense to list concentrations in micromolar as opposed to M?

We have provided the requested information as a third row in Table S2 showing our analysis of vanillic acid and vanillin in cultures containing glucose only, with G-diketone and GP-1 absent. We have also revised the concentrations to micromolar.

Table S4 - Strains listed in this study should include strain number designations

We have updated the table to include the strain number designations in the strains.

Reviewer #2 (Comments for the Author):  
  
The authors describe experiments to determine the protein responsible for metabolism of Hibbert ketones produced by chemical pretreatment of lignocelluosic biomass. This area has much activity in it and the conversion of depolymerized lignin to products is valuable for valorizing plant biomass.   
  
I have some concerns about the interpretations of the data presented in manuscript. The authors demonstrate that the metabolism of N. aromaticivorans produces reduced forms of the Hibbert ketone and upregulated dehydrogenases are capable of reducing these ketones. However, it is not clear that these are the in vivo enzymes that perform this reaction. As the authors describe the model in the Discussion, there are many unknown enzymatic reactions in the conversion of the Hibbert ketones. Therefore, it is possible that there are other dehydrogenases that may be performing this reaction.

The reviewer is correct that our experiments cannot rule out the existence of additional dehydrogenases or other enzymes that participate in G-diketone metabolism. We have modified the text to recognize this possibility. The revised manuscript now includes the following sentences:

L323: Strains containing single deletions of *ligLNDO* did not have a fitness defect in the presence of G-diketone; therefore, while our *in vitro* assays suggest that these dehydrogenases can complement functions, we cannot rule out the presence of other enzymes *in vivo* that may also be able to reduce G-diketone.

L349: We also acknowledge that our experiments do not preclude the presence of additional enzymes with redundant functions in our proposed model and are primarily based on in vitro data, as strains containing individual deletions of each aromatic dehydrogenases showed no fitness defect in comparison to the parent strain (Figure S6).

L359: As enzyme redundancy is a described characteristic of *N. aromaticivorans’* dimer degradation pathway (12, 17), it is also possible that additional, undiscovered enzymes can also reduce diketones in addition to the Lig dehydrogenases.

What particularly concerns me is that the deletion studies described at the end of the Results (lines 229-231). The authors report that these deletions do not have any defects, However, these mutants could be used to look at the profiles of compounds in the supernatant and if they are altered. Without this information, the assertions in the manuscript are somewhat weak. The authors should acknowledge this in their revision.

The deletion studies referenced above were designed to monitor the impact of individual gene deletions on growth in the presence of G-diketone. We would ultimately like to perform other tests that could be done with these and additional mutant strains, but because the G-diketone must be custom synthesized, we do not have sufficient materials to perform the additional experiment suggested by the reviewer (see our response to comment 4 of reviewer 1).

However, the reviewer’s point is valid, and we have modified the text to clarify what can and cannot be concluded from the data that is available on this set of mutant strains. The following sentences are now included in the revised manuscript:

L327: (Following previous revision in L323-327): We can also not conclusively demonstrate that LigL, LigN, and LigD perform the same functions *in vivo* as we observed *in vitro.*

Based on the results of our in vitro analyses of these aromatic dehydrogenases, we propose a model for metabolism of G-diketone by *N. aromaticivorans* (Figure 6)*.*

L349: (also included in previous comment) We also acknowledge that our experiments do not preclude the presence of additional enzymes with redundant functions in our proposed model and are primarily based on in vitro data, as strains containing individual deletions of each aromatic dehydrogenases showed no fitness defect in comparison to the parent strain (Figure S6).