Research Note

Microbial Quality of Bagged Baby Spinach and Romaine Lettuce: Effects of Top versus Bottom Sampling

JULIE A. KASE,1* STACEY BORENSTEIN,2 ROBERT J. BLODGETT,3 AND PETER C. H. FENG1

1Division of Microbiology, Office of Regulatory Science, and 2Division of Public Health and Biostatistics, Office of Food Defense, Communication, and Emergency Response, U.S. Food and Drug Administration, College Park, Maryland 20740; and 3Oak Ridge Research Institute for Science and Education, Oak Ridge, Tennessee 37830, USA

MS 11-097: Received 28 February 2011/Accepted 21 August 2011

ABSTRACT

Contamination with Escherichia coli O157:H7 and Salmonella have called into question the safety and microbial quality of bagged ready-to-eat leafy greens. This study expands on previous findings that these goods have high total bacteria counts (TBC) and coliform counts, variation in counts among different lots, that Escherichia coli is present, and disparities in counts when bags are top or bottom sampled. Nearly 100 bags of baby spinach and hearts of romaine lettuce from a single brand were subjected to both top and bottom sampling. Product was blended, and a portion serially diluted and plated to obtain TBC. Total coliform and E. coli levels were estimated by the most-probable-number (MPN) technique with ColiComplete discs. Top-sampled TBC from bags of baby spinach (48 bags, 13 different lots) ranged from 3.9 to 8.1 log CFU/g and bottom-sampled TBC ranged from 4.0 to 8.2 log CFU/g, with 52% of the bags (or 39% of the lots) producing TBC higher in bottom samples. For hearts of romaine (47 bags from 19 different lots), top-sampled bags had TBC ranging from 2.4 to 7.0 log, and bottom-sampled bags had TBC from 3.3 to 7.3 log, with 64% of the bags (or 63% of the lots) showing higher TBC in bottom samples. However, we are unable to reject the hypothesis that the top and bottom samples from either commodity contain the same TBC (P ≥ 0.08). No E. coli was detected and total coliform bacteria counts were, with few exceptions, ≤210 MPN/g, irrespective of TBC. In general, lots with the most number of days before the printed “use-by” date had lower TBC. However, the R2 values for either baby spinach (0.4085) or hearts of romaine (0.2946) suggest that age might not be a very good predictor of higher TBC. TBC varied widely between lots and even more so within same-lot samples, as indicated by the sum of squares results. This finding, along with higher TBC in bottom samples, suggests further consideration when a microbiological sampling scheme of bagged produce is designed.

Bagged leafy green commodities offer a convenient and presumably healthy way of increasing one’s consumption of fresh produce. Sales of ready-to-eat (RTE) salad are increasing, with several billions of dollars of product sold every year in the United States (6, 9). However, recent outbreaks involving Escherichia coli O157:H7 (2006, bagged spinach), E. coli O145 (2010, bagged, shredded romaine lettuce), and product recalls involving Salmonella (2009, loose spinach) and E. coli O157:H7 (2010, bagged baby spinach) have called into question the hygienic quality of leafy greens (4, 5, 19, 20). Both of the E. coli O157:H7 outbreaks involved products grown in the Salinas Valley of California, where approximately 80% of the leafy greens sold in the United States are grown (17). In between the 2006 and 2009 outbreaks and in the absence of U.S. government standards, voluntary, state-level marketing agreements in California and Arizona were reached, and they now cover over 90% of the leafy greens grown in either state (3, 18). Recently, a National Leafy Greens Marketing agreement that closely mimics the state agreements was proposed by the U.S. Department of Agriculture’s (USDA’s) Agricultural Marketing Service (18). The focus is on good agricultural practices such as water quality and shielding field contamination from wildlife or surface runoff, postharvest processing, and maintenance of proper sanitation and condition through distribution. Compliance is monitored through audits conducted by a USDA-authorized inspector, but product testing for indicator organisms or pathogenic organisms is not addressed (18).

RTE bagged leafy greens are of a particular concern, given that they are marketed as thoroughly washed, and consumers often consume these products directly out of the bag, without further treatment. Previous studies using indicator organisms (e.g., E. coli, and total and fecal coliform bacteria) have indicated that any concerns on sanitary quality are justified. Recently, the Consumer Union tested 208 bags including 16 different brands of salad greens collected in the New York City Metro area over a period of 8 days in the summer of 2009. While no pathogens were found (E. coli O157:H7, Salmonella, or Listeria), 39% of the bags tested had more than 10,000 CFU/g total coliforms, and approximately 5% of bags contained E. coli (7). Moreover, a previous study found highly variable bacterial contamination with E. coli O157:H7 and Salmonella have called into question the safety and microbial quality of bagged ready-to-eat leafy greens. This study expands on previous findings that these goods have high total bacteria counts (TBC) and coliform counts, variation in counts among different lots, that Escherichia coli is present, and disparities in counts when bags are top or bottom sampled. Nearly 100 bags of baby spinach and hearts of romaine lettuce from a single brand were subjected to both top and bottom sampling. Product was blended, and a portion serially diluted and plated to obtain TBC. Total coliform and E. coli levels were estimated by the most-probable-number (MPN) technique with ColiComplete discs. Top-sampled TBC from bags of baby spinach (48 bags, 13 different lots) ranged from 3.9 to 8.1 log CFU/g and bottom-sampled TBC ranged from 4.0 to 8.2 log CFU/g, with 52% of the bags (or 39% of the lots) producing TBC higher in bottom samples. For hearts of romaine (47 bags from 19 different lots), top-sampled bags had TBC ranging from 2.4 to 7.0 log, and bottom-sampled bags had TBC from 3.3 to 7.3 log, with 64% of the bags (or 63% of the lots) showing higher TBC in bottom samples. However, we are unable to reject the hypothesis that the top and bottom samples from either commodity contain the same TBC (P ≥ 0.08). No E. coli was detected and total coliform bacteria counts were, with few exceptions, ≤210 MPN/g, irrespective of TBC. In general, lots with the most number of days before the printed “use-by” date had lower TBC. However, the R2 values for either baby spinach (0.4085) or hearts of romaine (0.2946) suggest that age might not be a very good predictor of higher TBC. TBC varied widely between lots and even more so within same-lot samples, as indicated by the sum of squares results. This finding, along with higher TBC in bottom samples, suggests further consideration when a microbiological sampling scheme of bagged produce is designed.

Bagged leafy green commodities offer a convenient and presumably healthy way of increasing one’s consumption of fresh produce. Sales of ready-to-eat (RTE) salad are increasing, with several billions of dollars of product sold every year in the United States (6, 9). However, recent outbreaks involving Escherichia coli O157:H7 (2006, bagged spinach), E. coli O145 (2010, bagged, shredded romaine lettuce), and product recalls involving Salmonella (2009, loose spinach) and E. coli O157:H7 (2010, bagged baby spinach) have called into question the hygienic quality of leafy greens (4, 5, 19, 20). Both of the E. coli O157:H7 outbreaks involved products grown in the Salinas Valley of California, where approximately 80% of the leafy greens sold in the United States are grown (17). In between the 2006 and 2009 outbreaks and in the absence of U.S. government standards, voluntary, state-level marketing agreements in California and Arizona were reached, and they now cover over 90% of the leafy greens grown in either state (3, 18). Recently, a National Leafy Greens Marketing agreement that closely mimics the state agreements was proposed by the U.S. Department of Agriculture’s (USDA’s) Agricultural Marketing Service (18). The focus is on good agricultural practices such as water quality and shielding field contamination from wildlife or surface runoff, postharvest processing, and maintenance of proper sanitation and condition through distribution. Compliance is monitored through audits conducted by a USDA-authorized inspector, but product testing for indicator organisms or pathogenic organisms is not addressed (18).

RTE bagged leafy greens are of a particular concern, given that they are marketed as thoroughly washed, and consumers often consume these products directly out of the bag, without further treatment. Previous studies using indicator organisms (e.g., E. coli, and total and fecal coliform bacteria) have indicated that any concerns on sanitary quality are justified. Recently, the Consumer Union tested 208 bags including 16 different brands of salad greens collected in the New York City Metro area over a period of 8 days in the summer of 2009. While no pathogens were found (E. coli O157:H7, Salmonella, or Listeria), 39% of the bags tested had more than 10,000 CFU/g total coliforms, and approximately 5% of bags contained E. coli (7). Moreover, a previous study found highly variable bacterial

* Author for correspondence. Tel: 301-436-2923; Fax: 301-436-2644; E-mail: julie.kase@fda.hhs.gov.
counts in commercially available bagged varieties of lettuce and spinach mixes (21). An additional finding was the preliminary observation of higher counts for samples taken from the bottom of the bags versus the top. The authors suggested that this might be partly explained by condensation moisture noted at the bottom of many bags, which might provide a better opportunity for microbial growth. However, the actual effect of bottom-only sampling was not determined, as the same bag was not both top and bottom sampled (21). As a follow-up to those findings and to better assess the homogeneity of bacterial content within a bag, the current study sampled nationally distributed bagged cut spinach and lettuce mixes, with all bags opened and sampled both at the top and at bottom.

MATERIALS AND METHODS

Types and origin of produce tested. Forty-eight bags of baby spinach and 47 bags of hearts of romaine lettuce were purchased between June and October 2009 from stores in the Washington, DC, metropolitan area. All bags were of a single, market-leading national brand and labeled ‘‘thoroughly washed.’’ If not tested on receipt, bags were stored at 4°C and sampled within 24 h.

Sample preparation and method for total bacterial count (TBC). All bags were tested on or before the recommended used by date printed on each bag (with one exception). Separate but equal sized samples were obtained from the top and bottom of each bag and analyzed for total coliform and E. coli bacterial counts to assess their microbiological quality. The total count of the samples was determined with the method described in the U.S. Food and Drug Administration’s (FDA’s) Bacteriological Analytical Manual (10). In brief, 50 g of the product was blended with 450 ml of buffered peptone water (prepared in-house (10)), from which 1:10 serial dilutions were made. TBC were done by standard plate count, where 0.1 ml of each dilution was spread plated in duplicate on Trypticase soy agar (TSA; BD, Franklin Lakes, NJ).

Bacteria identification. To assess the types of bacteria that are present in bagged produce, approximately 20 colonies obtained from different samples were identified with the automated bacterial identification systems Vitek 2 Compact (bioMérieux, Inc., Hazelwood, MO) and Biolog (Biolog, Hayward, CA). To prepare for analysis by using the Vitek 2 GN Colorimetric Identification Card, a sterile swab was used to transfer a colony from the TSA plate to 3.0 ml of sterile saline in a clear plastic (polystyrene) test tube. The turbidity is measured with a turbidity meter and adjusted per manufacturer specification. For analysis with the Biolog GN2 MicroPlate, manufacturer instructions were followed (www.biolog.com/pdf/GN2b_Brochure.pdf).

Total coliform and E. coli counts. The total coliform and E. coli counts were determined with the most-probable-number (MPN) method by using ColiComplete discs (Biocontrol, Bellevue, WA) to assess β-galactosidase (total coliform) and β-glucuronidase (E. coli) activities. Specifically, after 48 h at 37°C, a combination of total coliform–positive (blue) and E. coli–positive (fluorescence) tubes was used to estimate the levels of each indicator from the MPN table (10).

Data analysis. Sample suspensions were diluted for counting colonies. However, instead of ignoring plate counts outside of an acceptable range and potentially introducing bias into the estimation of concentration, total bacterial plate counts for a particular sample were combined as follows: For all plates within a countable range (approximately <250 colonies), their colony count was divided by the sum of the mass of all their inoculants; for plates with colonies deemed ‘‘too numerous to count’’ (TNTC) (approximately >250 colonies), a mathematical procedure as detailed by Blodgett was used (1). In brief, a lower bound was estimated for each TNTC plate, based on the counts closest to 250 obtained from the plates. This value was divided by its volume of the initial sample to generate the lower-bound ratio and was then compared with a preliminary ratio. This preliminary ratio was calculated by summing all the non-TNTC counts and dividing this sum by the volume of their initial samples. If a lower-bound ratio was reached that was larger than the preliminary estimate ratio, the lower bound amount was included in the bacterial count calculation. Zero count plates were also included in TBC calculations.

A rarity index, a statistical diagnostic tool designed for microbiological experiments, was used to account for data variability possibly associated with laboratory error or perhaps aggregation of bacteria, and to separate that from true differences in bacterial counts. The rarity index equals the probability of the actual outcome divided by the probability of the most probable outcome. Blodgett provides further explanation and mathematical proof (2). For the current study, bags were grouped into three categories based on the calculated rarity index ratio: not explained by dilution alone (ratio near zero), mostly explained by dilution (ratio near one), or in between these two categories.

Statistical analysis. The sign test was used to test the hypothesis that there is no difference between the values obtained from the top and bottom of individual bags. Data analyzed with the sign test paired counts obtained for individual top and bottom samples. P values were calculated and evaluated. A nonparametric test like the sign test was appropriate, given the high variability in the data and lack of fit to a normal distribution. For a comparison of the amount of variability within a particular lot to the amount of variability between lots, the sums of squares between the lots and the sum of squares within the lots were calculated.

RESULTS

Consistent with the data from the previous study (21), total bacteria counts varied widely between lots and in same-lot samples. Top-sampled TBC from bags of baby spinach (48 bags, 13 different lots) ranged from 3.9 to 8.1 log CFU/g, and bottom-sampled TBC ranged from 4.0 to 8.2 log CFU/g, with 52% of the bags (or 39% of the lots) producing TBC higher in bottom samples. For hearts of romaine (47 bags from 19 different lots), top-sampled bags had TBC ranging from 2.4 to 7.0 log and bottom-sampled bags had TBC from 3.3 to 7.3 log, with 64% of the bags (or 63% lots) showing higher TBC in bottom samples. Interestingly, for baby spinach, the same bag produced the lowest TBC for top and bottom samples, as was the case for the highest TBC counts for the top and bottom portions. However, the opposite was true for hearts of romaine. Scatterplots (shown in Figs. 1 and 2) offer a graphical representation of the differences between top and bottom amounts for each bag sampled. The greater the agreement between top and bottom samples from the same bag, the greater the points concentrate around the identity line (y =
Examination of the raw counts for each dilution plate suggested that other factors beyond dilution (i.e., bacterial aggregation or analyst error) might be influencing the results in the cases where counts increased with increasing sample dilution. But, as Table 1 indicates, the quantity of bags where the counts can be explained by dilution does not seem to differ from those bags where the counts are not explained by only dilution in terms of the number of bags with higher top-sample counts. The largest difference was seen with the hearts of romaine bags, which could be explained by dilution. In this case, only 4 of the 16 bags had counts where the top sample exceeded the bottom sample. However, with the sign test, the difference between top and bottom was not significant ($P \approx 0.0768$). This was true even if all 47 bags were combined ($P \approx 0.0789$). Analysis on all combinations of the 48 bags of baby spinach produced $P$ values close to 1 (or exactly 1 in the case of the 9 bags, with top counts highest), disallowing the rejection of the hypothesis that the top and bottom samples contain the same TBC.

The amount of variability within a particular product lot and the amount of variability between lots was also examined by calculating the sum of squares. For baby spinach, the sum of squares between the lots was $1.16 \times 10^{16}$, and the sum of squares within the lots was $1.44 \times 10^{13}$. For romaine, the sum of squares between the lots was $5.97 \times 10^{13}$, and the sum of squares within the lots is $9.41 \times 10^{13}$.

Many varieties of bacteria colonies were observed on TSA after plating and incubation. Various colonies chosen from different samples were characterized. The most frequently identified bacteria were *Bacillus subtilis*, *Pseudomonas fluorescens*, *Pantoea agglomerans*, and *Sphingomonas paucimobilis*.

### TABLE 1. Rarity index results showing the grouping of the 48 bags of baby spinach and the 47 bags of hearts of romaine lettuce

<table>
<thead>
<tr>
<th>Greens type</th>
<th>Baby spinach</th>
<th>Romaine lettuce hearts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not explained by dilution alone</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>Possibly explained by dilution</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Mostly explained by dilution</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>

*For each bag, individual plate counts were assessed in terms of expected decrease in counts with increasing dilution and bags categorized accordingly.*

### FIGURE 1. Scatterplot showing the comparison of top versus bottom sampling of 47 bags of hearts of romaine lettuce. Log CFU values for top and bottom samples are presented for each sample. An identity line is drawn for reference. For example, if top and bottom values were equal, the data point would fall on the line.

### FIGURE 2. Scatterplot showing the comparison of top versus bottom sampling of 48 bags of baby spinach. Log CFU values for top and bottom samples are presented for each sample. An identity line is drawn for reference. For example, if top and bottom values were equal, the data point would fall on the line.

### FIGURE 3. Scatterplot showing the influence of "age of product" on overall log CFU averages for each bag of hearts of romaine tested. Least-squares regression lines were found by using the equation $\log \text{CFU} = 5.4114 - 0.17161 \times \text{number of days before the use-by date}$. $R^2 = 0.2946$.
No *E. coli* was detected in any sample and total coliform bacteria counts were, with few exceptions, $\geq 210$ MPN/g, irrespective of TBC. The age of the product (as judged by the days until recommended use-by date) was also examined as a predictor of higher TBC. Figures 3 and 4 show overall log CFU amounts for each of the bags tested in this study. In general, the lots with the most number of days before the printed use-by date had lower TBC (as indicated by the negatively sloping regression line). However, the $R^2$ values for either baby spinach (0.4085) or hearts of romaine (0.2946) suggest that age might not be a very good predictor of higher bacterial counts.

**DISCUSSION**

The microbial quality of RTE bagged leafy greens are of a particular concern given that they are often labeled as “thoroughly washed” or “triple washed”; as a result, consumers often consume these products directly out of the bag without further treatment. In fact and despite recurring recalls, there is no guidance from public health agencies to wash further such products because of the lack of documented evidence of washing efficacy.

The results of this study confirmed that these products contain high TBC levels, which varied widely between lots and even more so within same-lot samples, as indicated by the sums of squares results, lending credence to the belief that these products are difficult to render microbe free with washing. Despite presumed changes in industry practices in the last decade, total bacteria levels seem to have changed little; a previous retail survey of packaged salads found total mesophilic bacteria plate counts ranging from $2 \times 10^5$ to $8 \times 10^6$ CFU/g (12), similar to our study, wherein levels ranged from $2 \times 10^5$ to $2 \times 10^6$ CFU/g.

Previous findings by Luo et al. (14) noted increases in native microflora and *E. coli* O157:H7 after storage at $\geq 8{^\circ}C$ during the labeled shelf life of bagged baby spinach. In this study, there was a noted absence of moisture condensation in the bags, suggesting proper product handling and storage or perhaps a change in industry practices to better control the amount of moisture in the bags that was noted in the previous study (21). Moisture enters bags as a residual from product washing or possibly as condensation of the water vapor trapped within the polyethylene film bag. Time and temperature abuse has been noted to occur sometimes within the retail environment, but it is difficult to document. The uncertainties of this variable could explain the difficulties we experienced in this study in trying to attribute higher TBC to older products (Figs. 3 and 4).

*B. subtilis, P. fluorescens, P. agglomerans,* and *S. paucimobilis* were the commonly identified bacteria associated with both baby spinach and hearts of romaine lettuce. All these organisms are ubiquitous in nature and they can be isolated from soil, water, and plants, and usually do not cause disease in a healthy human. Previous studies have recovered these same organisms on minimally processed produce such as green peppers, cabbage, and lettuce (8). While the presence of these organisms is expected and largely harmless to consumers, many of the microflora can either stifle or enhance the survival and/or colonization of human pathogens like *E. coli* O157:H7, *Listeria monocytogenes,* and *Salmonella* Montevideo on leafy produce (8).

Hygienic standards for RTE bagged leafy greens exist in other countries, but they are not currently used in the United States. Moreover, hygienic standards for bagged, RTE, leafy greens are unlikely to be enforced in the United States within the future. In some cases, standards are set with indicator organisms such as fecal and total coliforms. A recent retail survey of leafy salad items in Brazil revealed that nearly three-fourths of the products tested exceeded the Brazilian standard for fecal coliforms (maximum of 100 CFU/g) (11). In the same study, only 3% of the samples tested positive for *Salmonella;* however, half of these did not exceed the 100-CFU/g fecal coliforms limit, thus raising concerns over the value of fecal coliforms as reliable indicators of pathogen presence. Similarly, *E. coli* is frequently used as an indicator of fecal contamination, as it is commonly found in the lower intestines of animals. However, like coliform bacteria, it can be isolated in nature from nonfecal sources. A recent field survey of both organically and conventionally grown lettuce grown in Spain found nearly 35% of samples collected tested positive for *E. coli,* with 20% exceeding the European Union guidelines of 100 CFU/g in RTE precut fruits and vegetables (16). Alternatively, a study conducted in the United States on 308 leafy green and herb samples found relatively low levels of *E. coli* from production through packaging for retail distribution. The levels found were close to 10 CFU/g, with total coliform levels more than 10 times higher (13). Another retail-level study on bagged spinach and lettuce mixes found only 16% to contain *E. coli,* but at levels less than 10 MPN/g (21). No *E. coli* was detected in the current study, but every sample tested had total coliforms. The levels of total coliform bacteria, however, did not correlate with high TBC, but were comparable (>210 MPN/g) to those found in a previous study (13).

In France, the standard for bagged salad is a maximum mesophilic count of $5 \times 10^5$ CFU/g (15). A previous U.S. study found 6 of 18 lots surveyed surpassed this maximum limit (10), and a survey done over a decade later noted that...
both spinach and lettuce mixes continued to exceed this limit (21). Our results showed that 4 of 95 bags tested had TBC greater than this limit, with another 11 bags approaching this maximum, with counts greater than \(5 \times 10^6\) CFU/g.

Care should be taken when making comparisons between these different studies, especially when relying on total bacterial plate count data. This study attempted to address the types of bias that can be introduced when data is ignored (e.g., zero counts, TNTC) or when laboratory errors or potential bacterial aggregation is not factored into data analysis. Using recently published work (1, 2), the FDA is currently revising the procedure used for handling plate count data (10). Primarily, the notion that only counts in the acceptable range of 25 to 250 should be used. While the overall outcomes of this study were not affected by this newer analysis, TBC counts for individual bags were adjusted, illustrating complexity when setting a regulatory maximum based on bacterial plate counts, namely, how to ensure best practices are being employed when raw counts are being combined and final concentrations calculated.

In summary, this study found samples taken from the bottom of bags to have higher TBC the majority of the time, and bags with the most number of days before the printed use-by date had generally lower TBC. Nonetheless, a statistically significant difference between top and bottom TBC was not found. TBC varied widely between lots and even more so within same-lot samples, as indicated by the sums of squares results. This finding, along with higher TBC in bottom samples, suggests further consideration when a microbiological sampling scheme of bagged produce is designed. However, the issue of best sampling practices when testing bagged produce remains largely unaddressed in the scientific literature and certainly warrants further examination.

ACKNOWLEDGMENT

We thank Lt. Kari Belin for her assistance in procuring the produce samples used in this study.

REFERENCES


